

**Reading Disability, Visual Stress, and Coloured Filters:  
A Randomised Controlled Trial**



**Examination Number 6267540**

**Psychology (Master of Science by Research)**

**The University of Edinburgh**

**August 2010**

## *DECLARATION*

I have read and understood The University of Edinburgh guidelines on Plagiarism and declare that this written dissertation is all my own work, except where I indicate otherwise by proper use of quotes and references.

## *ABSTRACT*

Coloured filters, in the form of plastic overlays or tinted spectacle lenses, are in widespread use to alleviate visual stress (also known as Irlen Syndrome), a disorder posited to be a major cause of reading disability (Irlen, 1991; Wilkins, 2003). However, a recent systematic review (Albon et al., 2008) concluded that the evidence for the efficacy of coloured filters is insufficient to recommend the treatment. The existence of visual stress as a diagnostic entity has also been questioned (Royal College of Ophthalmologists, 2009). This thesis first describes the various theoretical perspectives behind the use of coloured filters, and provides an in-depth review of the current evidence. A combined crossover study and randomised controlled trial of the coloured filters used by the Irlen Institute, the major proponent of the treatment, is then described. This experiment, which set out to avoid the methodological problems observed in previous trials - most importantly, double-blinding was employed - failed to find any evidence of visual stress, or for the statistically or clinically significant benefit of coloured overlays for reading rate or comprehension on two separate reading tests, in a sample of 61 Primary School-age children with reading problems. This was despite 77% of the sample having been diagnosed with visual stress by an Irlen diagnostician and prescribed coloured overlays. Visual stress theory and the current prevalence of the coloured filter treatment are discussed in the light of this new evidence.

# CONTENTS

<b>1. INTRODUCTION</b>	<b>5</b>
1.1 Reading disability	5
1.2 The history of colour as a treatment	7
1.3 The discovery of visual stress	8
1.4 Current methods of treatment	10
1.5 Summary	13
<b>2. THEORIES OF COLOUR AND READING</b>	<b>14</b>
2.1 Magnocellular deficit theory	14
2.2 Cortical hyperexcitability theory	21
2.3 Noise exclusion deficit theory	26
2.4 Summary	28
<b>3. EVIDENCE FOR COLOURED FILTER USE</b>	<b>30</b>
3.1 General comments	30
3.2 Non-randomised controlled trials	33
3.3 Randomised controlled trials	42
3.4 General discussion	49
3.5 Summary and conclusions	51
<b>4. THE PRESENT STUDY: A RANDOMISED CONTROLLED TRIAL</b>	<b>52</b>
4.1 Method	55
4.2 Results	60
<b>5. DISCUSSION</b>	<b>66</b>
5.1 The present study	66
5.2 The visual stress diagnosis	71
5.3 Directions for further research	72
5.4 Effective treatments for reading disability	76
<b>6. CONCLUSIONS</b>	<b>78</b>
ACKNOWLEDGEMENTS	79
REFERENCES	80

## **1. INTRODUCTION**

### **1.1 Reading disability**

Reading disability is subject to controversy in numerous ways. Definitions and diagnostic measures, aetiology and brain mechanisms, and treatments and interventions are all hotly debated in the psychological, physiological, and medical literature.

Perhaps due to the definitional issues, widely varying estimates of the prevalence of reading disability can be found in the literature. Figures of the percentage of school-age children affected, which show sex differences, range from as low as 3-6% (Hulme & Snowling, 2009) to as high as 22% in boys and 8% in girls (Rutter et al., 2004). It is clear, however, that even at the lowest estimate, reading disability affects a very substantial number of children. The disability persists into adolescence and adulthood (Hulme & Snowling, 2009).

Both the ICD-10 (World Health Organisation, 1992) and the DSM-IV-TR (American Psychiatric Association, 2000) use what has come to be known as the 'discrepancy' definition of reading disability - if an individual is found scores on a reading test below the level predicted by their age and IQ, they are diagnosed with reading disorder. Thus, the DSM-IV-TR characterizes reading disability, which it refers to as 'reading disorder', as 'reading achievement... that falls substantially below that expected given the individual's chronological age, measured intelligence, and age-appropriate education' (APA, 2000, p. 51).

However, recent studies have questioned the usefulness of measured intelligence in such diagnoses – IQ does not appear to predict responsiveness to teaching (e.g. Hatcher & Hulme, 1999). It has been suggested that IQ may explain 10-30% of variation in reading ability (Hulme & Snowling, 2009), so clearly it is not a full explanation. A discrepancy-based definition also has the potential to overlook specific reading problems in individuals with below-average IQ. Given that this is the case, many educational systems now offer remedial intervention for all children who show reading problems, irrespective of IQ (Albon et al., 2008).

Reading disability has been shown to be a heritable condition in studies from both quantitative and molecular genetic approaches (Pennington & Olson, 2005). Heritability estimates are variable; recent twin studies have estimated that genetic factors account for approximately 60% of the average group variance in reading deficits (Wadsworth et al., 2000; Hawke et al., 2006). Molecular studies have recently focused with some success on single nucleotide polymorphisms in genes such as *DYX1C1* (e.g. Bates et al., 2009), though it is highly likely that reading disability involves a complex interaction of several such genes (Parrachini et al., 2007). Interestingly, since

several of these candidate genes are implicated in normal, unimpaired reading, the findings provide evidence for the notion that dyslexia is possibly the lower tail of a normal distribution of reading ability (Bates et al., 2009; Shaywitz et al., 1992).

Reading disability is associated with a wide variety of adverse effects and life outcomes. Students with learning difficulties have been shown to have a higher incidence of stress, depression and suicide than average (Bender et al., 1999) and reading disability in particular has been associated with low self-esteem and social status (Humphrey, 2002; Nabuzoka & Smith, 1993; Riddick et al., 1999). In addition, several studies have linked reading problems at a young age to conduct disorder in adolescence, perhaps at least partially as a consequence of low self-esteem (Fergusson & Lynskey, 1997; Williams & McGee, 1994), and to offending behaviour (Kirk & Reid, 2001; though see Samuelsson et al., 2003).

These associations underline the importance of both gaining a full understanding of the aetiology of reading disorder, and thoroughly evaluating the treatment methods that have been suggested to ameliorate or alleviate the condition.

A great deal of evidence has accumulated which supports the 'phonological deficit' theory of dyslexia (Gabrieli, 2009; Hulme & Snowling, 2009; Ramus, 2001, 2003, 2004; Snowling, 1998; Vellutino et al., 2004; White et al., 2006). Individuals with reading disability are characterized as having poorly specified or fragile phonological representations, causing difficulty in several phonology-related cognitive processes essential for fluent reading, such as letter-sound decoding and verbal short-term memory (Vellutino & Fletcher, 2005). The neurobiological aetiology of this phonological deficit is not fully understood; some studies point to differences in the left temporoparietal cortex of disabled readers (Hoeft et al., 2006; Temple et al., 2001), but more work is necessary to fully clarify its nature. The phonological deficit, and its accompanying neural anomalies, are hypothesized to arise mostly from a genetic susceptibility, as was mentioned previously (Ramus, 2003).

Evidence for the phonological theory also comes from a different type of experiment: phonological ability in young children has been shown to be a good predictor of later reading ability in several longitudinal studies (Badian, 1994; Elbro et al., 1998; Hulme et al., 2002, Muter et al., 2004; see Castles & Coltheart, 2004 and Hulme et al., 2005 for detailed discussion). Both its explanatory ability, predictive power and wide evidence base make the phonological theory the current 'orthodoxy' in dyslexia research (Singleton & Trotter, 2005).

Yet not all researchers are convinced by the phonological deficit theory. Deficits in visual (Lovegrove et al., 1986), auditory (Tallal, 2003), and motor/cerebellar (Nicholson & Fawcett,

1999) systems have all been independently proposed as the underlying cause of the phonological problems observed in reading-disabled individuals. These theories, all of which are more controversial than the phonological theory (Ramus, 2003; White et al., 2006) have been unified under the general magnocellular theory of dyslexia (e.g. Stein, 2001), which will be discussed in detail below, especially with reference to its visual aspects.

Since this thesis considers reading problems which are characterized by some (e.g. Irlen, 2010, Wilkins, 2003) to be caused by a separate condition from developmental dyslexia, and since 'dyslexia' does not normally include reading comprehension difficulties (Hulme and Snowling, 2009), which will be discussed here, the term 'dyslexia' will not be used unless specifically referring to that disorder. Instead, the terms 'reading disability', 'reading disorder' and 'reading difficulty' will be used interchangeably henceforward.

The literature review which follows, and the randomized controlled trial which is described, investigate a controversial intervention for reading disability, based on a visual theory of reading disability: the use of coloured filters.

### *1.2 The history of colour as a treatment*

The use of light and colour for their apparent healing properties has a history stretching back to at least the ancient Greeks and Egyptians (Gottlieb & Wallace, 2001). Many authors have gone beyond the effects colour is popularly considered to have on mood (for which surprisingly little scientific evidence can be found - see e.g. Elliot et al., 2007 for an example), and made further claims. Several physicians in the latter part of the 19<sup>th</sup> Century, for instance, proposed many colour-based therapeutic interventions.

Collins (2002) recounts the story of A. J. Pleasonton, who, in the mid 19<sup>th</sup> Century, became famous for his attempts to cure many illnesses (and aid plant growth) using light shone through blue glass. Babbitt (1878) proposed that coloured light, dyes and lenses could be used not only to heal various disorders of the eye, but also a wide range of other medical conditions. Babbitt's pseudoscientific volume 'Principles of Light and Colour' advocated the use of colour as a panacea, a tradition that has been continued in various forms of 'complementary/alternative' and 'New Age' healing techniques; see Wauter (1999) for one example, involving the creation of homeopathic remedies using coloured light shone through water. Needless to say, more recent colour-based interventions are, *prima facie*, far more plausible than these historical treatments.

More focused interventions for learning disabilities began with Henning's 'chrome-orthoptics', (Henning, 1936) and involved the use of light therapy, along with lenses and prisms, to affect the

accommodative reflex of the eye. This was hypothesized to affect the autonomic nervous system, resulting in better concentration and improved learning. Although no scientific evidence was provided by Henning for the efficacy of this intervention or its mechanisms, the basic treatment methodology forms the basis for contemporary 'syntonics' (also referred to as 'light therapy' or 'photoretinology' – Howell & Stanley, 1988; College of Syntonic Optometry, 2009), a type of vision therapy named by Spitler (1941) which purports to 'balance' the visual system (Barrett, 2009).

Several studies have been undertaken to investigate the effect of coloured light shone into the eye on visual functioning, specifically the size of the visual field and how it relates to reading. For instance, Kaplan (1983) and Liberman (1986) hypothesized that reading disorder is caused by a reduced visual field, causing deficits in peripheral vision and thereby inefficient reading. Both studies reported visual field size increases and improved reading ability after syntonics treatment. The interventions in these studies involved participants looking through various coloured filters, either prescribed or individually selected by preference, at an incandescent light source. Unfortunately, neither study reported the details of reading tests administered to the participants, nor did they control for differing pre-treatment baselines in visual field size. These methodological issues, among others (Howell & Stanley, 1988) render their contribution to the evidence base for syntonics exceedingly weak.

### ***1.3 The discovery of visual stress***

While previous colour-based interventions focus on remediating various symptoms that are causing learning inefficiency in standalone treatment sessions, the majority of the recent literature focuses on interventions *in situ*, those which are used to treat symptoms at the same time as learning or reading is taking place. In a volume on dyslexia, Critchley (1964) noted that some children who had difficulty reading on white paper had their reading facilitated by coloured paper, but nothing further along these lines was reported until Meares (1980) described several children in a reading clinic who were reporting symptoms of discomfort when reading. Meares hypothetically linked these symptoms to contrast sensitivity.

Subsequently, Irlen (1983) posited that some individuals with reading difficulty suffer from a distinct syndrome, which was dubbed 'Scotopic Sensitivity Syndrome'. This choice of name is curious; some have questioned the appropriateness of the term 'scotopic', since it relates to conditions of low light which would involve the function of rod cells in the eye (Helveston, 1990). There are no rod cells in the fovea, where printed words are projected, so the extent to which the term relates to reading is unclear (Menacker et al., 1993). Nevertheless, several other names for

the proposed disorder have been used, including 'Meares-Irlen Syndrome' (occasionally 'Irlen-Meares' or 'Irlen Syndrome'), 'visual discomfort' and 'visual stress'.

The symptoms of the disorder (hereafter referred to only as 'visual stress' – the most neutral of the terms used) are described as various distortions and illusions of the text experienced to varying degrees by each individual. Irlen (1991) describes the five components of visual stress: light sensitivity (including difficulties reading under fluorescent lights), inadequate background accommodation (trouble reading in conditions of high contrast), poor print resolution (including illusory movements of text on the page), restricted span of recognition (also known as 'tunnel reading', a restriction on reading groups of words) and lack of sustained attention (p. 31). No indication is given as to the prevalence of each symptom within the population of individuals with visual stress.

The basis on which Irlen classified the symptoms as part of the visual stress disorder is not known. Information as to which of Irlen's patients were screened for other, possibly co-occurring, conditions such as attention deficit disorder or optometric, orthoptic, or ophthalmic abnormalities would be of interest in assessing the construct validity of visual stress (it has been noted by Scheiman, 2004, that many of the symptoms discussed by Irlen are remarkably similar to those reported by individuals with vision problems which are better understood by eye care professionals, such as accommodative anomalies or for a discussion of visual problems and their possible relation to visual stress). It should be noted that Irlen is of the opinion that dyslexia and visual stress are separate, but sometimes co-occurring, disorders (Irlen, 1991; 2010), both of which can contribute to reading disability.

Reading disability is by no means the only problem discussed by Irlen. Visual stress is also claimed to affect writing, depth perception, coordination, and motivation. Further, in a recent book, Irlen (2010) claims that conditions including, but not limited to ADHD, autism, chronic fatigue syndrome, epilepsy, Tourette syndrome, head injuries, agoraphobia, anxiety attacks, depression, and conduct disorder are related to visual stress. These disorders are, apparently, either worsened by the presence of visual stress, are in part caused by the self-esteem issues associated with being unable to read or pay attention in a school classroom, or are in actual fact simply visual stress which has been misdiagnosed as another disorder. No references to studies which confirm any link between visual stress and these other diagnoses are provided.

As we will see below, researchers such as Wilkins (e.g. 2003) have hypothetically linked various other disorders to visual stress, but unlike Irlen they have tested these hypotheses and published their findings. Already noted above is the evidence linking reading disorders to adverse long-

term outcomes such as offending behaviour, but the connections between these findings and visual stress *per se* have never been investigated.

Irlen (1983, 1991, 2010) does not discuss the theoretical basis for visual stress in detail, nor does she make specific, testable claims about its aetiology. Visual stress is said to 'possibly [involve] a structural brain deficit involving the central nervous system' (Irlen, 1991, p. 57), resulting in full-spectrum light inhibiting perception. Irlen (2010) provides pictures from a SPECT scan of a visual stress sufferer which appear to show reduced activation in several brain areas when the individual is wearing Irlen coloured lenses (p. 101). However, no scientific details are provided about the nature of these scans, and it does not appear that they have been published in a scientific journal.

#### ***1.4 Current methods of treatment***

Irlen (1991) recounts her discovery of the effectiveness of colour on reading, noticing that, when a coloured, transparent plastic overlay was placed over the text, disabled readers tended to report fewer visual stress symptoms and had improved reading performance. Further to this, Irlen makes several intriguing claims, indicating that not only does each person have an optimal filter colour which is unique to them, but that only the particular set of colours provided by the Irlen Institute are effective in treating the disorder. This latter claim is due to the fact that 'the Irlen Method has been developed, refined, and researched over many years' (Irlen, 1991, p. 194). No specifics are given, and indeed, it could conversely be argued that other types of colour-based intervention (e.g. Wilkins' Intuitive System - see below) have been researched far more extensively and scientifically than the Irlen Institute's method. Nevertheless, it is claimed that 'inaccurate colour selection [i.e. that of non-Irlen Institute practitioners] can result in headaches, eye strain, and fragmented brain processing resulting in more distortions and reading problems' (Irlen Institute, 2010).

It is important to note that Irlen did not perform or publish any controlled clinical trials or provide any evidence in the initial descriptions of the syndrome and its treatment. Indeed, Irlen's book (Irlen, 1991) is almost entirely based on anecdotes from the author's experience as an educational psychologist (Helveston, 2001), and makes a wide range of claims about clinical symptoms, syndrome prevalence and treatment efficacy with little empirical support. For instance, Irlen claims that visual stress is present in 12% of the normal population, and 65% of individuals diagnosed with dyslexia (Irlen, 1991). It is remarkable that in her most recent volume, Irlen (2010) provides almost no more supporting evidence for these claims (or indeed any other claim) than in the earlier works (Irlen, 1983, 1991). As we will see, a substantial number of studies

have been published in the intervening time. It is striking that Irlen (2010) hardly makes mention of them.

The Irlen Method (along with other, similar treatments) has received a great deal of mass media publicity (Helveston, 2001), and the Irlen Institute now has clinics in many US states, several locations in the UK, and many other parts of the world (Irlen Institute, 2009). Clearly this high level of publicity warrants further scientific investigation into the efficacy of the treatment.

In spite of its unreliable nature, Irlen's work and the publicity surrounding it can be seen as an inspiration for many subsequent studies. For example, Wilkins (e.g. 2002, 2003) and his research team have been major contributors to the literature, with many published studies and reviews. Wilkins (1994) has, on the basis of this research, developed for purchase by disabled readers the 'Intuitive Overlays', an alternative set of coloured overlays. In addition, Wilkins is the inventor of the 'Intuitive Colorimeter' (Wilkins, 2003), an instrument used in both research and clinical settings to choose a more precise optimal colour for tinting lenses. The Intuitive Overlays, the Intuitive Colorimeter, and a reading measure designed by Wilkins, the Wilkins Rate of Reading Test, are important parts of Wilkins' research methodology (see e.g. Wilkins & Evans, 2009).

As mentioned above, Wilkins and colleagues (e.g. Wilkins & Evans, 2009) have made the case that some other conditions, such as migraine and multiple sclerosis, have some neurological features in common with visual stress. This often involves issuing Intuitive Overlays to sufferers of these disorders. These arguments will be examined in detail below.

#### *1.4.1 Coloured Overlays*

The overlays used by the Irlen method and the Intuitive system are A4 plastic sheets, with one 'shiny' and one 'matte' side. Wilkins, on the basis of his group's research, concurs with Irlen's opinion that filters should be individually-prescribed, as there exists an optimal colour for each individual which will reduce their visual stress symptoms and facilitate their reading to the greatest extent (Wilkins, 2003; Wilkins et al. 2005; though see Simmers et al., 2001). The basic set of colours provided by the Irlen method and the Intuitive system are broadly similar (the chromaticity of the Intuitive Overlays, along with other details, can be found in Wilkins, 1994). In both systems, a patient's optimal overlay is chosen using a process of elimination, often combined with questions relating to any visual stress symptoms the patient may be suffering from. The diagnostic procedure of the Irlen Institute is described in detail in Section 4.1.

#### 1.4.2 Coloured Spectacle Lenses

Coloured lenses usually involve more in-depth examination. In the Intuitive system, the aforementioned Intuitive Colorimeter allows a wider variety of colours to be chosen for lenses, with over 25,000 different colours having been prescribed, according to Wilkins et al. (2007). Using the colorimeter, the patient can themselves manipulate colour, hue and saturation independent of one another, and thereby find their exact preferred colour (Wilkins et al., 1992; Wilkins, 2003). A set of test lenses with a limited range of colours exists to use during diagnosis with the Colorimeter, and after diagnosis the practitioner can send away for the precise colour lenses required (University of Essex, 2010). There exists a 'Society for Coloured Lens Providers', which provides a code of conduct which practitioners must sign as well as instructional lessons and courses (Society for Coloured Lens Providers, 2010). On the other hand, diagnosticians and screeners from the Irlen Institute – who do not subscribe to the Society for Coloured Lens Providers – only require further diagnostic appointments before tinted lenses are prescribed, but no instruments similar to the Intuitive Colorimeter are used by their method.

Wilkins (2003) has criticised the Irlen method due to the fact two or more lenses are often used simultaneously; this may lead to complementary colours being used, which 'may counteract each other', at which point there is 'no guarantee that the colour formed by combining the trial lenses will be better than either of the lenses on their own' (p. 90). The Irlen Institute's method of tinting lenses is confidential, and no technical data have been made available for scientific scrutiny.

Both Irlen (1991) and Wilkins (2003) note that the chosen 'optimal' colour of tinted spectacle lenses is often different from that of overlays. Wilkins (2003) offers a possible explanation in terms of 'the state of adaptation of the eyes diff[ering] for overlays and lenses. Lenses have an effect similar to that of a coloured light: everything within the visual field is coloured, and the eyes adapt to that colour... Overlays, on the other hand... are viewed while the eyes are adapted to coloured light' (p. 92). The Irlen Institute gives a similar explanation (Irlen Institute, 2010).

Finally, in publicity on BBC 1 Television (*The One Show*, 19 January 2010) and BBC Radio 4 (*All in the Mind*, 26 May 2010), Stein has claimed to be able to treat up to 50% of individuals with dyslexia with yellow and blue lenses. The theory behind these claims is discussed in Section 2.1 below, but no details of, firstly, how many children are being treated using this method or, secondly, the precise chromaticity or nature of the lenses being used were provided.

### *1.5 Summary*

Currently, the cause of reading disabilities is mainly characterized as a cognitive phonological deficit. However, several theorists have suggested, controversially, that visual disorders may also be part of the explanation. The visual stress theory, pioneered by Meares (1980) and Irlen (1983), is one of these theories. Along with the theory comes the intervention – the use of coloured filters to alleviate visual stress symptoms. Like many treatments and interventions, both effective and otherwise, the use of colour as a treatment for various disorders is positively ancient, and has come in a wide variety of forms.

Although originally proposed in anecdotal and somewhat overstated fashion, more reliable researchers have since investigated the visual stress theory in far more depth. The following parts of this thesis will discuss some posited theoretical bases for the treatment method's effects, and will assess the currently available literature on its efficacy.

## *2. THEORIES OF COLOUR AND READING*

Researchers have put forward various theoretical mechanisms which potentially explain the effects of coloured filters on reading. This section discusses three such explanations: magnocellular deficit theory, cortical hyperexcitability theory, and noise exclusion deficit theory.

### *2.1 Magnocellular Deficit Theory*

Many authors have proposed magnocellular deficits as an explanation of reading disorder (Stein, 2001). This review will focus on the visual aspects of this magnocellular impairment, though the theory has been posited to explain the auditory and motor as well as visual deficits which have been associated with dyslexia by some authors (Ramus, 2003, Figure 2).

Visual magnocellular theories often relate to the three types of cell found in the lateral geniculate nucleus (LGN), part of the brain's visual pathway leading to the primary visual cortex. The upper four layers of cells, known as the 'parvocellular' neurons owing to their small cell bodies, are involved in colour vision and processing fine spatial detail, along with smaller, 'koniocellular' neurons which are found throughout the LGN (e.g. Chatterjee & Callaway, 2003). Theories are focused, however, on the lower two layers of larger cells with a higher conduction velocity, the 'magnocellular' neurons, which respond to low spatial resolutions and high temporal frequencies - for example those caused by movement (Palmer, 1999). Cells can be found in the retina which selectively project to parvocellular or magnocellular neurons (Shapley & Perry, 1986), underlining the importance of these two distinct types of cell for visual processing.

Post-mortem studies which showed evidence of abnormalities and ectopias in the brains of dyslexic, as compared to non-dyslexic, individuals (Galaburda et al., 1985; Humphreys et al., 1990; Livingstone et al., 1991; see Eckert, 2004, for a review) are routinely used as evidence for the magnocellular view. The theory proposes that these abnormalities, which can be caused by interrupted neuronal migration, disrupt the magnocellular system.

This disruption, claim magnocellular theorists, cause the magnocellular system to fail to suppress the action of the parvocellular system. In non-disordered individuals, this suppression 'causes the activity in the parvocellular system to terminate so as to prevent activity elicited during one fixation from lingering into that from the next fixation' (Skottun, 2000a, p. 111). However, without this suppression, input from one visual fixation would 'bleed' into the next, creating confused images and, by extension, disordered reading (Livingstone et al., 1991; Lovegrove et al., 1986). Subsequent research has, however, shown that it is the magnocellular system, not the

parvocellular system, which is suppressed at each saccade (Skottun & Parke, 1998), and as a result this hypothesis became untenable.

Newer formulations of the magnocellular deficit theory have been proposed (e.g. Stein & Walsh, 1997; Stein, 2001), positing that defective motion sensitivity in the magnocellular system leads to binocular instability – unsteady fixations when using both eyes simultaneously - and, therefore, to ‘the letters [dyslexics] are trying to read to appear to move around and cross over each other’ (Stein, 2001, p.12). This phenomenon is appreciably similar to one of the proposed symptoms of visual stress, described above.

It is worth noting at this point that the proponents of a magnocellular theory-related mechanism for the benefits (if any) of coloured filters make no reference to visual stress, or to any other terms used to refer to the disorder. Nor, puzzlingly, do they make significant reference to any of the other studies into the effects of colour on reading, such as those by Wilkins and colleagues (e.g. Bouldoukian et al., 2002; Wilkins et al., 1994). Presumably these authors are of the opinion that the symptoms of ‘visual stress’ are in fact the result of a faulty magnocellular system, and that visual stress either does not exist or is not a separate condition from magnocellular disorder. As we will see, some experiments into coloured filters from a perspective outside these theoretical ‘camps’ do make reference to the magnocellular theory (e.g. Christenson et al., 2001; Noble et al., 2004).

Skottun (2000a), in a review of the literature, found that the number of studies providing evidence in support of the magnocellular deficit theory are outnumbered by the number of studies finding evidence incompatible with the theory. Other reviewers note that in several demographic studies, sensory deficits are found only in a small proportion of disordered readers, whereas the vast proportion of these readers have a cognitive phonological deficit (e.g. Ramus, 2003; White et al., 2006; see Section 1.1). Magnocellular deficits are also sometimes found in individuals with no reading problems (Vellutino et al., 2004). This evidence reduces the likelihood that the magnocellular theory is an explanation of a high proportion of reading disorders. In addition, Ramus (2004) has proposed an interpretation of the evidence which explains the neuroanatomical abnormalities discussed above (e.g. Galaburda et al., 1985) in terms of the phonological deficit theory.

Skottun (2000a, 2004) and Skottun and Skoyles (2007b) have repeatedly pointed out that it is extremely difficult to isolate magnocellular function in any experimental setting. Many experiments risk stimulating both magnocellular and parvocellular systems (for further discussion, see Stein et al. (2000) and Skottun (2000b)). Further, it is unclear whether some experimental techniques, such as those which use motion sensitivity (see e.g. Talcott et al, 2000),

are involving an entirely separate (extrastriate) brain pathway instead of either the magnocellular or parvocellular systems (Skottun, 2000a; Skottun & Skoyles, 2007b). In addition, some experiments have failed to find any evidence of the expected abnormalities on magnocellular-based tasks in dyslexic individuals (e.g. Johannes et al., 1996).

Related to this, many of the studies which purport to provide evidence for the magnocellular theory are based on spatial and temporal frequency measures – the most frequently-utilized way of differentiating between the magnocellular and parvocellular pathways. One problem with such tests has been demonstrated by Heath et al. (2006). These authors found that while individuals do perform reliably on visual and auditory versions of these tests, (i.e. test-retest reliability tends to be high), there do not seem to be reliable correlations within individuals on different tasks (i.e. construct validity is low). Studies which attempt to provide evidence for the magnocellular view, then, do not necessarily assess magnocellular function in its entirety.

Even so, as mentioned above, the hypothesized deficit in the magnocellular system has been proposed by some researchers (e.g. Christenson et al., 2001; Noble et al. 2004; Whiteley & Smith, 2001) as a potential explanation of the effects of coloured filters on reading, perhaps by facilitating the performance of the poorly-functioning magnocellular pathway, thus aiding reading<sup>1</sup>. The magnocellular system's involvement with colour may at first seem counter-intuitive, for it is the parvocellular system which is involved in colour vision, as previously stated. Nevertheless, three studies have investigated the effect of colour on reading function in this context – the first indirectly, and the second and third directly.

First, Stein et al. (2000b) used yellow filters as controls in an experiment on monocular occlusion as a treatment for dyslexia. The hypothesis in this experiment, which was based on previous work (Cornelissen et al., 1992; Stein & Fowler, 1985), was that, if a substantial proportion of dyslexics suffer from binocular instability, a method of treating this may be to occlude one eye for some length of time in order for the dyslexic to gain a more fixed frame of reference.

To test this, Stein et al. (2000b) sampled 143 dyslexic children's reading age along with orthoptic measures, including stability of fixation. 71 of these children were then given spectacles tinted with a yellow colour (the control group), while a further 72 children were given spectacles with a yellow right lens and the left lens occluded with opaque tape (the treatment group). Curiously, a control group with colourless, non-occluded lenses in their spectacles was not included in the experiment. This may be due to the fact that this study was an attempted replication of previous

---

<sup>1</sup> It has been suggested by Stein that this is the mechanism for the effects of yellow coloured filters. However, blue coloured filters, which Stein also claims facilitate reading, are hypothesized to 'trick' the brain's internal clock functions into behaving as if it were early morning – the light of which is blue - leading to increased attention and thereby improved reading. There is, as yet, no published evidence for this hypothesis, though it is forthcoming (Stein, personal communication, June 2010).

studies which used colourless lenses, but it would have been nonetheless instructive to have such participants in this study.

Retesting of fixation stability and reading age occurred at 3, 6 and 9 months after the spectacles were issued. The treatment group were found to develop stable fixations more quickly across this time than the control group, and reading age progressed much faster in the treatment group. The control group's progression was attributed to the yellow lenses 'boost[ing]' (p. 168) the magnocellular system.

Previous work on the monocular occlusion hypothesis was criticized for basing conclusions on flawed methodology (Bishop, 1989; though see a response from Stein, 1989). The newer study has drawn criticism also; Fawcett (2000) notes that, while advances in binocular stability did occur faster in the treatment group, by the end of the 9 months, only 6 more children in the treatment group had achieved stable fixation than the controls, which was not a significant difference. This draws the efficacy of monocular occlusion for achieving stable fixations into question; it is unclear whether the fixations could be aided by the yellow lenses, or simply appear as a matter of course during the maturation process, which is known to be the case in normal development (Bishop, 1989).

In addition, Fawcett (2000) suggests that more specific details about the reading performance of the children are required – if it were found that the children who achieved stable fixations are making less reading errors of a visual nature (e.g. skipping lines or words) as opposed to a linguistic nature (e.g. incorrect sounds), it would be more convincing evidence for Stein et al. (2000b)'s hypothesis.

Stein et al. (2000b) originally took their sample of 143 from a group of 700 dyslexics, after they found that 80% of that larger group did not have unstable fixations. This, combined with the fact that a proportion of the remaining 20% will probably gain stable binocular fixations spontaneously (p. 168), makes it clear that the monocular occlusion method will not be applicable to the vast majority of those with reading disorders. Perhaps for this reason, Stein and colleagues have performed no more research into monocular occlusion, preferring to further investigate the effects of the coloured filters they previously used as a control.

The second study to be considered, Chase et al. (2003), used a variety of measures involving red and blue light to study the magnocellular pathway's contribution to reading. Previous research using filters had shown that the colour blue may facilitate reading (e.g. Iovino et al., 1998, though see Christenson et al., 2001) and suggested that the magnocellular pathway's function is inhibited by red light (e.g. Hughes et al., 1996) - though see Pammer & Lovegrove, 2001. These latter authors, on the basis of four separate experiments investigating red and blue light on

magnocellular function, are sceptical that separate colours can differentially activate the magnocellular system. They also suggest that differences in stimulus contrast, as opposed to colour, may be an explanation of the various inconsistent results from studies in this area.

Chase et al. (2003) attempted to differentiate between two possible mechanisms for blue light aiding reading: that the facilitation is due to either the presence of short wavelengths (blue light) or the absence of long wavelengths (red light). Their research involved the use of Irlen filters (Irlen, 1991) and isoluminant stimuli (those which are created specifically to stimulate only the colour-sensitive parvocellular pathway – see Lu et al., 1999). They concluded that reading is facilitated by the removal of long wavelengths, and that a grey filter would do just as well as a blue filter to remove these. Additionally, they showed that the presence of long wavelengths appears to interfere with reading, and, using the isoluminant stimuli, that the parvocellular pathway is less inhibited by the presence of red light than the magnocellular pathway.

This research has been shown to have several flaws, however. Skottun (2004) has demonstrated that not only is it incorrect to say that red light has an effect on the magnocellular system, also it is not necessarily possible, as Chase et al (2003) assumed, to isolate magnocellular function by using red light – indeed, ‘red light may have a profound effect on parvocellular neurons’ (Skottun, 2004, p. 67). In addition, Jordan et al. (2007) have shown that creating true isoluminant stimuli is no simple matter, and that the stimuli created by Chase et al. (2003) – where the letters used were the same luminance as the background - may not have been processed in the manner the authors had assumed, confounding their results. A similar set of experiments would have to be performed bearing these caveats in mind before stronger conclusions could be drawn.

The next study to be discussed is by Stein’s research group. Ray et al. (2005) hypothesized that, since blue light had been shown to inhibit magnocellular function through a mechanism involving delayed retinal S-cone cell stimulation (Stockman et al., 1991), the magnocellular pathway may be more sensitive to yellow light, even though it does not mediate colour vision. To investigate the hypothesis, the researchers had a group of children with reading disorder read through a yellow filter for 3 months, to release the magnocellular pathway ‘from the negative influence of the S-cones’ (p. 2), and compared them with a similar group of children receiving a placebo treatment. After this time, the researchers found that motion sensitivity, convergence, accommodation, and reading ability (the latter measured by the child’s British Ability Scales (BAS) reading score (Elliott et al., 1978)) had improved significantly more in the group using the yellow filters. See the next section for a more detailed discussion of these findings.

The theoretical explanations for this finding (involving reduced S-cone input), Ray et al. (2005) admit, are controversial. Indeed, Skottun & Skoyles (2007a) provide evidence that the inputs of S,

M, and L-cone cells are summed *before* communication to the magnocellular system, making it unlikely that the influence of one type of cone in particular can be ‘filtered out’. They also note that, when discussing mechanisms, Ray et al. (2005) seem to be proposing ‘two contradictory views with regard to the effect of yellow stimuli on the magnocellular system’ (Skottun & Skoyles, 2007a, p. 291*n*) – at one point suggesting that magnocells are inhibited by too much input from the S-cone cells in the eye, and at another suggesting the problem is due to an incorrect L/M-cone balance. This renders the theory somewhat confused. See Section 3.3 for further discussion of the partial aspects of Ray et al. (2005)’s study.

Due to these criticisms, and especially since no further experimental work has been done from this theoretical perspective, it is difficult to conclude that the magnocellular deficit theory has a great deal to say about the effects of colour on any visual process.

More generally, as noted by Skottun (2005), the issues of, firstly, the magnocellular pathway’s potential involvement in the reading process, and secondly, the magnocellular deficits potentially being a cause of reading disorder, are separate and should not be confused. It would be wholly unsurprising if the magnocellular pathway was involved in reading, along with the many other visual systems. More precise evidence would be required to show that a deficit in the magnocellular system alone was responsible for some reading problems.

It has been suggested that sensory and sensorimotor deficits may be a correlate, but not necessarily a cause, of dyslexia (Ramus, 2003; Vellutino et al., 2004; White et al., 2006), but it is unclear that these are due to magnocellular impairments in particular (Amitay et al., 2002). Indeed, Skoyles & Skottun (2004) note that ‘not only are there cases of dyslexia that are not attributable to magnocellular deficits, but also there are a substantial number of instances of magnocellular deficits that do not lead to dyslexia’ (p. 81). The hypothesis of magnocellular causality is therefore not assured. Finally, Castelo-Branco et al. (2007) found that magnocellular impairments are present in the developmental disorder Williams Syndrome, but they do not reliably predict any behavioural or perceptual outcome. This clearly challenges the magnocellular theory, which predicts a causal pathway between low-level visual impairments and high-level deficits.

From an entirely different perspective, Wilkins (2003) marshals four main lines of evidence against the magnocellular theory of coloured filters’ facilitation of reading. Firstly, experiments have found that individuals with no reading disability can also benefit from coloured overlays and lenses (e.g. Jeanes et al., 1997). However, we have already seen above that magnocellular deficits are sometimes found without reading disorder (Skoyles & Skottun, 2004; Vellutino et al., 2004), blunting the effectiveness of this line of criticism. Secondly, it is noted that the

magnocellular theory is a theory about dyslexia specifically, and since dyslexia is not necessarily associated with visual stress (Wilkins, 2003; though see Singleton & Trotter, 2005), we cannot necessarily expect the theory to explain visual stress. Thirdly, Wilkins (2003) argues that the magnocellular theory cannot explain the colour specificity found in several studies of colour and reading (e.g. Jeanes et al., 1997; Wilkins et al., 1994; though see Simmers et al., 2001).

Finally, Wilkins (2003) notes that Simmers et al. (2001) examined a group of children who regularly used (and, apparently, benefitted from) coloured lenses, and failed to find the magnocellular deficits which would be expected on the magnocellular theory (see also White et al., 2006). The authors suggest that this evidence shows that visual stress is separate from other disorders, such as dyslexia and attention deficits, which often contribute to poor reading. This is in line with the opinion of Irlen (1991).

## *2.2 Cortical Hyperexcitability Theory*

Since Wilkins (e.g. 2003) views the magnocellular theory as insufficient to account for the effects of colour on reading, he and his colleagues have tentatively proposed an alternative theory based on cortical hyperexcitability – certain areas of the cortex over-responding to particular visual stimuli. This theory draws comparisons between individuals with visual stress and those with several other potentially related disorders: migraine, epilepsy, multiple sclerosis and autism. This review will first look into how these connections have been made, and will then describe and discuss the neurophysiological basis of the theory.

### *2.2.1 Connections to other disorders*

Noting that one of the symptoms often described by visual stress sufferers is headache, Maclachlan et al. (1993) surveyed a sample of 74 children with reading disorders. Their results indicated that individuals who benefit from coloured overlays – i.e. those who, by the Wilkins Intuitive method, have visual stress - were almost twice as likely (60% to 31%) to have a family history of migraine than those who do not report any benefits from colour. Wilkins' theory of visual stress (e.g. Wilkins, 1995; Wilkins et al. 2004; Wilkins et al., 2007) uses this finding as a basis for making other comparisons between visual stress and migraine.

Wilkins' theory of visual stress (Wilkins, 1995) hangs on the assumption that migraine (and therefore, hypothetically, visual stress) is caused by cortical hyperexcitability. However, there is considerable controversy amongst researchers investigating migraine as to whether this is indeed the case. Studies which find cortical hyperexcitability in migraine (e.g. Mulleners, 2001; see Aurora & Wilkinson, 2007, for a review) are to be contrasted with those which come to the opposite conclusion – findings of cortical hypoexcitability (reviewed by Ambrosini et al., 2003).

Coppola et al. (2007) suggest the controversy is due to a semantic confusion, and propose a new term – 'cortical hyperresponsiveness' - which they suggest accounts for all the supposedly conflicting results. According to Coppola et al., (2007), the evidence suggests that different kinds of stimulation cause the migranous cortex to respond in different ways –repetitive stimulation will cause hyperexcitability, but for low numbers of stimuli, cortical excitability will be reduced (i.e. hypoexcitability). For this reason, the authors view 'hyperexcitability' as too limiting a term to describe the true nature of migraine. Stankewitz and May (2009), on the other hand, suggest that the contradictory results arise from methodological issues. For instance, very few studies control for the circadian nature of migraine – the authors suggest several endocrinal influences may confound the estimates of cortical excitability. For these reasons, the assumption that

migraine is simply caused by cortical hyperexcitability is not an entirely firm basis for building a theory.

Nevertheless, the use of coloured filters for individuals with migraine has been studied, and positive results have been found. Evans et al. (2002), in a sample of adults with migraine, found that pattern glare (defined as a hypersensitivity to certain visual patterns - see Chronicle & Wilkins, 1996; Wilkins and Nimmo-Smith, 1987) was the only reliable visual correlate of migraine from candidates of contrast reduction sensitivity, pattern glare, and a range of orthoptic and optometric disorders. In a follow-up randomized controlled trial using the same sample, Wilkins et al. (2002) found that coloured lenses of an optimal colour, as selected on the basis of assessment with the Wilkins Intuitive Colorimeter (Wilkins et al., 1992), caused a marginally significant reduction in headache. However, these studies are subject to sampling bias, since the criterion for selection of participants was that the individual reported some benefit from coloured overlays (none of the sample had previously used coloured lenses). For this reason, the results of the two studies should be regarded as provisional. Indeed, Evans et al. (2002) admit that 'there is still a need for a rigorous controlled trial to investigate the optometric correlates of migraine in a large, unselected, sample' (p. 140).

Other disorders have been similarly linked to visual stress. Wilkins et al. (1999) performed a longitudinal study into the use of coloured lenses in a population with photosensitive epilepsy. Having selected an optimal colour of lens, 17 patients were followed up an average of 2.4 years later. The authors found that 13 of the patients still used their coloured lenses, and they had noticed important improvements when wearing them, such as reduced dizziness. Three of the participants reported a reduction in number of seizures. Wilkins et al. (1999) argue that these effects cannot be attributed to the placebo effect, since the duration of continued lens use is very long. This is a common argument in this area (see below), but it is not clear that it is valid. Placebo effects are not simply tied to novelty; there is no firm reason for thinking a placebo effect should necessarily be short-lived. Wilkins & Evans (2009) suggest that, since epilepsy is associated with seizures, in this case the lenses can be shown to help another condition involving cortical hyperexcitability.

Another possible connection to visual stress comes from a study of lens use in multiple sclerosis (MS) by Newman Wright et al. (2007). Since estimates of the prevalence of visual deficits in MS range from 38-85% (Arnold, 2005), and a small proportion of MS sufferers experience seizures (Koch et al. (2008) give an average of around 3% from numerous studies), Newman Wright et al (2007) hypothesized that coloured overlays may elicit similar positive responses to those Wilkins et al. (1999) found with lenses in epilepsy.

The authors did indeed show that performance on a the Wilkins Rate of Reading test (WRRT; Wilkins et al., 1996) and a new visual search task, the 'Circles Search Task' was superior when participants used overlays of their chosen, 'optimal' colour. All but one of the 26 participants also reported a reduction in subjective visual stress symptoms. Newman Wright et al. (2007) note with surprise that, after several weeks the participants' performance on reading and visual search tasks *without* an overlay had also increased. They rule out practice as a sufficient mechanism for this result since performance with an overlay was invariably superior, but unfortunately do not consider a possible explanation based on a *combination* of practice and placebo effects.

The final disorder linked to visual stress by Wilkins and colleagues is autism. It has been observed that a third of individuals with autism develop seizures in early life (Gillberg, 1991), and various disorders of visual perception have been found in children with autism (Dakin & Frith (2005) for a review of the literature). With this in mind, Ludlow et al. (2006) used the WRRT to assess the effects of coloured overlays on the reading ability of children with autism (presumably high-functioning autism; no further definition is given) compared with that of a sample of neurotypical children and those with moderate learning disability. It is to be noted that the autistic children participating in this experiment were not poor readers. The researchers found that reading rate with overlays increased only for the autistic children (indeed, the overlays appeared to inhibit the reading of the control group).

The conclusions that can be drawn from this study are limited, however. Since the autistic children in the sample were not those who suffered from seizures, the connection to cortical hyperexcitability was not necessarily present in the sample. Since the phenotype of autism is extremely heterogeneous (see, e.g. Folstein & Rosen-Sheidley, 2001; Frith, 2006), there is some uncertainty as to whether this study assessed the correct subtype of autism for any concrete conclusions about links to visual stress to be drawn. In addition, Ludlow et al. (2006) acknowledge that their study did not sufficiently control for novelty effects, which may have influenced the results.

It can be seen, then, that the connections drawn between visual stress and migraine, epilepsy, MS and autism are at present speculative and based on a small number of studies. More evidence will need to be gathered to make more solid associations between the disorders.

### 2.2.2 Proposed mechanism

Having provided these tentative lines of evidence which may show that visual stress involves cortical hyperexcitability, the theory posits that visual stimuli involving high spatial frequency and high contrast will cause an abnormal reaction in individuals with a particular kind of

hyperexcitability. Lines of text have been compared to high-contrast visual gratings (e.g. Wilkins & Evans, 2009). Since this kind of visual stimulus has been shown to cause hyperneuronal activity leading to seizures in those with photosensitive epilepsy (Wilkins et al., 1980) and, more pertinently, perceptual distortions in a study of individuals with migraine (Huang et al., 2003), it is suggested that a similar phenomenon can be found in individuals suffering from visual stress.

It is therefore predicted that the main marker of visual stress will be found to be pattern glare, and this has been found to be the case (e.g. Hollis & Allen, 2006). Pattern glare is defined by Evans & Stevenson (2008) as being caused by stimuli ‘with spatial frequency of about 3 cycles per degree (cpd), even width and spacing (duty cycle 50%), high contrast and be viewed binocularly’ (p. 296 – see also Wilkins et al., 1984). Evans & Stevenson (2008) went on to define clinical thresholds for the Evans and Wilkins Pattern Glare Test (see Wilkins & Evans, 2001), and suggested that anyone scoring above a certain threshold on the test may be suffering from visual stress in everyday life.

Some researchers have attempted to tease apart two of the aspects of pattern glare mentioned above. Williams & LeCluyse (1990) had shown that blurring text (reducing both the spatial frequency and the contrast level) facilitated reading in disabled readers, and as a result Williams et al. (1995) used a search task in an effort to differentiate between the effects of spatial frequency and contrast reduction. Groups of individuals with reading disorder, those with reading disorder and attention deficit disorder, and a control group of normal readers were asked to search for a particular letter in arrays which had been altered visually in various ways. Intriguingly, the group with both ADD and reading disorder had their reading facilitated by both spatial frequency and contrast reduction, but in the reading disorder-only group, only contrast reduction appeared to aid reading – the poor readers had their search time reduced to almost the level of the controls with reduced contrast, but no significant effect was found for changes in spatial frequency. The reasons for the differences between the two groups of poor readers are uncertain, but Williams et al. (1995) suggested that further work should investigate why contrast reduction aided reading rather than spatial frequency.

If, as in the Williams et al. (1995) study, contrast reduction is the only mechanism for reading improvement, it could potentially explain any benefits found by users of coloured filters. However, this would not explain the findings of several studies which appear to show colour specificity, that is, an ‘optimal’ colour for each individual with which they report benefits that they do not report for any other colour. On the basis of seminal work by Zeki (1983) on neurons in the visual system of monkeys which respond to particular colours and the finding by Xiao et al. (2003) that the cortex of monkeys is spatially arranged into areas for separate colours, Wilkins & Evans (2009) hypothesize that looking through colour can ‘redistribute excitation’ in the cortex

(p. 5). This leads to the reduction of visual stress symptoms as described in studies such as Wilkins et al. (1994), and potentially improved reading results.

While this line of reasoning could explain the benefits of colour found in some experiments, the authors never explicitly state the reasons behind individual differences in colour specificity. Indeed, a study examining the low-level visual aspects of visual stress – that is, aspects which are objective - found no evidence for this specificity (Simmers et al., 2001). In this small-scale study, the experimenters compared the accommodation reflex of the eye of 5 participants who used coloured lens. They were unable to find a difference in the size of the accommodative reflex (which involves the eye's lens changing size to maintain a clear image) when the participants used a lens of the participant's favoured colour compared to a lens of a complementary colour.

Finally, future research could examine one potential prediction of the cortical hyperexcitability account. A recent study, operating on the assumption that migraine is caused by cortical hyperexcitability, suggests that transcranial magnetic stimulation (TMS) is potentially a method to restore normal cortical excitability (Brighina et al., 2009). If this is the case, an interesting question could be asked: could a similar methodology alter symptoms in individuals with visual stress? If Wilkins' theory of abnormal cortical hyperexcitability in visual stress is correct, a potential prediction would be that the visual distortions associated with visual stress would be reduced with TMS.

### *2.3 Noise Exclusion Deficit Theory*

A third theory which has some bearing on the use of coloured lenses for reading is that poor readers have deficits in visual and auditory noise exclusion when compared to normal readers. Sperling et al. (2005) attempted to differentiate between this hypothesis and the magnocellular deficit theory (see above). They assessed contrast sensitivity to both magnocellular- and parvocellular-activating stimuli in 28 dyslexic children and compared their performance to that of 27 controls. In the presence of visual white noise, dyslexic participants were found to have significantly higher contrast thresholds to detect a signal.

In a further study, Sperling et al. (2006) found the same effect for both children and adult dyslexics on a random-dot kinematogram task – in the absence of noise, dyslexic participants performed similarly to controls, but in the presence of noise, dyslexics required significantly higher coherence in the dots before motion was recognized. The authors' conclusion from both these studies was that dyslexics have particular problems with extracting salient stimuli from irrelevant noise. Sperling et al. (2006) suggest that this, and not a magnocellular deficit, may be the basic cause of dyslexic symptoms.

Noise exclusion theory has been directly related to the use of coloured filters in a study by Northway et al. (2009). Dyslexic adults performed a symbol discrimination task in the presence and absence of Gaussian noise in three further conditions – without a filter, while wearing a neutral-density filter, and while wearing an optimal coloured filter (selected using the method described by Wilkins (2003)). Dyslexics were found to have higher discrimination times in the presence of noise, as predicted by Sperling et al. (2005, 2006). Their performance while wearing the 'optimal' filter was found to be significantly better only in the presence of noise – indeed, their discrimination time was reduced to that of a set of non-dyslexic controls. In the absence of noise, there were no significant differences between dyslexic and control performance.

According to Northway et al. (2009), the noise exclusion theory is not incompatible with cortical hyperexcitability theory. Indeed, they explain their results by suggesting that coloured filters help reduce cortical suppression, which may be the root cause of the dyslexics' inability to efficiently exclude noise. Further research would have to be performed to directly test this hypothesis.

There are a number of problems with the noise exclusion hypothesis, however. Normally, reading is not performed in conjunction with white or Gaussian noise, and it is therefore unclear whether the results found in Sperling et al. (2005, 2006) and Northway et al. (2009) can generalize to everyday reading performance. This is especially true since all the studies discussed so far

assess single-word or single-symbol reading in noise, and not reading lines of text. Since the cortical hyperexcitability theory is at least partially based on the assumption that the lines of text act similarly to a visual grating illusion, one would not necessarily expect white or Gaussian noise to cause the same effects. In addition, the participants in the studies by Sperling et al. (2005, 2006) were not asked to undergo any optometric or orthoptic examination, and therefore undiagnosed oculomotor or orthoptic problems may have influenced the participant's performance. Finally, the noise exclusion theory does not explain the specific phonological problems found in most dyslexics (e.g. Hulme & Snowling, 2009; White et al., 2006; see also Section 1.1) which are present with or without visual noise.

Not all studies have found noise exclusion deficits in disordered reading. Shovman & Ahissar (2006) compared the performance of 20 adult dyslexic participants to 20 controls on a variety of visual measures, including signal-noise discrimination. They found no differences between dyslexics and controls on any measure, leading them to conclude that visual deficits, while possibly present in some of the dyslexic sample, do not impinge upon reading performance.

Northway et al. (2009) explain the lack of findings in Shovman & Ahissar (2006) by noting that in that study, the authors did not select visually symptomatic dyslexics, as in the Northway et al. (2009) study. However, given that previous studies of noise exclusion in dyslexics (Sperling et al., 2005, 2006) chose samples which were not necessarily only individuals reporting visual symptoms and nonetheless found noise exclusion deficits, this objection is unconvincing.

While the status of the visual noise exclusion deficit in poor readers is contested, the evidence for an auditory noise exclusion deficit in that population is much stronger (e.g. Chait et al., 2007; Ziegler et al., 2009). However, Ramus & Szenkovits (2008) note that all these deficits can be considered as part of the phonological deficit model of dyslexia (see Section 1.1). In this interpretation, the noise exclusion difficulties are caused by the individual's poor access to phonological representations, and not a low-level sensory problem. If this explanation is correct, the invocation of cortical hyperexcitability, or a separate and as yet unknown stimulus-in-noise extraction problem, as an explanation for the noise exclusion deficits is rather unparsimonious. It is safer to state that the existence, prevalence and cause of noise exclusion deficits in poor readers has not been uncontroversially established.

## *2.4 Summary of Theories*

This section reviewed the viability of three theories – magnocellular deficit theory, cortical hyperexcitability theory, and noise exclusion theory – in explaining the effects of coloured filters on reading disorder. It is certainly true to say that none of them is firmly established – indeed, proponents of all three theories admit that they are tentative and require a good deal more research and improvement to be confirmed. However, it can also be seen that some significant problems exist in the formulation of the three theories as they stand.

Magnocellular deficit theory as an explanation of dyslexia is highly controversial (Skottun, 2000), becoming even more controversial when it is applied to the effects of coloured filters. Most poor readers appear to have no magnocellular deficits (Ramus, 2003) and many people with magnocellular abnormality have no reading disorder, or any other visual symptoms (Castelo-Branco et al., 2007; Skoyles & Skottun, 2004; Vellutino et al., 2004). Experiments attempting to link magnocellular function to the effects of colour on the visual system have been deeply flawed methodologically (see e.g. Skottun, 2004), and even proponents of the use of coloured filters have noted that magnocellular function seems to be an entirely separate variable from an individual's receptivity to coloured filters (Wilkins, 2003).

In the time since Wilkins cautiously proposed the theory of cortical hyperexcitability to explain his findings about coloured filters (see Wilkins, 1995), conspicuously little has been added – the most recent formulations of the theory (e.g. Wilkins & Evans, 2009) do not differ in any major details. This is perhaps due to the small number of studies performed in the area, mostly analyses of the effects of coloured filters in populations such as those with autism or multiple sclerosis, who may display symptoms the researchers have likened to those of visual stress (e.g. Ludlow et al., 2006; Newman Wright et al., 2007). There is certainly value in this line of research - any new interventions for these conditions and disorders would be of great importance - but at the present stage, the majority of such studies remain unreplicated and therefore the theory drawn from them can only be uncertain. A great deal more work is required to improve cortical hyperexcitability theory.

While not incompatible with cortical hyperexcitability theory, noise exclusion theory has been proposed as a separate explanation of the effects of coloured filters (Northway et al., 2009). However, while there is moderately strong evidence that poor readers suffer from an auditory noise exclusion deficit, the data collected thus far relating to a similar deficit in the visual modality is much more equivocal (e.g. Shovman & Ahissar, 2006; Sperling et al. 2006). Noise exclusion deficit theory was proposed relatively recently (Sperling et al., 2005), and even more recently as a mechanism for coloured filter's effects (Northway et al. 2009). Clearly the theory is

in its initial stages, and the basic hypotheses will require corroboration by a larger and more wide-ranging body of evidence.

Of the three theories discussed in this review, which represent the entirety of attempts to explain the effects of coloured filters with reference to previously-existing evidence, none seem at present completely convincing. The strongest of the theories, possibly due to the fact that it is aimed specifically at the effects of coloured filters as opposed to being a theory of reading disability in general, is the cortical hyperexcitability theory, but as noted above there is rather a dearth of solid, replicated evidence to support this theory. It remains to be seen whether this theory, or any of the others, will increase in plausibility with future research.

One last point worth making is that the researchers from the various theoretical perspectives rarely appear to make reference to each other's work. Proponents of the magnocellular theory, for example, do not make reference to visual stress as a discrete condition. Proponents of cortical hyperexcitability criticize the magnocellular theory, but make no attempt to explain the results found in magnocellular-based experiments, preferring to criticize the theory in general. Representatives of the Irlen Institute do not refer to any of the competing theories, other than to state that the treatment methods stemming from them are inferior to the Irlen Method (Irlen Institute, 2010). There are clear contradictions; for instance, the colour specificity found by Wilkins and Irlen is inconsistent with the research of magnocellular theory proponents, who claim that only yellow and blue filters are effective.

It is clear that the somewhat confused literature would benefit and be clarified significantly from a greater level of engagement with – even enlightening criticism of – opposing theoretical positions from each group of proponents.

### 3. EVIDENCE FOR COLOURED FILTER USE

#### 3.1 General comments

This section will review the current evidence base for the use of coloured filters, both plastic overlays and tinted lenses, and will attempt to answer two separate, though most likely related, general questions which are sometimes conflated in the literature when asking if coloured filters ‘work’. First, do coloured filters reduce reported symptoms of visual stress? Secondly, and most importantly for the current purposes, do coloured filters improve skills such as reading speed, accuracy, or comprehension in individuals with reading disability? This section will not concern itself with potential mechanisms for the reading improvements or visual stress reduction – these have already been discussed. It will focus instead on practical and clinical applications.

The questions have been approached in a wide variety of ways. Case studies, tests which assess the immediate effects of coloured filters, studies with a longitudinal design, and full randomised controlled trials have all been employed to investigate the effects of coloured filters. Some, but not all, studies have introduced elements of blinding, controls, and placebo treatments.

Unfortunately, as well as in methodology, the studies vary considerably in quality. Albon et al. (2008) have produced an excellent systematic review of this literature pertaining to the second question above, entitled ‘The effectiveness and cost-effectiveness of coloured filters for reading disability’ from which this review will take a lead. Before examining the studies, however, it is useful to consider some general concerns about the literature.

A serious issue with many of the studies, is the choice of sample population. Several studies have used samples of individuals who report visual stress symptoms, but do not necessarily have reading disorder (e.g. Hollis & Allen, 2006; Jeanes et al. 1997; Noble et al., 2004; Waldie & Wilkins, 2004). The reliability of these studies to answer the main question about the effectiveness of coloured filters in alleviating reading *disability* is necessarily limited.

Related to this, Singleton & Trotter (2005) have noted that the definition of ‘visual stress’ is somewhat confused between what we may refer to as a ‘response’ versus a ‘symptomatic’ definition. The ‘response’ definition characterises visual stress as a response to coloured filters (that is, an increase in reading rate or some other objective measure when the filters are used), and can be found as the operating assumption in studies such as Evans & Joseph (2002) and Kriss & Evans (2005). Typically, a 5% increase in reading rate on the Wilkins Rate of Reading Test (Wilkins et al., 1996) when using the ‘optimal’ overlay is regarded as an indicator that the subject has visual stress (Kriss & Evans, 2005, Table 1). Aside from its arbitrariness, and the fact that in

most cases an increase of 5% is hardly vast, there is a circularity in this diagnostic criterion: it both overlooks the fact that visual stress may not be the only reason for the apparent benefits of coloured filters (there could be placebo effects to consider), and ignores important questions about the efficacy of the filters in a wider sample of individuals.

Some randomised controlled trials (Bouldoukian et al. 2002; Wilkins et al. 1994) have only considered as participants individuals who had been using coloured filters for some time before the trial, and so also fall victim to this circularity in their method. While they can show the more fine-grained effects of the filters in the sample who already receive benefit from them, and can potentially indicate avenues for theoretical investigation of these participants, their generalisability to other populations is low. Thus, the ability of these studies to answer questions about how useful coloured filters are for the general population of individuals with reading disability is restricted.

The second, 'symptomatic' definition is based on reported visual stress symptoms. This has been assessed using visual stress questionnaires such as that of Irlen (1991), using equipment such as the Intuitive Colorimeter (see above) or asking associated questions or by investigating symptoms reported in an educational setting. This definition has its own caveats, which are discussed alongside the evaluation of the study by Wilkins et al. (1994) in the 'Randomised Controlled Trials' subsection below.

To assess which of the definitions is the more useful, the age of the experimental participants is important; Singleton & Trotter (2005) point out that the responses of children to visual stress questionnaires have the potential to be very unreliable, as children are easily led by the content and phrasing of questions. The authors see no reason to distrust the symptom reports of adults, however, and therefore recommend using 'symptomatic' measures where possible with adults. Some studies in this area (e.g. Blaskey et al., 1990; Bouldoukian et al. 2002) have used both children and adults of various ages in their samples, which may be inadvisable due to this problem of response reliability.

Finally, many authors (e.g. Albon et al., 2008; Blaskey et al. 1990; Scheiman, 2004; Singleton & Trotter 2005) note the critical proviso that anyone reporting visual stress symptoms may not necessarily be suffering from a distinct 'visual stress' condition – the presence of other orthoptic or optometric disorders such as strabismus or amblyopia can often account for the symptoms with an explanation which is already known to eye care professionals. It is therefore recommended that each individual reporting visual stress symptoms should be referred for a full optometric assessment in addition to taking part in any experiment, to rule out other causes of their symptoms. The results of this assessment should be reported along with the rest of the data.

In their systematic review, Albon et al. (2008) identified 8 randomised controlled trials and 15 other studies which were suitable for inclusion. A further 67 articles were rejected for reasons such as their being reviews, measuring inappropriate outcomes, or including inappropriate participant samples (such as those without reading disorder). This review will examine and evaluate a selection of notable non-randomised controlled trials and some randomized controlled trials as defined by Albon et al. (2008), with the addition of several trials which were not included in that review. The Albon et al. (2008) review and meta-analyses will then be considered in more detail. Finally, conclusions will be drawn about the current state of the evidence for the use of coloured filters.

### *3.2 Non-Randomised Controlled Trials*

Before the larger-scale studies are examined, a case study of two dyslexic children performed by Hannell et al. (1991) will be considered. These authors had encountered a great many anecdotal accounts of the efficacy of coloured lenses, and so resolved to perform a scientific investigation into the use of this treatment. The two participants, who were brothers aged 9 and 12 years, were given a variety of reading, writing, and perceptual tests both with and without their chosen colour of lenses. The colour was selected via a process of elimination from a choice of 5 colours, a process not as fine-grained as the Irlen procedure or later procedures involving the Wilkins Intuitive Colorimeter, but one which apparently still produced positive results for the two participants.

The experimenters found substantial improvements in oral reading, writing, and figure-ground discrimination in both participants, and a more accurate perception of the spaces in between printed words in one participant when using the coloured glasses. Since this was a case study, the authors did not attempt to control for placebo. They note that ‘...some degree of faith has to be put in the subjects’ integrity. In this case both boys were felt to be working honestly...’ (p. 69). For this reason, the study cannot be considered solid evidence for the efficacy of coloured filters; it instead acts as an interesting inspiration for further research in this area.

Menacker et al. (1993), performed a well-designed, if small-scale, cohort study into the effects of coloured lenses. The participants in this study were 24 children (aged 8-12 years) who had all been formally diagnosed with dyslexia by a qualified psychologist. Thus, while this study did not have a large sample size, it did not suffer from the same problems of heterogeneity in participant age and reading ability as discussed above.

After establishing baseline reading ability with the San Diego Quick Assessment reading test (LaPray et al., 1972), the experimenters tested the participants using reading level-appropriate passages from standard classroom reading textbooks in six randomly-ordered conditions, with participants looking through: a red lens, a blue lens, a green lens, a yellow lens, a colourless (neutral-density) lens and no lens. No significant trend in reading performance – measured by the error rates on the test – was found for any of the colours.

The participants were also asked which of the lens conditions they preferred – that is, which condition made reading most comfortable. Interestingly, the subjectively-preferred lens was not found to correlate with objective reading performance in any condition. The authors point out that it is difficult to reconcile these findings with the theory of Irlen (e.g. 1991) that an immediate increase in comfort owing to coloured lenses will go on to improve reading ability.

Menacker et al. (1993) note a possible criticism – their experiment did not take colour specificity into account. The four coloured lenses used were not individually-prescribed to the participant, nor were they from the colour charts of any of the colour systems, such as the Irlen system. However, the authors respond to this criticism by noting that they ‘used the four primary perceptual hue categories available to human vision (red, green, yellow and blue)’ (p. 218) and that human vision is too coarse to expect different results from a wider variety of colours. This explanation is incompatible with evidence from Wilkins (e.g. 2003), who finds colour specificity is important.

Kriss and Evans (2005) examined the relationship between dyslexia and visual stress, and in doing so tested the efficacy of coloured filters. They took a sample of 64 children aged 7-12 years, half of whom had a formal diagnosis of dyslexia and half of whom were age-matched controls. The participants chose the coloured overlay which gave them the most comfort when viewing text from a set of 30 Intuitive Overlays (including double overlays; Wilkins, 1994). They were then all tested on the Wilkins Rate of Reading Test (WRRT; Wilkins et al., 1996).

The WRRT consists of a passage of closely-spaced text, with 15 common words in random order repeated on each line. The participant reads these words out loud, and the number of words read out correctly in one minute is noted. The test, then, is specifically and purely designed to investigate the effects of overlays on rate of reading; any visual reading problems will be exacerbated by the layout of the test, and since the words are random and in no particular order, there is no comprehension element. In the Kriss and Evans (2005) study, it was found that, with the overlays, reading rate was significantly faster in the dyslexic group but not the control group.

Albon et al. (2008) point out that it should be taken into account, especially when considering some of the longitudinal studies considered below, that WRRT results have been shown to have a significant practice effect; at both immediate retest and retest after 8 weeks, participants read on average 2.9% faster than previously (Wilkins et al., 1996). This slightly reduces the extent of the conclusions which can be drawn from studies using the WRRT over a space of time, and which assume that an intervention (most often coloured filters) is responsible for the increase in reading rate.

As mentioned above, the authors in this study diagnosed a participant with visual stress if they showed a particular response to the filters as tested on the WRRT, in this case an increase in reading rate of greater than 5%. It was found that visual stress was present in the non-dyslexic control group at a prevalence of 25%, and in the dyslexic group slightly more commonly, at 47%.

On this basis, they conclude that visual stress is common in the general population, but more common in individuals with dyslexia.

Noted above are the concerns other authors have had with this method of diagnosis (Singleton & Trotter, 2005) – it contains a level of circularity. Kriss & Evans (2005) did not include a ‘symptomatic’ measure of visual stress, such as a symptoms questionnaire – this would have allowed a more complete indication of the prevalence of visual stress in their sample. Incidentally, it is also concerning that the participants in this study did not undergo an optometric examination to test for other visual problems, even though these are noted by the authors as a potential cause of visual stress symptoms (p. 352). For this reason, the authors risked confounding the prevalence of visual stress in their groups with the effects of other visual problems.

In addition, Kriss & Evans (2005) note that there were several methodological flaws in their study. Firstly, and most importantly, the method by which the coloured overlays were selected may have been imperfect – the authors note that ‘a marked number of participants in both groups selected mint-green overlays’ (p. 356). Noting that the mint-green and grey overlays (the latter of which were also chosen far more often than any other colour, aside from mint-green) were first shown to participants towards the end of the selection process. ‘It is possible’, they note, ‘that participants could see from the pile that there were not many overlays left, and therefore thought in order to complete the task successfully they must select an overlay’ (p. 360).

This observation speaks to a very important point in studies of this nature. Very few studies, most likely for reasons of brevity, report the colours of the filters chosen by each participant. However, if the selection process is leading participants to choose a colour arbitrarily, and not because it truly reduces their visual stress symptoms, the notion of colour specificity (e.g. Wilkins, 2003) could be called into serious question. Future studies should contain an appendix section noting the frequency of choice of each colour.

The results of other studies in this area have rather variable results. For instance, Kyd et al. (1992) studied the effects of Irlen overlays on 14 children who had been diagnosed as suffering from visual stress by an Irlen diagnostician, as compared to 14 control participants with no reading disorder. The Neale Analysis of Reading (Neale, 1989) was used as the measure, and the authors found that while reading rate increased significantly for the treatment group while they were using their chosen Irlen overlay, there was also an ‘unexpected’ (p. 29) detrimental effect on reading accuracy and comprehension in 43% of the treatment group. The authors do not discuss whether this result points to the overlays engendering a greater ability to read words quickly without properly processing them, or some other explanation.

Interestingly, the results of Kyd et al. (1992) can be contrasted with those of Iovino et al. (1998), who studied a sample of 45 children with reading disorder, in a rather different design. These authors tested the effects of blue overlays, as compared to red overlays and a 'no overlay' condition, on the Formal Reading Inventory (Wiederholt, 1986). Unlike Kyd et al. (1992)'s participants, in this study reading comprehension and accuracy increased significantly in the condition with blue filters, but reading rate decreased. The previous two studies are mentioned to highlight the differences in results which are perhaps to be expected due to the wide degree of variance in methodology in this area.

Many of the studies examined thus far have concentrated on the immediate effects of coloured filter use, occasionally including some small-scale measure of long-term effects as a secondary measure (e.g. Bouldoukian et al. 2002). Other studies, however, have had a longitudinal focus. One such example is Noble et al. (2004) who investigated the long-term reading performance aided by Irlen overlays of 71 children diagnosed with visual stress by an Irlen diagnostician. This study is notable since it tested the reading progress made by children using coloured filters, as opposed to comparing one-time reading ability with and without filters. Noble et al. (2004)'s study is not included in the systematic review by Albon et al. (2008), and no explanation for this omission is given.

The participants were split into two groups, and both were tested on the 4<sup>th</sup> Edition of the Gray Oral Reading Test (GORT; Wiederholt & Bryant, 2001). The first group was given immediate access to Irlen overlays, and the second had their treatment delayed by three months, at which point the two groups were tested again. Then, overlays were prescribed for the second group and after another three months both groups were tested a final time.

The results showed that the 'immediate treatment' group made more reading progress than the 'delayed treatment' group on all measures of the GORT (rate, accuracy, fluency and comprehension) in the first three months. After six months, the reading improvements of the immediate treatment group had reached a plateau, but the delayed treatment group had improved considerably, reaching the same level of ability. The authors interpret these results as showing that the Irlen overlays were able to remove an obstacle to reading, but that further tuition would have been required before the children's skills would have risen above the plateau.

The authors discount the placebo effect as an explanation of their results, stating that the children used the overlays without prompting for much longer than would be expected for a treatment which had no actual effect. As noted above, this assumption is not at all assured. Placebo effects, being evoked by belief, should be expected to last as long as that belief. The authors go on to

assert that 'the large gains in reading skills by both groups is also likely to be beyond what might be expected from increased parent and teacher interest and support, or the novelty of a new method' (p. 21). Again, no explanation or additional evidence is given for this statement. It could be argued that, since no blinding of the parents, teachers, or children was undertaken in this study, the combined effects of expectation, novelty, and other placebo-related responses (on the part of the participants or those instructing them) could potentially explain the findings, or at least a proportion of them.

It is to be noted that the participants in this experiment did not necessarily have reading disorder; indeed, the participants were classified in three education categories – those in regular education, those in transition from Spanish-language classes to English-language classes, and those in special education. The authors argue that this heterogeneity did not affect their results - statistical tests revealed that, over the time periods, these means stayed equally far apart, indicating that the effects of the overlays were the same in each education category.

One major concern with longitudinal studies, especially of children in an educational setting, is controlling the amount of input given to the children by their teachers in the time periods between experimental tests. Noble et al. (2004) mention only that the teachers in their study were instructed not to encourage the participants to use the overlays (in order to assess whether or not the effects of the overlays were due to their novelty). A major flaw, then, of the study is that the first group of children were all at a different school to the second group. If the school teachers and reading curriculum were different in both schools, it is hard to see how the experiment could have been adequately controlled.

Related to this, we have seen previously when discussing the study of Stein et al. (2000b) that, if one's hypothesis is based on a property of children's visual system, one must be extremely careful in a longitudinal study of imputing any improvements in visual properties to a treatment intervention when in actual fact these improvements could have come about with the normal process of maturation.

One longitudinal study which better controls for the problems of other experiments is that of Scott et al., (2002) who, like Noble et al., (2004), included children from two schools, but used a within-subjects design so that varying instructional input would not be a significant issue. In the first of two experiments, these authors tested 153 children – all of the children in one year of the school, aged 10-12 years - with coloured filters using the Intuitive Overlays method (Wilkins, 1994). Those reporting benefits – more clarity of text, less visual symptoms – were given their preferred colour of overlay to use in their regular home or classroom activities for 4 months. After this time, a random sample of 94 of these children was taken, and split into those who were

still using the overlay at the end of the 4 months (the 'overlay' group, consisting of 31 children), and those who were not (a control group, with 62 children).

The children in both groups were then tested on the WRRT and other reading measures, and then given an orthoptic assessment. The children in the overlay group read on average 8 more words per minute with the overlays, which was statistically significantly faster. The control group, who had faster speeds in general than the overlay group, did not read significantly faster with overlays. This is, to some degree, unsurprising – the children had been grouped on the basis of their preference to use overlays, the result showing that they did indeed benefit from using them is to be expected. Several other studies in this area have been noted by Albon et al. (2008) to involve a selection bias of this nature, as we will see later in this section.

Most interesting are the findings from the orthoptic tests. The results showed that the children in the overlay group had significantly poorer amplitude of ocular accommodation than the controls, and that the children who favoured an overlay with a long-wavelength colour (colours on the red end of the spectrum) for their overlay tended to have a greater amplitude of accommodation – though only on binocular tests - than those who chose a short-wavelength colour (those on the blue end of the spectrum). This was, as the authors note, a small effect – a difference between a mean amplitude of 14D in the long wavelength users and 12.5D in the short wavelength users. The authors do mention that 'a mean amplitude of accommodation of 12.5 D should be more than adequate to allow normal reading' (p. 162), perhaps reducing the clinical significance of their findings.

In the second experiment described by Scott et al. (2002), 104 children aged 7-11 years from a different school were assessed with Intuitive Overlays, and given their preferred colour, as in the first experiment. Some months later (most likely 4-5 months, but no exact times are given by the authors), the children were split into groups based on usage, as above (20 in the overlay group and 98 in the controls – the authors do not explain the disparities in sample size described in their report), and received an optometric test and the WRRT. No results whatsoever from the WRRT in this experiment are reported by the authors.

As in the first experiment, children who chose longer-wavelength colours for their overlays tended to have significantly higher binocular (though not monocular) amplitude of accommodation than those who chose shorter-wavelength colours. The tests in this second experiment that were different from the first – mean accommodative lag, refractive errors – did not show any significant differences between the two groups.

It is worth noting that in both the experiments of Scott et al. (2002), the sample numbers were widely different in the 'overlay' and 'control' groups (31 vs. 62, and 20 vs. 98), which may have influenced the results of the statistical tests performed. It would be interesting to see a similar analysis of orthoptic and optometric tests performed on two groups more matched on sample size.

The authors note that their results are consistent with those of Simmers et al. (2001), who also found abnormal accommodation in those reporting visual stress symptoms, in a sample of 5 adults aged 22-30 years. Simmers et al. (2001) showed that the magnitude of accommodative microfluctuations in the eyes of the participants was higher when they were not using any lens than when they were using a tinted lens of a preferred colour. However, the microfluctuations were just as low in conditions which involved a lens of complementary colour or a colourless, neutral-density filter.

Simmers et al. (2001) conclude that, firstly, accommodative microfluctuations may be an indicator of visual stress, and secondly, that lenses used to alleviate these fluctuations do not need to be chosen on the basis of individually-preferred colour. Clearly, these results conflict with those of Scott et al. (2002), who found accommodation differences relative to chosen colour. The wide differences in sample size in these two studies may be an explanation of this difference, but the disparity confirms once again the confused nature of the evidence on the subject of coloured filter use.

Nevertheless, in concurrence with the opinion of Simmers et al. (2001), Scott et al. (2002) note that 'accommodative mechanisms are unlikely to be a sufficient explanation for the individual choice of colour and its effect on reading fluency' (p.164), suggesting instead that such mechanisms and any problems with them are a correlate, rather than a cause, of visual stress. This conflicts with the results of Blaskey et al (1990), discussed in Sections 3.3 and 5.2 below. Scott et al. (2002) conclude by favouring an explanation in terms of the cortical hyperexcitability theory, discussed above.

Some general methodological criticisms apply to this study. As well as the fact that, curiously, only one set of WRRT results from this experiment are reported, this study's particular usage of the test gives some cause for concern. The WRRT is designed in such a way that two passages are read by the participants under each condition (overlay and no overlay). It is potentially concerning that Scott et al. (2002) note that 'the Rate of Reading test was administered *once* with and *once* without the overlay, in random order' (p. 159; emphasis added). While this may have been convenient so as not to tire the children in the study, it may also mean that practice and fatigue effects were not adequately controlled for.

Finally, and a major problem for the generalisability of the results from the study, we are not aware of how many children in the experimental samples had reading problems of any kind. No baseline reading results were taken, and two school years were sampled in their entirety. The lack of formal diagnoses of, for instance, dyslexia in samples has been a concern of some authors in this area (e.g. Christenson et al., 2001).

In a study of a somewhat different nature, Northway (2003) studied the value of two different tests, the WRRT and the Developmental Eye Movement test (DEM; Garzia et al., 1990), for predicting long-term use of the overlays. This was in an attempt to investigate the extent to which the overlays affected the different aspects of reading that the tests assess – reading rate and eye movement scanning/tracking automaticity, respectively.

It has been suggested that an individual showing a 5% increase in rate of reading on the WRRT is a good predictor of continued use of coloured overlays (Wilkins et al., 1996; see also Kriss and Evans, 2005, Table 1). The other test used, the DEM, involves participants reading down a vertical row of random numbers, then along a horizontal row. The ratio of vertical to horizontal reading time is then calculated, and matched to a table of average ratios by age. It can therefore be used to determine if a participant is below the appropriate age bracket for automaticity in eye movements (Garzia et al., 1990).

The Northway (2003) study had a good sample size – 64 children who had a formal diagnosis of dyslexia and no visual anomalies were prescribed an overlay using the Intuitive Overlay method (Wilkins et al., 1996), and were then assessed on both the WRRT and the DEM, with and without their chosen overlay. After a period of 12 weeks, the subjects were split into three groups – those who continued to use their overlay for this time period, those who stopped using their overlays before the end of the time period, and those who did not select an optimal overlay (and were therefore prescribed a random overlay).

A significant improvement in DEM performance was found when optimal overlays were used. In addition, performance on the DEM was found to be a good predictor of the length of overlay use – a better predictor than WRRT performance change. Northway (2003) was the first study in this field to use the DEM, and suggested that other future studies use it, as it may be a good indicator of how long children will use the overlays.

Explanations for this finding – that in many participants reading rate does not necessarily improve significantly, but automaticity does - are mysterious; Northway (2003) suggests several possible reasons, for example the fact that the DEM consists of rows of single numbers rather

than words. This 'may result in this test being cognitively less demanding and small improvements in visual perception are more readily seen' (p. 463). It is also suggested that 'the crowded textual appearance in the [W]RRT may be a significant and more difficult obstacle in the presence of poor eye movement control in those children found to have poor scanning technique in the DEM, and the benefits of overlays are less obvious' (p.463). A finer-grained test of the efficacy of coloured-filters is certainly welcome, but since Northway (2003) is the only study to use the DEM, it would be prudent to wait for replications before recommending the use of the DEM as a regular tool in clinical trials.

Since the Northway (2003) study did not have a control group, any blinding, or a placebo comparison, it is not high-quality evidence to suggest that overlays have an effect on reading performance. However, the study is valuable since it emphasizes the importance of testing participants on a variety of different measures, including - but not limited to - measures of reading rate such as the WRRT.

### *3.3. Randomised Controlled Trials*

The first randomised controlled trial of the efficacy of coloured filters was that of Blaskey et al. (1990), which specifically focused on the Irlen method and how it compared to vision therapy. This latter intervention involves participants attending a once- or twice-weekly meeting with a clinician who employs various training techniques to treat orthoptic disorders.

To recruit participants, the experimenters advertised for those interested in the Irlen method. They then accepted only those individuals who had been diagnosed with visual stress by an Irlen diagnostician as well as on the basis of a visual stress symptom questionnaire, and who had also been given a detailed optometric examination to reveal any ocular problems, such as those mentioned above. The 30 participants, aged 9 to 51, were split into the three groups, the first of which received Irlen lenses, the second of which had unrelated vision therapy (which involved an average of thirteen sessions specifically targeted to treat the problems identified in the optometric examination – however, no details of the content of these sessions are given), and the third of which was a control group which received no treatment.

An element of blinding was introduced into the Irlen lenses group, who were given trials for two weeks each of both their optimal lens (as prescribed by an Irlen diagnostician) and a random, ‘placebo’ lens, before choosing which of the lenses they preferred and using them for a further two weeks. Neither the participants nor the experimenter who performed the testing knew which was the optimal lens.

Both before and after their treatment, the participants were administered a battery of reading tests which included parts of the GORT (this study specifically used the results for word recognition in context and speed; see Gray, 1963) and parts of the Woodcock Reading Mastery Test (those which examined recognition of isolated words; see Woodcock, 1987). They also filled out visual stress symptom questionnaires after treatment. The experimenters found that subjective visual stress symptoms seemed to decrease in both the treatment groups, but participants showed very little improvement on the reading measures in the Irlen group. In the vision therapy group, the participants did improve their GORT scores significantly.

One possible interpretation of these findings is the following: the Irlen lenses reduce only subjective visual stress symptoms through either an unexplained mechanism or the placebo effect, but are unable to produce any objective reading improvement. The vision therapy, on the other hand, does begin to treat the visual problems which cause visual stress symptoms, and therefore has more impact on the reading measures. Indeed, after vision therapy treatment, the scores of the ‘vision therapy’ group on the Irlen scotopic sensitivity test had dropped to the point

where they would no longer have been considered candidates for the Irlen method (Scheiman, 2004). However, no further research directly comparing vision therapy and coloured filters has been performed, so this interpretation is necessarily tentative (though see Discussion section, below).

The most important aspect of the Blaskey et al. (1990) study is its highlighting of the point that the symptoms which are often classified as part of the 'visual stress' syndrome may in actual fact arise from, or be confused with, a range of vision problems. Studies that do not take this into account risk rendering their results questionable. Scheiman (2004) underlines the critical need, mentioned above, for thorough assessment of refraction, accommodation (amplitude, frequency and response), binocular vision (eye alignment, fusional vergence, and vergence facility), and ocular motility (saccadic speed and accuracy)' (p. 40) in all participants in visual stress-related experiments, in order to ensure that those apparently suffering from visual stress are not simply those who have received poor vision care.

While the Blaskey et al. (1990) study is of good quality, and included adequate eye examination for its participants, there are a number of other potential methodological issues which, had they been addressed, would have improved the validity of the results. For instance, the authors themselves point out that the sample size was rather small, and a larger group would have allowed better comparisons of the treatments. Similarly, the reading abilities of the sample before treatment may well have been highly heterogeneous – the only criterion for inclusion related to reading was that the participants felt 'they were not reading up to their potential' (p. 606).

Lastly, the exclusion of any participants not displaying visual problems as measured by an optometric assessment gives some cause for concern; showing that visual stress symptoms are sometimes caused by other visual problems is not the same as showing that visual stress is not also an independent disorder. This question could not be addressed by the Blaskey et al. (1990) study.

The next randomised controlled trial to be considered is that of Wilkins and his research group, who examined two potential effects of coloured filters: firstly, reducing visual stress symptoms, and secondly, improving reading ability (Wilkins et al., 1994). After a 'full optometric assessment' about which the authors provide no details, participants were assessed using the Intuitive Colorimeter, and were prescribed coloured lenses (and a pair of control lenses) on the basis of their colour preference. The participants were not made aware of which were the optimal lenses, and indeed it was shown that they could not reliably distinguish between the two sets of lenses.

The experimental methods were as follows. The children were assessed on the Neale Analysis of Reading (a standard reading test which gives scores for reading speed, accuracy and comprehension after reading several passages; see Neale, 1989) once after wearing the optimal (experimental) lenses for three weeks, and once after wearing the control lenses for three weeks. They were also asked to keep 'symptom diaries' of their visual stress symptoms during both of these time periods.

To be included in the experiment, the participants had to have been using an optimally-coloured overlay (chosen by a process of elimination from a selection of ten colours) for three months prior to testing. It is also worth noting that the participants in this study were deliberately not a homogenous sample – the children from one of the schools used to recruit participants were those with reading disorder, and a second set from another school were those who had reported visual stress symptoms, but not necessarily reading disorder. The authors explain that the study was designed to investigate whether *any* category of participants can benefit from coloured lenses. The study began with a sample of 68 children (of average age 12 years 2 months), but had a high attrition rate - 15 participants did not complete the reading experiment, and a further 16 failed to complete their symptom diaries.

Nevertheless, Wilkins et al. (1994) found that the coloured lenses did have a significant effect on visual stress symptoms – the number of days with reported symptoms in the symptom diaries was significantly lower for time period when the optimal lenses were worn than it was for the period with the control lenses. The study did not, however, find any significant effect on reading ability for the optimal lenses, an outcome which the authors explain with reference to the lack of homogeneity in the sample. They state that their sample size was too small to isolate the subgroup of participants (i.e. those with headache, asthenopia, or reading difficulty) who benefitted most in their reading score.

Due to the high attrition rate leading to a small sample size, the heterogeneity of the sample, and the fact that the participants had been using coloured filters prior to testing, it is difficult to draw any firm or generalisable conclusions from the Wilkins et al. (1994) study.

Incidentally, an interesting question arises from the methodology of the Wilkins et al. (1994) study, and applies to other studies in which visual stress is diagnosed on the basis of symptoms, especially in the case of participants with reading disorder. Participants may be willing to attribute their reading difficulties to anything an interviewer or diagnostician suggests, even if the symptoms are not present in actuality. For this reason, during diagnosis, asking precise 'leading' questions such as 'do the words and letters appear to move around on the page?' as opposed to more neutral questions such as 'what do you see when you look at the page?' seems

inadvisable. At this point, no reliable objective physiological correlate of visual stress has been found; until such a correlate is revealed, and since the responses to visual stress questions and questionnaires are purely subjective, it is sensible to keep diagnosis as open and neutral as possible.

As part of the same research team, Bouldoukian et al. (2002) followed up the Wilkins et al. (1994) study by focusing on the participants' reading ability. Participants were chosen on the basis that they reported a reduction in visual stress symptoms when using coloured overlays, and the results of the Wilkins et al. (1994) paper - that visual stress symptoms are significantly reduced by optimal colour - were assumed.

The Bouldoukian et al. (2002) study used a rather heterogeneous sample, as previous studies. 20 of the subjects in study had been diagnosed with dyslexia, 3 more had 'specific learning disability', and the remainder had an unspecified diagnosis (in general, all participants were attending the Institute of Optometry because of 'suspected or diagnosed specific learning disability.' The authors provide no more information about). The sample was also somewhat heterogeneous in age: 4 adults aged 18-40 years and 29 children with an average age of 11 years. Since the experiment was a within-subjects design, the age and reading ability of the subjects being so varied is not a serious fault. However, it does make generalising the findings to other populations - for example, those with and without dyslexia - more difficult.

One interesting aspect of the Bouldoukian et al. (2002) study is the use of a placebo overlay along with coloured overlays with the Wilkins Rate of Reading Test. This placebo overlay was a pale yellow ultraviolet blocking filter, with minimal spectral absorption and therefore no significant effect on participant's vision of the text through it. In order to maximise the potential placebo effect from this overlay, the participants were informed that this overlay was 'a wonderful discovery to help people with learning disabilities' and the words 'Research Model' were inscribed in the top corner (p. 57). The study, then, did not include a 'no overlay' condition, and so was comparing the effects of the optimal overlay (once again chosen using the Intuitive Colorimeter) to controls.

Unfortunately, as well as allowing no baseline 'no overlay' comparison, this design made it impossible to blind either the experimenters or the participants to the nature of the overlays. In each case the experimenter knew which overlay was the individual (optimal) overlay and which was the control ('the examiners... knew the identity of the patient and of the overlays that were being used' (p. 58)), and the participants were aware of which overlay they had chosen, and which had been given to them. The inclusion of a second 'non-optimal' overlay, as in the Blaskey et al. (1990) study detailed above, would have allowed for at least partial blinding. Further, even

though the experimenters had, as mentioned previously, made substantial efforts to convince the participants that the placebo overlay was 'as special' as the chosen one, they could not change the fact that one overlay was selected in a process (using the Intuitive Colorimeter) and the other was simply given to the participant without the same kind of elaboration. The experimenters assume the perceived equivalence of a personal, individual treatment with a general one, and this assumption can certainly be called into question.

The results of the study showed a significant effect of the optimal overlays compared to placebo overlays on the participants' rate of reading on the WRRT. On average, rate of reading was 4% faster with the optimal coloured overlay than with the control. 66% of the participants who had a particular subjective preference after reading with both overlays preferred their optimal overlay to the placebo, and those who preferred the optimal overlay read on average 5% faster with it than the placebo, whereas those preferring the placebo overlay only read an average of 1.7% faster with it than the optimal overlay.

A further analysis was performed by splitting the participants into two groups, those who went on to use the filters for up to 6 weeks after the experiment, and those who only had an immediate (same-day) benefit. While the former group did receive a greater effect from the filters, it was not shown to be statistically significantly different from the effect received by the latter group.

While better-designed than several other studies in this area, Bouldoukian et al.'s (2002) experiment did suffer from some significant flaws. The study would, for example, have benefitted from a larger sample, and a design which included at least partial blinding of both participants and experimenters. In addition, the inclusion of only those participants who already report benefits from coloured filters, some of whom had been using the filters for some time previous to the study, is an unnecessary confound which prevents conclusions from being drawn about the effects of the filters (immediate or otherwise) in other groups of individuals. Indeed, the study can only be used to comment on those who already receive benefit from coloured filters - the study shows there is an effect in this group, but how that effect can generalise to other populations is unclear.

Finally, unlike other studies where participants with visual anomalies were excluded from the final analysis so as not to confound the data (e.g. Wilkins et al., 1994), Bouldoukian et al. (2002) note that a third of the participants had visual anomalies corrected before continuing on to receive treatment. While the authors note that, on occasion, 'symptoms are still present in people with specific learning difficulties whose ocular motor function is within normal limits' (p.56), it would have been interesting to see a baseline reading rate both before and after these visual anomalies were corrected. This would allow a more detailed comparison of the extent to which

the overlays in particular, as opposed to the visual correction, benefitted the individuals, and would have been a partial follow-up of the methodology of Blaskey et al. (1990), discussed above.

A longitudinal study which has already been discussed for its theoretical implications, is that of Ray et al. (2005), which was performed from the perspective of the magnocellular deficit theory. In it, 38 children aged 7 to 14 with severe reading difficulties were issued a baseline reading test with the British Ability Scales-II reading measure (Elliot, 1996), and then issued with either a yellow overlay or the experimental placebo – a piece of card with a ‘letterbox’ space cut out, so that only one line of text could be read at a time, for three months.

After this time, reading tests were taken again, and the group with coloured filters were found to have improved significantly more than the placebo group.

The main difficulty with the Ray et al. (2005) study, apart from the issues with all longitudinal designs as outlined above for the Noble et al. (2004) study, is the placebo, which could be argued to be somewhat inappropriate. It is not clear how two treatments that are so different from one another could be fairly compared. In addition, the lack of a control group, matched for educational input, receiving no treatment limits the conclusions which can be drawn about the effectiveness of either treatment.

Ray et al. (2005)’s statement that ‘all [participants] found that viewing through yellow filters improved the clarity of N5-sized text compared with blue, pink, green, red, or gray filters or no filter’ (p. 5/13) is interesting. Given previous work in the field by, for example, Wilkins (e.g. 2003), one would expect to find a variety of colours being chosen, as each individual is assumed to prefer their particular ‘optimal’ colour. The fact that this is not the case (and that yellow filters do not appear to be chosen more often than any other colour in other experiments) illustrates the confusion in this area.

One last study to be considered in this area is that of Christenson et al. (2001), which studied the effects of blue lenses on reading in a sample of individuals formally diagnosed with dyslexia using the Dyslexia Determination Test (Griffin & Walton, 1987). In a simple repeated-measures design, their results showed that the blue lenses had no significant effect on reading accuracy or comprehension, when compared to reading with no lens.

The study is, however, subject to many of the methodological concerns which have previously been noted for other studies. The sample size of 16 is very small, and the nature of the trial meant that no blinding for participants or experimenters was employed.

While the study is subject to these weaknesses, it does serve to illustrate an important point. Very few studies have taken into account the nature of the reading disorder of each individual, which, as Christenson et al. point out, may lead to 'a co-mingling of subjects who may have dyslexia and/or other causes of poor reading (e.g. vision problems and attention deficits)'. This obviously has the effect of making the studies less generalisable, but also may obscure any effects which the coloured filters may have on one particular subset of those with reading disorder. As we have seen above, the vast majority of those with reading problems appear to have dyslexia caused by a phonological deficit (Ramus, 2003), but other causes such as attentional deficits should certainly be taken into account. Future studies should at the very least report the number of individuals in their sample who have a formal diagnosis of dyslexia, and possibly those who are suspected to be poor readers due to an attentional problem or a visual deficit.

### *3.4 General Discussion*

Before concluding, this review will discuss the analyses and recommendations of Albon et al. (2008) after their systematic review of the literature on coloured filters, adding some further points about the status of the research in this area.

Albon et al. (2008) performed three meta-analyses focused on three outcomes of the randomized controlled trials included in the review. These are, to this author's knowledge, the first meta-analyses to be performed in this area. The first two meta-analyses, on reading accuracy (in which data from 3 studies were included) and reading speed (including data from 4 studies), showed that, while there may be a trend towards positive results for preferred coloured filters when compared to controls, the effect is not significant.

The third meta-analysis, on reading comprehension (which included data from 4 studies), found a slightly positive effect of preferred coloured filters over controls. Too few of the studies analysed subjective visual stress symptoms and their reduction by coloured filters for their inclusion in a meta-analysis.

The utility of meta-analyses is, however, limited when the methodological quality of the analysed studies is poor. As can be seen from the discussion above, the quality of much of the research into coloured filters is very low.

Albon et al. (2008) note various flaws in methodology in the reviewed studies. Importantly, they note that 'there are serious limitations in using the intervention under investigation to screen and include subjects for a study of effectiveness of that same intervention, since the study population has then been pre-sampled to show benefit from the intervention, which is likely to bias the results towards a positive effect.' (p. 82). This selection bias applies to the majority of randomized controlled trials in this area, including those often cited by proponents as evidence of the efficacy of coloured filters (e.g. Bouldoukian et al. 2002; Scott et al., 2002; Wilkins et al. 1994). It should also be noted that bias of this type confounds results by ignoring any improvements in reading that may have occurred in the pre-study period – individuals using overlays may attract more attention from parents or teachers to give them more reading instruction, or may simply have practiced reading more in this period due to a feeling of novelty caused the filters.

Related to this concern, it was noted at the start of this section that the definition of visual stress often varies from study to study, with some researchers assuming that visual stress is a response to coloured filters, and others asking participants for subjective measures of any reduction in visual stress symptoms. Studies that use the former, a 'response' definition (e.g. Kriss & Evans,

2005) risk confounding their results with the same problems as the aforementioned sampling bias.

Another problem pointed out by Albon et al. (2008) regards the sample populations in the studies. While the within-subjects (crossover) design of many of the studies is positive in that it means variation is less important, the vague definition of reading disorder given by most authors is concerning. While some participants were described as having a formal diagnosis of dyslexia, even within this definition there could potentially be considerable variation, again reducing the generalisability of results. The inclusion by future studies of baseline reading and reading age data would make interpretation of results considerably less difficult.

An intervention which necessarily involves 'optimal' versions of the treatment (in this case particular colours) for each participant is extremely difficult to control adequately, especially with a placebo. Albon et al. (2008) suggest that the best placebo would be 'a tint designed to be similar to the preferred colour, but lying outside the chromaticity shown to be effective in an earlier test of the intervention' (p. 81). However, it seems even more reasonable to design the experiment so that the participants are unaware of their 'optimal' colour (it is kept ambiguous by the individual involved in ascertaining the colour), so an unrelated placebo colour can be included in the experimental design without the participant automatically regarding it as inferior.

This discussion raises the question of diagnostic method. It is conceivable that some forms of diagnosis used in this area are more conducive to placebo effects than others. As discussed above, to reduce placebo effects, diagnosticians should avoid asking 'leading' questions of the participants. In addition, it is not implausible that the use of equipment such as the Wilkins Intuitive Colorimeter (Wilkins et al., 1992), a large, complicated-seeming machine, would lend the intervention an impression of 'importance' which the less dramatic process of selection from a set of coloured overlays would not produce. So far, no studies have been performed which compare different methods of treatment with coloured filters within one experiment. Such an experiment would be instructive in assessing the different placebo effects that particular forms of this intervention may produce.

While appropriate for testing the claims made by proponents such as Irlen (1991), performing trials measuring the immediate effects on reading of coloured filters may miss any longer-term effects the filters may produce. As noted in the discussion of Noble et al. (2004) above, longitudinal studies of a continually-developing variable such as reading skill are extremely difficult to design, involving as they do several other variables such as further interventions by parents and teachers which are near-impossible to completely control. However, it is not

impossible for more rigorous follow-up studies of any future randomized controlled trials to be performed in the future, and for great care to be taken to reduce confounding variables as much as is feasible.

Finally, Albon et al. (2008) note that reporting of the studies has in several instances been of low quality. Important information such as loss to follow-up and the method of calculation of some figures is often ambiguously reported (p. 47). Noted above is the concern that few studies report the colours chosen by their participants, thereby not allowing an appraisal of the colours chosen, which may be due to the nature of the selection process rather than a real 'optimal filter' effect. In future, researchers in this area should use the systematic review as a guide to designing and reporting studies.

Overall, the authors conclude that both the randomised controlled trials and other studies into the efficacy of coloured filters for reading disorder leave much to be desired, stating 'there was no convincing evidence to suggest that coloured filters can successfully improve reading ability in subjects with reading disability or dyslexia when compared to placebo, or other types of control' (p. 93).

### *3.5 Summary & Conclusions*

This section of the review began by asking two questions. Firstly, do coloured filters reduce visual stress symptoms beyond placebo? Secondly, do coloured filters improve reading skills beyond placebo?

For three reasons - problems discussed above with 'leading' questions, the fact that any visual stress symptoms are by their very nature subjective, and the fact that most studies have investigated reading disorder instead of focusing on visual stress symptoms - it is difficult to answer the first question. It is not safe to assume that because filters improve reading disorder, they must therefore have reduced visual stress symptoms or, indeed, vice versa. Studies which purport to show improvements in visual stress symptoms have had serious methodological problems, further obscuring an answer to the question.

In answer to the second question, although coloured filters are in use across the world as an alleged intervention for reading disorders, there exists no strong, unambiguous set of studies suggesting they are particularly efficacious, and the studies which do exist are bedevilled by both methodological concerns and poor reporting of results. It is therefore essential that more well-designed, well-reported randomised controlled trials are performed before coloured filters are recommended as a treatment for reading disability.

#### **4. THE PRESENT STUDY: A RANDOMISED CONTROLLED TRIAL**

After drawing the conclusions described in Section 3.5, we decided it would be very helpful to conduct a randomised controlled trial into the efficacy of coloured filters that, to the greatest extent possible, avoided the methodological problems of previous studies. This would contribute clearer and less ambiguous evidence either for or against the use of the filters.

For the study, we capitalised on an opportunity to integrate our experiments with the diagnostic procedure of an Irlen diagnostician active in the Inverclyde area of Scotland. The diagnostician is regularly active in various schools in the area, and we arranged to follow-up the diagnoses in Newark Primary School, Port Glasgow.

In this school, children who are noticed by their teachers to be below-average readers are as a matter of course recommended for Irlen assessment. If found to require coloured overlays or, later in the diagnostic process, coloured spectacle lenses, these overlays are provided free of charge to the children; the education department of the local authority, Inverclyde Council, finances the prescribed filters, as well as the costs of the Irlen diagnostic sessions (of which there may be several).

The Irlen practitioner agreed to, firstly, assess whether or not each child suffered from visual stress, secondly, prescribe a particular colour of overlays for each child, and thirdly – and most importantly – to ensure that each child was not aware of their diagnosis, their prescribed colour, and, to the greatest extent possible, the claimed effects of the filters. We believe the participant blinding is a notable improvement the present study made over previous methodologies.

All children assessed by the Irlen practitioner were measured on these three tests by the experimenter, who was also blind to the child's diagnostic status and prescribed colour. Firstly, all children were assessed using the Mini-Mental State Examination (MMSE; Folstein et al., 1975). This test was originally designed to assess the mental functioning of patients with age-related impairments such as dementia and is not ordinarily used in children. It has, however, been shown to correlate significantly with chronological age, mental age and reading age in children (Ouvrier et al., 1993). In addition to this, the test is extremely brief compared with other tests of mental functioning such as full IQ tests; it can be completed in less than 5 minutes. Thus, it was chosen for inclusion in the present study as a control measure for cognitive ability, and as a matching measure for the randomised controlled trial, described below.

The second measure used was the Wilkins Rate of Reading Test (WRRT; Wilkins et al., 1996), which was discussed above in the context of past experiments. It is specifically designed to test

immediate effects of coloured filters on reading rate, and deliberately omits any element of comprehension. Several changes were made to the procedure of the WRRT for the purposes of the present study; these are discussed in Section 4.1.2 below.

This part of the experiment was run as a crossover design, with all children given the WRRT in three conditions – their prescribed colour of overlay (or, if they did not have visual stress, a randomly-selected colour), a placebo colour, and a colourless overlay. The inclusion of three types of overlay allowed more informative comparisons to be made – is there, for example, a general effect of colour but no specificity? Or is there no difference between the colours?

We predicted that, if the overlays have the effect claimed of them by proponents, the double-blinding employed in this study should not reduce their alleged reading facilitation effects. In this case, we would expect to see the highest reading rate in the condition with the prescribed overlay than the colourless overlay condition. However, if the overlays work mainly by placebo effect, the blinding should ensure that the reading rate of our participants is not significantly greater when the prescribed overlay is used. In both cases, the effects of the placebo overlay are uncertain – it may be the case that any overlay will facilitate reading to some extent, but an incorrect colour may also, according to Irlen (1991), cause further problems with reading.

We also included a measure of general reading ability, the Gray Oral Reading Test. (GORT; Wiederholt & Bryant, 2001). This test gives results of reading fluency, reading comprehension, and a combined age-standardised Oral Reading Quotient (ORQ). This allowed us to investigate any immediate effects of the coloured filters on general reading ability.

A between-subjects randomised controlled trial design was used for this part of the study. The children were randomised into two groups, which could be matched on measures such as age and MMSE score. One group was tested on the GORT using only their prescribed overlays, while the other was tested using colourless overlays.

Again, our predictions were that, if the overlays were effective, fluency, comprehension and ORQ would be higher in the randomised group in which coloured overlays were used. If the overlays were ineffective, no differences between the two groups would be found.

On the recommendations of authors such as Blaskey et al., (1990) and Scheiman (2004), we included an orthoptic assessment for all our participants. Each child was assessed by a trained orthoptist on a battery of measures. As noted above, the evidence on orthoptic correlates is somewhat confused. We therefore did not make any specific predictions about orthoptic results

and how they would relate to chosen colour, visual stress diagnosis (see, e.g. Scott et al., 2002), or any other of the study's measures.

## 4.1 METHOD

### 4.1.1 Participants

Participants were a sample of 75 children who had been referred for Irlen diagnosis by their class teacher, on the basis of their being perceived to be below-average readers. Children were not included in the sample if they had a formal diagnosis of Autistic Spectrum Disorder. Parent/guardian's consent was sought and received for all participants.

13 children were deemed unsuitable for diagnosis by the Irlen practitioner, and were therefore removed from the study. 1 child left the school before testing could be completed. After attrition, therefore, 61 children remained (39 male, 22 female), aged 7-12 years (mean = 9.54 years, SD = 1.19).

### 4.1.2 Materials

*Irlen Coloured Overlays* are available in 10 colours<sup>2</sup>, with both 'matte' and 'shiny' sides, and were used in conjunction with the reading tests as detailed below. The overlays were provided by the Irlen diagnostician, and come with the Irlen name and other details written along the top and bottom edges. These edges were removed prior to the study.

The 'colourless overlay' condition was created by placing a clear cellulose acetate sheet over the text to be read (both sides of this sheet were 'shiny').

The *Mini Mental State Examination* (MMSE; Folstein et al., 1975) includes questions assessing orientation to time and place, registration of objects, attention, and language. The maximum MMSE score attainable is 30.

The 'serial 7s' task of the MMSE was thought to be overly difficult for the ability range of the sample, and so was replaced by the task in which the participant spells the word 'WORLD' forwards and backwards, as recommended in the MMSE instruction manual.

It is possible that this word-based task may not have been appropriate for a sample of children with reading problems. However, we believe the 'Serial 7s' task would in all likelihood have been performed even more poorly.

---

<sup>2</sup> Aqua, Goldenrod, Grey, Blue-grey, Rose, Peach, Yellow, Green, Turquoise, Purple.

The *Wilkins Rate of Reading Test* (WRRT; Wilkins et al., 1996) consists of lists of 15 familiar words<sup>3</sup> repeated 10 times each. These lists are arranged in 10 lines, each with a different random word order, in closely-spaced type and a small font size, designed to visually stress the reader. The non-sentential nature of the lines removes any comprehension aspect from the test, instead assessing more basic visual abilities.

The design of the present study necessitated the use of a somewhat modified WRRT. Due to the young age and poor reading ability of the participants involved, the sheets used in the present study had a larger font size (Helvetica, 16pt) than in the original test, as advised by Wilkins (personal communication).

Participants were first given a 'practice' sheet, which they read for 30 seconds, and which was not scored. They then read 6 variations (Versions A-F) should perhaps mention that only 4 versions are provided with official test, and that two additional orders were created to allow a unique version for each administration of the test of the test, two with each of the three overlay conditions - prescribed overlay, placebo overlay, and colourless overlay - in the format ABCCBA to control for practice and fatigue effects (see also 'randomisation', below). Participants read these 6 sheets aloud for one minute each. Deviations - incorrect words, mispronounced words and skipped words or lines - were noted, resulting in a score of number of words correctly read in one minute. These scores were then averaged for each of the three overlay conditions.

The *Gray Oral Reading Test - 4<sup>th</sup> Edition* (GORT; Weiderholt & Bryant, 2001) is a full test of reading ability, standardized on 6,000 children. Participants read a sequence of stories aloud, and answer 5 comprehension questions on each. The experimenter reads out the questions, and the participant reads along and gives answers. Stories are read until a 'ceiling' score is reached, at which point the test is complete.

After noting questions answered correctly, deviations from the text and the time taken to read each story, age-standardised 'fluency' and 'comprehension' scores are calculated. These are summed to produce an 'Oral Reading Quotient' (ORQ) for each participant.

In the present study, the GORT was performed with either a coloured overlay or the colourless overlay covering all the stories and the questions - see 'randomisation', below for details of how this was arranged.

---

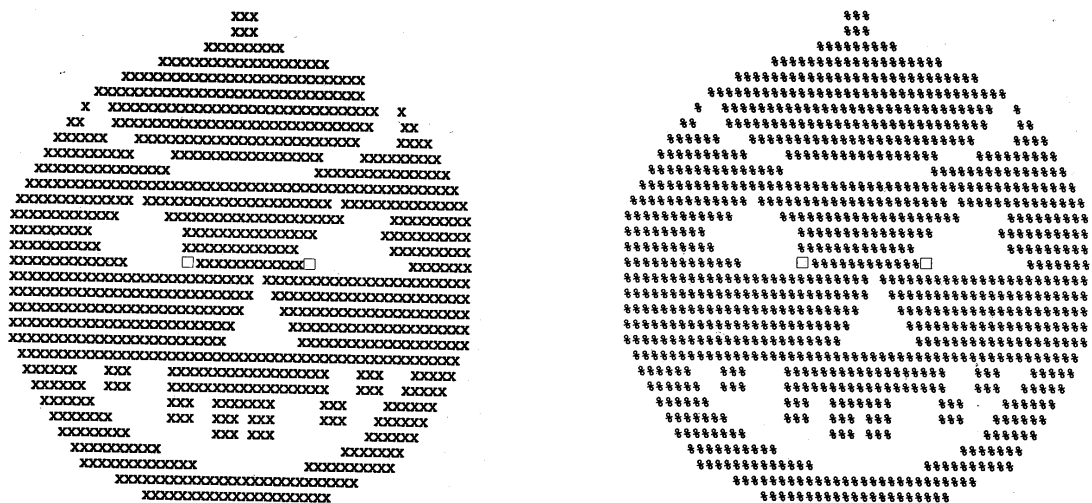
<sup>3</sup> My, play, see, to, for, you, is, the, look, up, cat, not, dog, come, and.

### 4.1.3 Irlen Screening

To assess each participant's visual stress status, the Irlen diagnostician used the Irlen Differential Perceptual Schedule (IDPS; Irlen, 1987). As discussed above, the Irlen Institute literature calls the disorder 'Irlen Syndrome', but we will continue to use the term 'visual stress', as before. The IDPS is in four parts:

1. Participants are asked a series of questions regarding how often they encounter adverse effects (visual distortions, light sensitivity, headaches, eyestrain, tiredness, etc.) while reading.
2. Participants are shown a series of pictures made up of lines and symbols (Figure 1). They perform tasks involving counting the numbers of lines and symbols in specific parts of the pictures. As these tasks are being performed, participants are asked to report any visual distortions (blurring, movement, flickering, glowing etc.) that they experience, and their level of comfort when looking at the page.

Figure 1. Visual stress diagnostic tasks from the IDPS (Irlen, 1987, Pictures A and B). Participants count the symbols (either the letter 'X' or '%' signs) between the central markers. Each picture appears on a separate 21.5cm x 21.5cm page.



3. Participants are shown a page of text in Dutch (a language foreign to them). One half of the page is covered with a Yellow overlay, and the other with a Turquoise overlay. The participant is asked which of the colours, if either, makes the text more comfortable to look at. The less comfortable colour is removed and replaced with another of the overlays (the order of subsequent overlays is not fixed), and the participant is asked which of these is more comfortable. This procedure is continued in a process of elimination until an optimally-coloured overlay is selected.

4. Participants read a section of prose in English (their own language) with and without the selected overlay. They are again asked about their visual symptoms.

Ordinarily, after the screening session, participants would be made aware of their optimal colour, given an overlay for immediate classroom reading use, and instructed in how to use it. However, due to the requirement for blinding in the current study, the participants were not informed which colour was their optimal, and were instead informed that their diagnosis was not yet complete. The details of diagnosis, both the presence of otherwise of visual stress and the optimal colour (if visual stress was diagnosed) were then passed on to the supervisor for randomization.

#### *4.1.4 Randomisation*

The randomisation was performed by the supervisor, who was informed by email of each child's study code, visual stress diagnosis (i.e. visual stress or normal) and prescribed overlay (if visual stress). Children in the normal group were each assigned a nominal 'prescribed' colour, selected to mirror, approximately, the frequency of the different prescribed colours in the group of children diagnosed with visual stress. For each prescribed colour, the placebo colour was determined by pre-fixed pairings of the ten overlays. Where possible, each colour had been assigned a partner from the complementary part of the spectrum. The pairings were: Aqua/Goldenrod, Yellow/Blue-grey, Green/Rose, Peach/Turquoise, Grey/Purple.

The trial order list was prepared such that, within each sequential series of 12 participants, the six possible trial orders for the WRRT (six A-B-C orderings of prescribed colour, control colour and colourless) were counterbalanced against the two possible overlay assignments for the GORT (prescribed colour, colourless). Identical trial order lists were used for the Visual Stress and Control groups, and each child was assigned the next available trial order for their group.

For each study code, the supervisor specified the trial order for the WRRT in an email to the experimenter. The overlay to be used for the GORT was specified in a separate email, which was not opened by the experimenter until immediately before the GORT was administered, and after the WRRT had been completed. Details of the trial orders were known only to the supervisor. The experimenter was thus blind to the child's Irlen status and to the prescribed overlay colour for all participants.

#### *4.1.5 Procedure*

The experimental procedure for each participant was carried out on a separate day, at least 2 days after their Irlen diagnosis. Participants sat at a table in a quiet room in the school, with

fluorescent lights switched on as well as daylight from a window, as advised by the Irlen diagnostician. All the children were tested by the same experimenter. Participants first completed the MMSE in the standard manner described above.

Participants then completed the WRRT practice sheet, followed immediately by the 6 versions of the test. Participants were informed when they had reached the halfway point.

The GORT, which as stated above lasts as long as it takes for the participant to reach a 'ceiling' level, was then administered.

Throughout the procedure, no reference was made to the claimed effects of the overlays, and any questions asked by the children on this subject were redirected. The overlays were simply placed over the text before each child started reading, with no explanation given or attention drawn to them. The overlays were thus imbued with as little significance as possible.

The entire experimental procedure took approximately 40 minutes, after which the participant returned to their class.

#### *4.1.6 Orthoptic Assessment*

On another separate day, each participant was assessed by a qualified orthoptist from Glasgow Caledonian University's Eye Clinic.

The following ocular functions were assessed: uniocular near and distance visual acuity, ocular motility, convergence, accommodation, fusional reserves, and saccades. A cover test (to detect the presence of squint) was also performed.

Each orthoptic assessment lasted no more than 30 minutes. The orthoptist performing the assessment was blind to each child's Irlen diagnosis, prescribed overlay colour, and performance in the reading assessments.

## 4.2 Results

### 4.2.1 Group Characteristics

Of the sample of 61 participants included in the study, 47 (77%) were diagnosed with Irlen Syndrome (this group will be referred to as the 'Visual Stress' group, with the remainder referred to as the 'Control' group). The frequencies of overlay colours prescribed for each of the 47 are detailed in Figure 1, while Figure 2 details the 'prescribed' colours allocated in the randomization process to the Control group. All overlays were used on their 'matte' (as opposed to 'shiny') side.

Figure 2: Prescribed colours, Visual Stress group

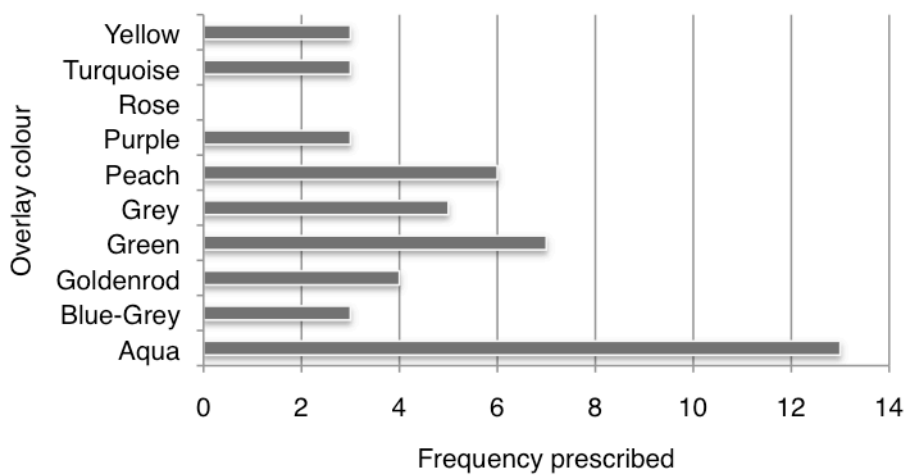
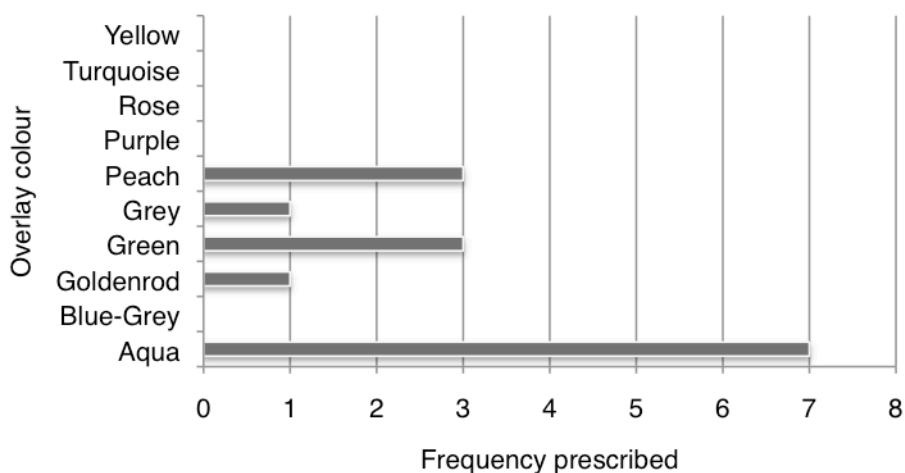


Figure 3: Randomised 'prescribed' overlays, Control group



A chi-square analysis for the Visual Stress group showed that the selection of colours chosen was significantly different from chance ( $\chi^2 (9, n= 47) = 5.2, p = .005$ ). It can be seen from Figure 1 that Aqua, at 13 prescriptions, was significantly over-represented and Rose, at zero prescriptions, was significantly under-represented.

59 of the 61 participants were assessed by the orthoptist - due to long-term absence from school, participant 5 (Control group) and participant 49 (Visual Stress group) were not assessed.

Table 1 compares the characteristics of the Visual Stress group and the Control group: mean age, frequency of each sex, mean MMSE score, percentage of each group where any orthoptic tests were failed (a broad measure of those who were found to have orthoptic problems), and the mean percentage of orthoptic tests failed.

Table 1: Participant characteristics

Group	Visual Stress	Control
Mean age	9.6 years	9.2 years
Sex frequency	28 male, 19 female	11 male, 3 female
Mean MMSE score	23.94	24.79
% in which any orthoptic tests failed	54.30%	38.50%
Mean % orthoptic problems	14.87%	11.40%

Mann-Whitney U tests showed no significant differences in the percentage of individuals in which orthoptic problems ( $U = 251.50, n_1 = 46, n_2 = 13, p = .32$ ) or for the mean percentage of orthoptic problems ( $U = 262.5, n_1 = 46, n_2 = 13, p = .48$ ) were found between the Visual Stress and Control groups. The Visual Stress group, then, did not suffer from more detectable orthoptic problems than the Control group.

Mean number of orthoptic tests failed did, however, correlate negatively with baseline reading rate (score on the 'colourless overlay' condition of the WRRT) at a level which reached significance,  $r(59) = -.27, p = .04$ . More orthoptic problems were, then, associated with slower reading.

#### 4.2.2 Individual orthoptic tests

Table 2 compares the percentages of the Visual Stress group and the Control group who failed each individual orthoptic test ('cover' refers to the 'cover test', to detect the presence of squint).

Table 2: Individual orthoptic tests failed

Test	Near acuity	Distance acuity	Cover	Ocular motility	Convergence	Accommodation	Fusional reserves	Saccades
% VS group failed	21.74%	34.78%	4.35%	6.52%	4.35%	41.30%	2.27%	2.17%
% Control group failed	15.38%	15.38%	7.69%	0.00%	15.38%	30.77%	0.00%	0.00%

Since visual stress sufferers had been found to have differences in accommodation in previous experiments (e.g. Scott et al., 2002; Simmers et al., 2001), an independent-samples t-test was used to test for differences in accommodation between the two groups. The difference was not found to be significant ( $t(57) = -.68, p = .50$ ). Our two groups did not, then, differ in ocular accommodative ability. However, this result must be viewed with caution as, firstly, the groups differed so widely in size (Visual Stress group  $n = 46$ ; Control group  $n = 13$ ) and because it is uncorrected for the number of comparisons in orthoptic results made between the groups.

#### 4.2.3 MMSE Scores

Means and standard deviations of MMSE scores for all participants, and mean scores by group are shown in Table 3.

Table 3: MMSE scores

Group	Mean Score	MMSE	Std. Deviation
All ( $n = 61$ )	24.13		3.82
Visual Stress ( $n = 47$ )	23.94		3.96
Control ( $n = 14$ )	24.79		3.33

An independent samples t-test showed that there was no significant difference between the visual stress and control groups on MMSE score,  $t(59) = -.73, p = .47$ .

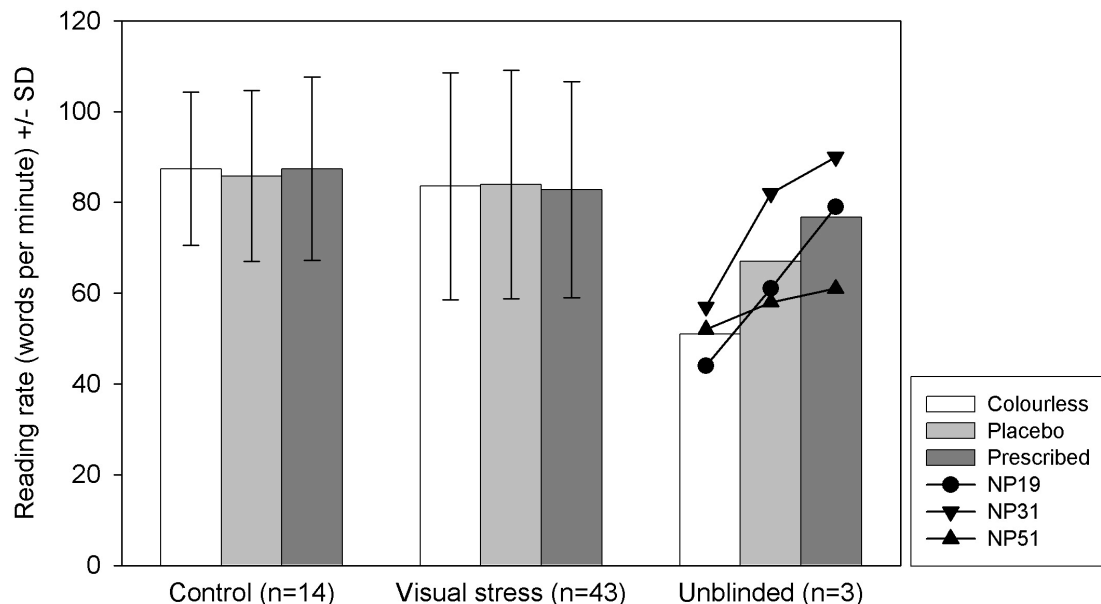
As expected (on the basis of Ouvrier et al. 1993), the scores on the MMSE were significantly correlated with chronological age,  $r(59) = .36, p = .002$ . MMSE scores were also significantly correlated with baseline WRRT scores (those in the 'colourless overlay' condition) for all participants,  $r(58) = .40, p = .002$ , and with GORT ORQ averaged across the two experimental conditions,  $r(58) = .41, p = .001$ . This suggested the MMSE could be useful for matching reading ability in the randomized controlled trial (see GORT results, below).

#### 4.2.4 WRRT – Crossover Study

60 of the 61 participants successfully completed the WRRT (participant 64 (Visual Stress group)'s WRRT performance was seemingly random). It became obvious during testing that 3 of the children (participant codes NP19, NP31 and NP51; all in the Visual Stress group) knew their filter colour, and in two of these cases, had been using their filter for some time prior to the experiment. This compromised both participant and experimenter blinding; data from these three children was separated from the main analysis as detailed below.

Figure 4 illustrates the mean WRRT scores (with standard deviations) for the Visual Stress group ( $n = 43$ ), the Control group ( $n = 14$ ) for each experimental condition (colourless overlay, placebo overlay, and prescribed overlay). The means, as well as individual scores, for the three unblinded participants are also included, for illustrative purposes.

Figure 4. WRRT reading rate for Control group, Visual Stress Group, and unblinded participants.



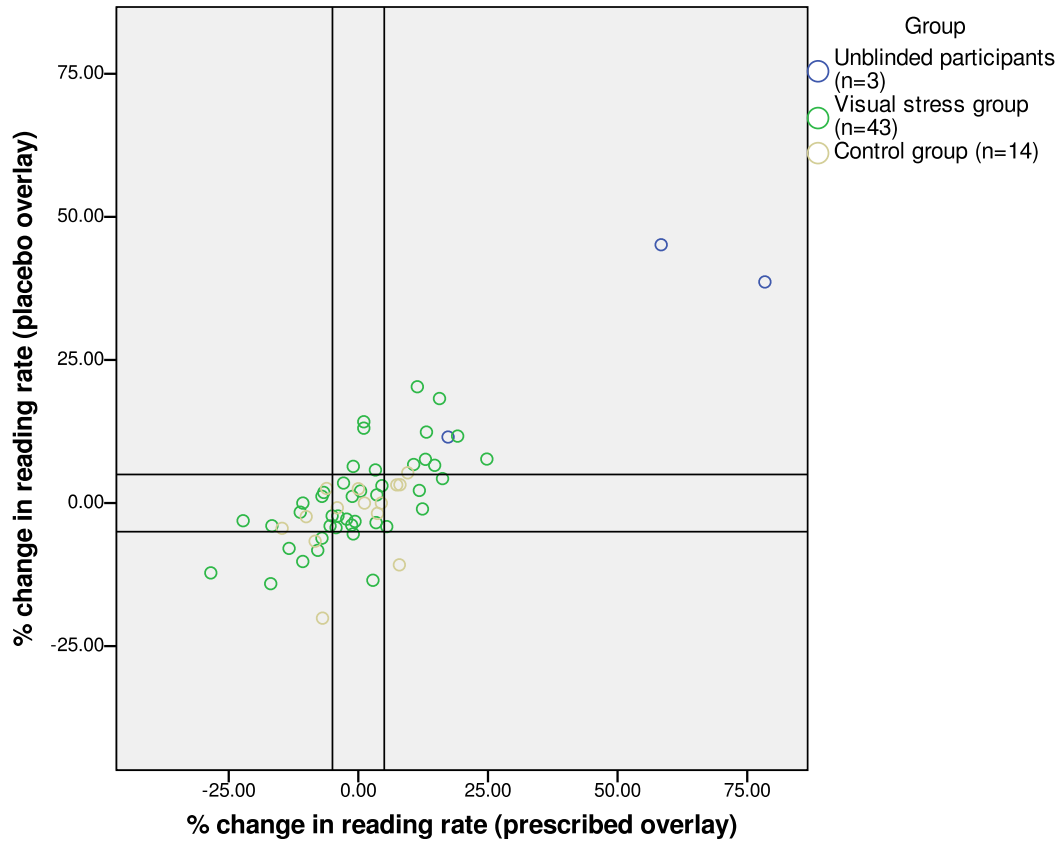
For the Visual Stress and Control groups, excluding the Visual Stress-unblinded group (total  $n = 57$ ), a two-way repeated measures ANOVA was used to test between the three experimental conditions. With sphericity assumed ( $c^2(2) = 4.35, p = .11$ ), this test showed no significant effect of overlay condition  $F(2, 110) = .16, p = .85$ . No significant interaction effect was found for WRRT score and group,  $F(2, 110) = .90, p = .40$ . These results suggest that neither prescribed nor placebo overlays significantly facilitated reading rate when compared to a colourless overlay.

Although the mean ‘baseline’ reading rate, that is, the reading rate with the colourless overlay, was lower in the Visual Stress group (incorporating those in the Visual Stress-unblinded group) than the Control group, this difference was not reliable,  $t(59) = .98, p = .33$ . The two groups were of similar reading ability, even though one had been diagnosed with visual stress.

Figure 5 overleaf illustrates the percentage increase in reading rate on the WRRT for all participants, for both the prescribed overlay (x-axis) and the placebo overlay (y-axis) condition when compared to the colourless overlay condition (the baseline reading rate). Since a typical criterion for diagnosing visual stress in the Intuitive method is an increase in reading rate of 5% (Evans & Joseph, 2002; Kriss and Evans, 2005; Wilkins et al., 1996), lines are added to both axes to denote a 5% increase in reading rate. Also added are lines to show a 5% decrease in rate, since

many of the participants experienced this when using the overlays. These results indicate that the clinical significance of both the placebo and prescribed overlays is low.

Figure 5: Percentage change in WRRT reading rate (from baseline) for placebo and prescribed overlays



#### 4.2.5. GORT – Randomised Controlled Trial (RCT)

60 of the 61 participants successfully completed the GORT (participant 19 (Irlen group) left the test room after only partial GORT completion).

The mean ORQ of these 60 children was 80.00, s.d. = 13.64. The fact that the standardized mean ORQ is 100 emphasizes the poor reading ability of our sample.

The randomization method described above arranged the children into two RCT groups, which were matched as described in Table 4.

Table 4: GORT RCT matched group characteristics

Group	<i>n</i>	Sex Freq.	No. in vis. stress group	Mean Age	Mean MMSE score	Mean WRRT (colourless) score
Prescribed Colour	31	21 male, 17 female	23	9.58 years	24.16	81.10
Colourless	29	17 male, 12 female	23	9.48 years	24.21	83.09

Independent samples t-tests showed no significant differences in age ( $t(58) = -.31, p = .76$ ), MMSE score ( $t(58) = .05, p = .96$ ), or baseline WRRT score ( $t(58) = .30, p = .77$ ) between the two groups, underlining the well-matched nature of the groups.

Table 5 shows the mean scores for GORT Fluency, Comprehension, and ORQ for each group.

Table 5: Mean GORT scores for RCT groups

Group	Fluency	Comprehension	ORQ
Prescribed Colour	5.16	8.00	79.48
Colourless	5.52	7.93	80.55

Independent samples t-tests showed no significant differences between the two groups on Fluency ( $t(58) = .47, p = .64$ ), Comprehension ( $t(58) = -.11, p = .91$ ), or ORQ ( $t(58) = .30, p = .77$ ). The prescribed filters, then, did not significantly facilitate reading on any of the GORT measures.

## 5. DISCUSSION

### 5.1 The Present Study

The study described here was a rigorously-designed randomized controlled trial of the efficacy of Irlen coloured overlays for treating reading disability. Importantly, it involved a type of participant blinding which was unusual for studies of this nature: the children were unaware of their chosen overlay colour, their visual stress diagnostic status, and were to the greatest degree possible unaware of the supposed effects of the coloured overlays. This, combined with the experimenter blinding, meant that any placebo or expectation effects were minimized.

In the sample of children who were - to the best of our knowledge - blinded in this manner, our results clearly showed no effect of either placebo overlay or coloured overlay on the Wilkins Rate of Reading Test.

Though the sample size of children who were *not* blind to their visual stress diagnosis was extremely small ( $n = 3$ ), it is striking to see the difference in WRRT scores observed. These children did show a clear difference in overlay condition – exactly the kind, and extent, of difference that the theories of Irlen (2010), Wilkins (2003) and others would predict<sup>4</sup> to be found for all the participants diagnosed with visual stress (Figure 4). This result underscores the need to have careful participant blinding in studies of this nature.

Although the wide difference in group size hampers a fully accurate comparison, we can attempt to contrast the ‘visual stress’ group with the ‘control’ group. No effect of coloured overlays on WRRT scores was found, but are there factors which may explain this? Certainly, the groups did not differ on MMSE score, though this is not a fine-grained measure of IQ. Testing intelligence with a full IQ test would have been preferable, but would have resulted in a testing time of over one hour for each child, likely resulting in fatigue and loss of attention. The fact that no reliable differences were found between the groups on the various orthoptic tests implies that either there were no differences in the visual properties of the two groups, or that the orthoptic test was not sensitive or specialised enough to discern visual stress (see Section 5.2 for further discussion of the orthoptic results).

As would be expected since MMSE score correlated with baseline WRRT score (that in the ‘colourless overlay’ condition) and there were no between-group differences on MMSE score, the visual stress group did not have lower baseline scores on the WRRT than the control group. This

---

<sup>4</sup> These theorists would have predicted the large difference between the ‘colourless’ and the ‘prescribed’ overlay conditions. However, as discussed below, the ‘placebo’ overlay may have been predicted to benefit to a small extent (as was found), or to interfere further with reading ability.

finding tells us only that the Irlen diagnostician was not diagnosing visual stress on the basis of reading ability.

In the end, no reliable differences could be found between the Visual Stress group and the Control group, on a wide range of characteristics or on the effects of the intervention. This is obviously contradictory to the view of those who suggest that visual stress is a reliable construct, the symptoms of which can often be alleviated by coloured overlays. As mentioned above, the caveat to this conclusion that the Control group in our study was quite small compared to the Visual Stress group. More equal groups would potentially have allowed for differences to be uncovered.

Table 4 and the accompanying statistical tests above show that the two randomized groups (those given a coloured overlay and those given a colourless overlay) for the GORT were very well matched on a number of measures such as MMSE score and baseline reading rate on the WRRT. They also included an equal number of children diagnosed with Irlen syndrome. That we found no differences on any of the three GORT measures is more evidence that, under double-blind conditions, the coloured overlays do not have an effect. It has been suggested that not only reading rate, but reading comprehension can improve when coloured filters are worn (e.g. Wilkins et al., 1994, p. 359; Sandra Irlen, personal communication, October 2009). Our results are inconsistent with this claim.

Although considerable effort was expended to minimize placebo effects in this study, it is highly unlikely that they were removed completely. The novelty and possible subjective attractiveness of the overlays, which none of the children included in the analysis had seen previously, would almost certainly have contributed to increased concentration and attention from at least some of the children to the reading task when they were being used. This makes the lack of significant differences observed all the more conspicuous; even the residual placebo effect, which would have been impossible to remove, did not affect the children's reading performance. In addition, the lack of positive results confirms that the blinding must have been sufficient; no significant effects, placebo or otherwise, were found.

We believe our sample size was sufficient to test the claims of the Irlen Institute (2010) which claims that the Irlen method benefits 46% of individuals with 'reading problems, dyslexia and learning difficulties' (Irlen Institute, 2010). While we were not aware of the numbers within our sample who had a formal diagnosis of dyslexia, we believe our sample of children with reading problems would have included a substantial number of children with such a diagnosis. Indeed, the fact that the mean GORT Oral Reading Quotient score in our sample was substantially lower than the standardized average gives us reason to characterize our sample as reading disabled. In

any case, 77% of the sample were diagnosed with visual stress. That no significant differences in reading with and without the overlay were found even in this 77% is striking.

The colour specificity which is discussed by Wilkins (2003) and Irlen (2010) was not supported by our research. It is claimed by these authors that one specific colour and no others will facilitate each individual's reading. Indeed, the Irlen Institute claims that colours which differ from this 'optimal' chromaticity may cause additional reading problems (Irlen, 1991). In our unblinded children, we found that the placebo overlays – which were spectrally very different from the prescribed colour - benefitted reading more than the colourless baseline condition. We would suggest that, given a lack of blinding, the expectation effects associated with the presence of any coloured filter will increase an individual's attention and concentration, and lead to improved reading.

Finally, our analysis of the prescribed colours in our sample may at first glance support a prediction of Stein (personal communication), once which has seen recent media publicity, that 25% of dyslexic children should benefit from blue filters. Aqua, the deepest blue available in the Irlen colour system, was prescribed significantly more often than expected by chance. However, the other part of Stein's theory, that another 25% of dyslexics should benefit from yellow filters (see, e.g. Ray et al., 2005, though in that study the entire sample preferred yellow lenses), was not supported by our research. Yellow was not prescribed significantly more often than would be expected by chance.

Not only is there a lack of statistical significance in our findings, but as shown in Figure 5, the overlays failed to cause clinically significant increases in our participants' reading rate. The diagnostic criterion for visual stress, discussed in section 3.1 above, of a 5% increase in WRRT reading rate (e.g. Wilkins et al., 1996), would suggest that, putting aside the Irlen diagnosis, 19 children from our sample had visual stress, as their reading rates increased over 5% with the prescribed overlay. However, the same number of children had a 5% decrease in reading rate with their prescribed overlay. Slightly more children – 16 versus 11 - had an over 5% increase than an over 5% decrease in reading rate with the placebo overlay. Figure 5 clearly demonstrates that the increases and decreases in reading are clustered around zero, with only two children – both of whom were unblinded - showing dramatic differences. The clinical significance of the overlays is patently very small. Interestingly, previous papers (Evans & Joseph, 2002; Kriss & Evans, 2005; Wilkins et al., 1996) have reported the number of children in their sample whose reading increases by over 5%, but not those whose reading decreases by the same margin.

It remains possible that, in a small subset of reading-disordered individuals, coloured filters may make a clinically significant improvement to reading, and this subset is too small for randomised

controlled trials of the nature described here to detect. However, many of our sample were diagnosed with visual stress, so the prediction made by proponents of coloured filters would have been that effects would be found. Why, then, might previous studies have found reliable effects in their samples? We would suggest that the sampling bias, discussed in Section 3.4 above, may be the reason. By selecting for inclusion in the studies only individuals who had previously been using coloured lenses, the authors biased their results towards the positive (Albon et al., 2008).

At the very least, the results of our study suggest that the current widespread diagnostic sessions for visual stress and provision of coloured filters in schools are in many cases inappropriate, and other avenues – those with better evidence – for ameliorating reading disabilities should be considered (see Section 5.4).

#### *5.1.1 Limitations and concerns*

Naturally, since we investigated only the Irlen system of coloured filters, our results are not fully generalisable to other systems. However, while the diagnosis and method to choose a colour differ between the Intuitive system (Wilkins, 1994; 2003) and the Irlen system (Irlen, 1991; 2010), and the methods of tinting lenses may be different, the nature of the coloured overlays available from both perspectives does not differ to a large degree. As noted above, the appearance of the overlay is the same (one ‘matte’ side, one ‘shiny’) and the colours are broadly similar, both with ten colours, with only the colours mint green, blue, and pink in the Intuitive system replacing turquoise, blue-grey and peach in the Irlen system (though it does appear that the Intuitive colours are more often used in combination than the Irlen overlays – in the study of Scott et al., 2002, many participants were prescribed combinations of colours, whereas in our study, all participants had a single colour).

Related to the first point, it must also be taken into account that our results are necessarily reliant on the expertise of one individual, the Irlen diagnostician. Since Irlen practitioners use purely subjective measures to diagnose visual stress (and do not use equipment such as the Intuitive Colorimeter), we were required to base our studies entirely on her testimony. If the diagnostician incorrectly diagnosed some children, our results would have been impacted commensurately. We have, however, no reason to think that the diagnostician’s efforts in our study were unusual in any way: as stated previously, we followed up on the diagnostician’s regular work with the Inverclyde children.

As noted by Albon et al. (2008), testing for the immediate effects of coloured filters is not a full investigation of their efficacy. It is possible that the filters remove a barrier to good reading by

ameliorating visual stress symptoms, and allowing normal reading skills to develop over time. For this reason, we intend to carry out a follow-up test at one year after the overlays were issued. This test will again involve the GORT, which is age-standardised, allowing developmental comparisons to be made. If the children who still use their overlays after one year have, at that point, made more reading progress than we would expect on the basis of their previous reading scores, we would have tentative evidence that the overlays are useful in the long-term. However, it must be said that, given the theories of visual stress claim an immediate effect of the filters will be found in most sufferers of the condition, and we failed to find this in our study, we would suggest that a longer-term reading improvement would be unlikely.

In our study, the WRRT was performed in a modified version, leaving our results open to the challenge that the larger typeset did not adequately visually stress the readers. After all, the small text of the original WRRT is deliberately designed to be very difficult for visual stress sufferers to perceive.

In response, we would point out that such changes were recommended for the age group of our sample by the inventor of the test (Wilkins, personal communication, January 2010), and the standard WRRT test materials include a version in larger type, that is aimed at young children who are used to have appreciable difficulty with the small type version. Certainly, very many of the children in our sample found the task very difficult, and read at a very slow rate. Had we given these children the standard WRRT, their results would most likely have been poor to the point of being uninterpretable. Finally, given the treatment is recommended for classroom use, we should surely have observed some reliable difference in reading rate when using the overlays at any type size.

Finally, one last minor concern with our study, which we do not believe would have influenced our results to any great extent, is nevertheless worth mentioning. While Irlen overlays are often prescribed to be used on particular sides – ‘shiny’ or ‘matte’, the Irlen diagnostician we worked with only prescribed overlays on their matte side. However, the acetate sheet used for the ‘colourless overlay’ condition on the WRRT was shiny on both sides, and therefore could only be used in this form. While we do not regard this as overly concerning, it is conceivable that a shiny overlay would create more glare than a matte overlay. However, if the reflective nature of the colourless overlay impeded reading, we would have expected to see results indicating a positive effect of the overlays when compared to it. This was not the case.

## 5.2. *The visual stress diagnosis*

The existence of visual stress as a diagnostic entity has not been confirmed or clarified by the results of our study. Indeed, we did not find any evidence of the disorder either in the reading test results or the orthoptic examination.

Unlike previous studies, such as Simmers et al. (2001) and Scott et al. (2002), we did not find any reliable orthoptic correlates of visual stress. Simmers et al. (2001) found increased accommodative microfluctuations in individuals who reported visual stress symptoms when coloured lenses were not worn. Our orthoptic examination did not specifically test for this, but no significant accommodative differences were found in the visual stress group.

This lack of accommodative differences between groups is also inconsistent with the findings of Scott et al. (2002), who showed an association between monocular accommodative ability and preferred colour. Our study failed to replicate these results, which is in line with the findings of Ciuffreda et al. (1997), who found no accommodative differences when their participants wore Irlen lenses. The differences may be at least partially explained by the fact that our study, and that of Ciuffreda et al. (1997) used Irlen diagnosis and Irlen lenses, whereas Simmers et al. (2001) and Scott et al. (2002) used the Wilkins Intuitive system (Wilkins, 1994). However, our study and that of Scott et al. (2002) have more in common as they used coloured overlays (as opposed to lenses), on the nature of which the Irlen and Wilkins systems differ less, as noted above.

Another limitation of our study is that we did not include a subjective measure of visual stress symptoms (the Irlen diagnostician collected this information for each child upon diagnosis, but we did not have access to it). If, as in the study of Wilkins et al. (1994), reading skill did not change but subjective visual stress symptoms were ameliorated, the overlays may well have had some utility. However, it is the case that visual stress symptoms and poor reading are linked, theoretically, in all characterizations of the disorder (e.g. Irlen, 2010; Wilkins, 2003). If a treatment claimed to reduce these symptoms does not bring about any significant improvement in reading, one has reason to doubt that the individuals studied do, in fact, have the disorder.

To conclude this section, it must be pointed out that though the evidence for visual stress as a discrete entity is equivocal, this does not imply that optometric or orthoptic problems cannot interfere with reading. Indeed, our study found a significant negative correlation between the number of orthoptic tests failed and performance on the baseline measure of the WRRT. Could it therefore be possible that the symptoms claimed to be due to 'visual stress' are in actual fact due to other visual problems?

It is, after all, intuitively obvious that more visual difficulties would have a detrimental effect on reading. Howell & Stanley (1988) were of the opinion that 'the visual symptoms claimed by Irlen [and, by extension, other visual stress theorists] can be divided into two groups; those associated with the symptoms of abnormal glare sensitivity, and those which would appear to have an oculomotor basis' (p. 70). They go on to point out that optometrists can prescribe tinted 'sunglass' lenses (which are not required to be of any particular colour) to alleviate glare sensitivity, and orthoptists have methods of treating oculomotor disorders. Scheiman (2004) reminds us that the only study which has directly *compared* coloured filters to treatment is that of Blaskey et al. (1990), discussed in detail above. In that experiment, individuals who received 13 weeks of vision therapy for their ocular disorders were found to have significantly lowered their score on a test used by Irlen practitioners to test for visual stress, despite the fact they had not been exposed to coloured filters. Studies such as that of Scott et al. (2002), as previously discussed, concluded that oculomotor disorders are merely a correlate, not a cause, of visual stress, but it would nonetheless be useful if more studies of the same style as Blaskey et al. (1990) would be carried out, to clarify this important issue.

### ***5.3 Directions for further research***

Further trials of coloured filters for reading disability should be carried out with extreme care, to avoid the methodological problems outlined in the literature review above. Just mentioned was the potential for a further trial, preferably with a much larger sample size, using a similar methodology to Blaskey et al. (1990), comparing visual stress treatments (coloured filters) to treatments for oculomotor disorders.

As well as performing trials, however, new and different methodologies could be utilized. This section outlines some potential ideas for future research in this area, and points out areas that have not been investigated, or have been neglected so far. The most useful future studies in this area should attempt to find objective evidence of visual stress and of any differences made by coloured filters. This section ends with a discussion of the potential for a cost-benefit analysis into coloured filters as a treatment.

Discussed in Section 3.1 was the notion that symptom-based definitions of visual stress are purely subjective, and are subject to concerns about 'leading questions', especially when the patients are young children. Characterizing visual stress as merely a response to coloured filters is a circular definition which does not throw any light on the causes of the disorder. Clearly, solid evidence of visual stress as a disorder is critically important if the condition is to be recognized from a scientific standpoint. As a corollary, if the coloured filter treatments are efficacious, we should presumably see evidence of this in more objective studies. Studies of this nature would, at

this point, be more useful than further trials of coloured filters; the theoretical basis of visual stress should be clarified so that a more well-defined disorder can be described, and more specific treatments designed.

While some brain imaging evidence exists for individuals with migraine, showing that they have abnormal blood oxygen level dependent signals when viewing aversive visual stimuli (Huang et al., 2003; see Wilkins et al., 2007 for discussion), no studies have been performed which investigate visual stress or the use of coloured filters at a brain level. Only if the connections between migraine and visual stress are solid – we saw above that this is not necessarily the case – would this be evidence for a brain-level change in visual stress. A rigorous, well-controlled imaging study similar in nature to the unpublished one referenced by Irlen (2010) – a study comparing brain activity when individuals diagnosed with visual stress are reading with and without filters, though with the proviso that they are compared to ability-matched controls – would be illuminating in this regard. Both theoretical and treatment effectiveness questions could potentially be addressed by such a study.

A large number of studies have investigated eye movements while reading. In a review of the area, Rayner et al. (2005) note that it is well-established that poor readers, when compared to normal readers, have in general longer eye fixations, increased regressions (moving back, rightwards, along the line to previously-read text) and shorter saccades (eye movements) when reading. It has been suggested that these faulty eye movements are a cause of reading disability, and are perhaps themselves caused by impaired magnocellular function (Stein & Talcott, 1999). However, Vellutino & Fletcher (2005) discuss studies showing that reading disabled individuals do not show anomalous eye movements when viewing nonverbal stimuli (see also Rayner, 1998 for a review). If the eye movement differences are restricted to the verbal modality, it is far more probable that they are a consequence, rather than a cause, of reading disability. Indeed, it is likely the case that 'less fluent eye movements reflect the difficulty that disabled readers are having understanding the text they are reading' (Rayner et al., 2005, p. 83). In any case, Hutzler et al. (2006) show that magnocellular function in particular – as measured on motion-based tasks – does not appear to be related to eye movement accuracy while reading (they are, at best, a correlate of poor reading, and for this reason could be used as a marker of reading disability, as hypothesized by Frith & Frith, 1996).

Nevertheless, it would be interesting to see such a study performed involving coloured filters. This study could ask two questions. Firstly, are the eye movement differences in visual stress sufferers different from those of other disabled readers who do not report visual stress symptoms? Data from an experiment of this nature could influence theoretical views of visual stress. Secondly, if visual stress sufferers, or any other disabled readers, show abnormal eye

movements while reading (as either a cause or consequence of their reading failure), does the use of coloured filters improve these movements towards a more normal level? Indeed, it is surprising that such an experiment has not yet been carried out, given that – properly double-blinded and controlled - it would provide convincing objective evidence for visual stress as a diagnostic entity.

In addition, studies that show the clinical or educational utility of the visual stress diagnosis would be of great interest. Discussed above is the evidence that phonological skills in early childhood are a good predictor of reading ability in later school years (Hulme & Snowling, 2009). If similar studies could relate later reading ability to the severity of visual stress symptoms, then the clinical value of the diagnosis would increase. The visual stress symptoms could be measured either subjectively or objectively, the former with the current methods – questions and questionnaires, noting the caveats discussed above - and the latter with some of the eye movement or brain imaging methods outlined in this section. On the basis of previous studies (Ramus et al., 2003; White et al., 2006) it is not clear that visual properties would have a great deal predictive power relating to reading ability. However, since visual stress was not directly assessed in these studies, the question remains open.

Discussed in Section 2.2.1 was the fact that proponents of cortical hyperexcitability theory have investigated the effects of coloured filters on a wide range of other, hypothetically linked disorders such as migraine, epilepsy, multiple sclerosis, and autism. Eperjesi et al. (2004) have looked into the effects of coloured filters for treating age-related macular degeneration, and other conditions that cause low vision (see Eperjesi et al. 2002 for a review). At this point in time, such experiments have found no evidence that the coloured filters are an effective intervention – they fail to produce a clinically significant improvement in reading rate in samples of individuals with low vision. Eperjesi et al. (2002) also note that a great deal of the research into this area is of poor quality. If coloured filters are to be used to treat low vision, along with other disorders, a great deal more rigorous evidence is required. This should include unselected, double-blind, randomised controlled trials.

Finally, rather like the evidence for their efficacy, there is confusion about economic aspects relating to the coloured filter treatment. To investigate these, we made several requests under the Freedom of Information (Scotland) Act 2002 to Scottish NHS Health Boards. We found a wide variation in costs, with some Health Boards (such as NHS Fife and NHS Lothian) not offering the treatment at all, while others, such as NHS Ayrshire & Arran, which spends approximately £33,650 annually on coloured overlays, lenses, and the requisite staffing costs. While this level of spending is not high in comparison to other NHS costs, we would suggest that it may be larger than one would expect for a treatment for which there is such uncertain evidence (Section 3). In

addition, the variation in spending on the treatment across health boards is once again illustrative of the confusion in this area.

Naturally, the NHS is not the only avenue through which visual stress screening and coloured filters are available. Irlen Scotland (2010) operate in a number of areas, and some other organizations (e.g. Crossbow Education, 2010; Rainbow Readers, 2010) offer private screening and treatment. The cost to the patient varies across these organizations, and costs tend to be much greater if tinted spectacles, as opposed to coloured overlays, are said to be required. For tinted spectacle lenses, Irlen are generally recognized to be the most expensive organisation, with the cost of screening, diagnosis, and issuing of the lenses somewhere in the region of £350-£450, and overlays costing substantially less than this (Albon et al., 2008).

A full economic analysis of the costs and benefits of the coloured filter treatment should be performed. Albon et al. (2008) attempted to perform such an analysis, but were unable to do so due to a severe lack of information. They could find no study that took into account the costs of the treatment and weighed them against the benefits, though they do note that 'the difficulty of such primary research should not however be underestimated.' (p. 88).

Discussed in the next section are some evidence-based treatments for dyslexia which involve teaching phonological skills. It may be possible, given information about the cost-effectiveness of these forms of treatment, to be able to compare them to the cost effectiveness of one or other of the forms of treatment with coloured filters along the lines of Scott et al. (2003), who analysed the cost-effectiveness of cognitive behavioural therapy for preventing relapses of clinical depression. These authors produced a cost-effectiveness acceptability curve, a graphical method of showing 'the probability that an intervention is cost-effective compared with the alternative, given the observed data, for a range of maximum monetary values that a decision-maker might be willing to pay for a particular unit change in outcome' (Fenwick & Byford, 2005, p. 106).

Due to the confusion about the causes of visual stress symptoms (i.e. the disagreement over whether or not it constitutes a separate syndrome), such a comparison between treatments would be difficult to carry out on the basis of current data. This suggests that, if the treatment continues to be used by educational psychologists and healthcare providers, a specific study of its cost-effectiveness versus other common interventions in randomised groups of appropriate children is essential. Indeed, a further step would be to compare coloured filters both to treatment for dyslexia – the phonological interventions discussed in the next section – and to vision therapy and oculomotor treatment - such as that discussed in Blaskey (1990) and Scheiman (2004). Comparing and contrasting the cost-effectiveness of all three types of intervention in a large population of disabled readers, perhaps giving details of the percentages of the group who

would expect each to be effective, would be of great utility for potential users of these treatments. Given that the present study and the preponderance of evidence seems to suggest that coloured filters are not an effective treatment for dyslexia, we would expect such an analysis would not be in favour of use of coloured filters.

#### *5.4 Effective treatments for reading disability*

It has been shown in this study and the literature review preceding it that there is very little convincing evidence for the efficacy or effectiveness of coloured filters for reading disorder. It is worth, then, briefly considering some treatments which are backed by more convincing evidence. This section will only discuss interventions for dyslexia, and not for orthoptic problems, which have been discussed above.

Hatcher et al. (1994) in a now-classic study from the standpoint of the phonological theory of dyslexia, gave four groups of children aged 6-7 years different types of reading intervention for periods of 20 weeks at a time, and compared the progress made by each group. The reading interventions were: 'phonological training', which involved only intensive instruction in identifying various aspects of the internal sound structure of words such as rhyme, segmentation and sound substitution; 'reading training', which involved only reading books and building up vocabulary; 'reading plus phonology', which involved a combination of both of the previous interventions; and lastly a control group, who received only their normal classroom teaching.

At the end of the 20 weeks, the authors found that the 'reading plus phonology' group had made the greatest gains in reading ability, followed by the. The authors note that this is 'an important, and not at all obvious, result' (p. 52). Given that the children in the 'reading plus phonology' group actually received *less* reading instruction than those in the 'reading training' group and yet still showed superior reading performance, the results confirmed that explicit phonological instruction is of great importance.

Torgesen et al. (2001) report evidence that different amounts of phonological instruction can have similar results as long as they are intertwined with reading training, while Hatcher et al. (2004), performed a study evaluating differing types of phonological intervention. Temple et al. (2003) performed a brain imaging study that found changes in activity in the left temporo-parietal regions (among other areas) of the brains of dyslexic children who had been receiving phonological training. This fulfils a prediction made by the phonological theory, which posits that these areas are functioning abnormally in dyslexics (Ramus, 2003). Torgesen (2002, 2005) has reviewed dyslexia interventions and offers some suggestions for future studies evaluating such treatments.

The option to predict and prevent reading disability has also become increasingly available to educators. Using a variety of methods including genetic risk factors, socioeconomic status, brain-based measures such as event-related potentials, and mostly importantly, tasks assessing phonological ability, it may become possible to accurately predict the severity of dyslexia in individual children, and target them with specific interventions (Gabrieli, 2009).

The interventions described thus far relate to the remediation of reading disorder itself. However, while these interventions are taking place, it would be desirable to have evidence-based interventions for some of the indirect adverse effects of reading disorder described in the Introduction section. It has been noted studies investigating this particular area are few and far between (Alexander-Passe, 2006). Burton (2005) describes 18-month discussion- and activity-based groups created for dyslexic children aged 11-12 that had a positive effect on a variety of measures of self-esteem. This study was rather small-scale, however, so more work on this interesting topic is required.

What is obvious from all the studies of interventions discussed here is that ameliorating reading disability is a very long and very intensive process, which needs careful planning and often involves teaching children a variety of different skills. As Hulme & Snowling (2009) point out, 'the miracle cure so far eludes us' (p. 89). Proponents of coloured filters should be wary of presenting their treatment as such a cure. It has been argued that 'complementary' and 'alternative' treatments for medical conditions, while often not inherently harmful, can be detrimental to health if used in place of, rather than alongside, conventional evidence-based treatments (Singh & Ernst, 2008). Though, of course, using coloured filters in place of evidence-based reading interventions would not have immediate health effects, it may delay the development of reading skills unnecessarily, potentially leading to many of the adverse social consequences of reading disability discussed in Section 1.1.

## 6. CONCLUSIONS

As we have seen, the use of coloured filters has been subject to considerable controversy. Helveston (2001) claims that the Irlen Institute displays ‘...classic group behaviour. The concept has a strong charismatic personality [Helen Irlen] as originator and sustained leader. The supporting evidence is almost entirely anecdotal... The procedure... remains a type of “trade secret”. Finally, a financially rewarding franchise activity is at the basis of the Irlen Institute activity.’ (p.13).

Judging whether these social criticisms of the Irlen Institute’s operation are valid is beyond the scope of a review of this nature; it can however be said confidently that the claims made by Irlen (e.g. 1991, 2010) are not consistent with the available evidence. Other claims, such as those of the research groups of Wilkins (e.g. 2003), Northway, (e.g. 2005) and Stein (e.g. Ray et al., 2005) are backed by some experimental evidence, but it has been argued here that a great deal of that evidence is either of poor quality or currently only tentative. Our own study, which we believe to have been of higher quality than previous work in the area, provided unambiguously negative results.

Thus, our study, and the review within this thesis, is in line with the conclusions drawn in recent reviews by Albon et al. (2008), the American Academy of Paediatrics (2009), the Royal College of Ophthalmologists (2009), and Wright, (2007), all of whom see no convincing evidence to recommend treating reading disability with coloured filters. Several of these reviews call for more basic research to be performed to carefully explore the effects of this treatment. We submit that our study is such research, and we found data which only reinforced the reviewers’ conclusions. Coloured filters, then, should not be recommended for alleviating reading disability until such a time that substantial, high-quality evidence is available.

## *ACKNOWLEDGMENTS*

My grateful thanks go to both of my academic supervisors.

Liz McKelvie from Irlen Scotland, Nadia Northway and Jillian Campbell of Glasgow Caledonian University Eye Clinic were also all extremely obliging, and I greatly appreciated their advice and input to various aspects of the experimental process and design.

Last, but certainly not least, all the pupils, teachers and support staff at Newark Primary School, Port Glasgow were unfailingly helpful and welcoming throughout the process of the experiment described herein, which could not have been performed without their kind assistance.

The study was supported by a grant from the Moray Endowment Fund.

## REFERENCES

- Albon E., Adi Y. & Hyde C. (2008). *The Effectiveness and Cost-effectiveness of Coloured Filters for Reading Disability: A Systematic Review*. Birmingham: University of Birmingham Department of Public Health and Epidemiology.
- Alexander-Passe, N. (2006). How dyslexic teenagers cope: an investigation of self-esteem, coping and depression. *Dyslexia*, 12 (4) 256-275.
- Ambrosini, A., De Noordhout, A. M., Sandor, P. S. & Schoenen, J. (2003). Electrophysiological studies in migraine: a comprehensive review of their interest and limitations. *Cephalalgia*, 23 (1), 13-31.
- American Psychiatric Association (APA) (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Washington, DC: American Psychiatric Association.
- Amitay, S., Ben-Yehudah, G., Banai, K. & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain*, 125 (10), 2272-2285.
- Arnold, A. C. (2005). Evolving management of optic neuritis and multiple sclerosis. *American Journal of Ophthalmology*, 139 (6), 1101-1108.
- Aurora, S. & Wilkinson, F. (2007). The brain is hyperexcitable in migraine. *Cephalalgia*, 27 (12), 1442-1453.
- Babbitt, E. D. (1878) *The Principles of Light and Colour*. London: Babbitt & Co.
- Badian, N. A. (1994). Preschool prediction: orthographic and phonological skills, and reading. *Annals of Dyslexia*, 44 (1), 1-25.
- Barrett, B. T. (2009) A critical evaluation of the evidence supporting the practice of behavioural vision therapy. *Ophthalmic and Physiological Optics*, 29, 4-25.
- Bates, T.C., Lind, P. A., Luciano, M., Montgomery, G. W., Martin, N. G., Wright, M. J. (2009) Dyslexia and DYX1C1: deficits in reading and spelling associated with a missense mutation. *Molecular Psychiatry*, Article in Press.
- Bender, W. N., Rosenkrans, C. B. & Crane, M. K. (1999). Stress, depression, and suicide among students with learning disabilities: assessing the risk. *Learning Disability Quarterly*, 22 (2), 143-156.
- Bishop, D. V. M. (1989). Unfixed reference, monocular occlusion, and developmental dyslexia - a critique. *British Medical Journal*, 73 (3), 209-215.
- Blaskey, P., Scheiman, M., Parisi, M., Ciner, E. B., Gallaway, M. & Selznick, R. (1990) The effectiveness of Irlen filters for improving reading performance: a pilot study. *Journal of Learning Disabilities*, 23 (10), 604-612.
- Bouldoukian, J., Wilkins, A. J. & Evans, B. J. W. (2002). Randomised controlled trial of the effect of coloured overlays on the rate of reading of people with specific learning difficulties. *Ophthalmic and Physiological Optics*, 22 (1), 55-60.
- Brighina, F., Palermo, A., Daniele, O., Aloisio, A. & Fierro, B. (2009). High-frequency transcranial magnetic stimulation on motor cortex of patients affected by migraine with aura: a way to restore normal cortical excitability?. *Cephalalgia*, Article in Press.
- Burton, S. (2004). Self-esteem groups for secondary pupils with dyslexia. *Educational Psychology in Practice*, 20 (1), 55-73.
- Castelo-Branco, M., Mendes, M., Sebastião, A. R., Reis, A., Soares, M., Saraiva, J., Bernardes, R., Flores, R., Pérez-Jurado, L., & Silva, E. (2007) Visual phenotype in Williams-Beuren syndrome challenges magnocellular theories explaining human neurodevelopmental visual cortical disorders. *Journal of Clinical Investigation*, 117 (12), 3720-3729.
- Castles, A. & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, 91 (1), 77-111.
- Chait, M., Eden, G., Poeppel, D., Simon, J. Z., Hill, D. F. & Flowers, D. L. (2007). Delayed detection of tonal targets in background noise in dyslexia. *Brain and Language*, 102 (1), 80-90.
- Chase, C., Ashourzadeh, A., Kelly, C., Monfette, S. & Kinsey, K. (2003) Can the magnocellular pathway read? Evidence from studies of color. *Vision Research*, 43 (10), 1211-1222.
- Chatterjee, S. & Callaway, E. M. (2003). Parallel colour-opponent pathways to primary visual cortex. *Nature*, 426 (6967), 668-671.
- Christenson, G. N., Griffin, J. R. & Taylor, M. (2001) Failure of blue-tinted lenses to change reading scores of dyslexic individuals. *Optometry*, 72 (10), 627-633.
- Chronicle, E. P. & Wilkins, A. J. (1991). Colour and visual discomfort in migraineurs. *Lancet*, 338, 8771.
- College of Syntonic Optometry (2010) *Website of the College of Syntonic Optometry*. Available at <http://www.syntonicphototherapy.com/online/> (accessed 10<sup>th</sup> August 2009).
- Coppola, G., Pierelli, F. & Schoenen, J. (2007). Is the cerebral cortex hyperexcitable or hyperresponsive in migraine? *Cephalalgia*, 27 (12), 1427-1439.

- Cornelissen, P., Bradley, L., Fowler, M. S. & Stein, J. F. (1992) Covering one eye affects how some children read. *Developmental Medicine & Child Neurology*, 34 (4), 296-304.
- Critchley, M. (1964) *Developmental Dyslexia*. London: Whitefriars Press.
- Crossbow Education (2010) Website of Crossbow Education. Available at <http://www.crossboweducation.com/> (accessed 10<sup>th</sup> August 2010).
- Dakin, S. & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, 48 (3), 497-507.
- Eckert, M. (2004). Neuroanatomical markers for dyslexia: a review of dyslexia structural imaging studies. *The Neuroscientist*, 10 (4), 362-371.
- Elbro, C., Borström, I. & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: the importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, 33 (1), 36-60.
- Elliot, A.J., Maier, M. A., Moller, A. C., Friedman, R. & Meinhardt, J. (2007). Colour and psychological functioning: the effect of red on performance attainment. *Journal of Experimental Psychology (General)*, 136 (1), 154-168.
- Elliot, C. D., Smith, P., McCulloch, K. (1996). *British Ability Scales – Second Edition*. Windsor: NFER-Nelson.
- Elliott, C. D., Murray, D. J. & Pearson, L. (1978). *British Ability Scales*. Windsor: NFER-Nelson.
- Eperjesi, F., Fowler, C. W. & Evans, B. J. W. (2002). Do tinted lenses or filters improve visual performance in low vision? A review of the literature. *Ophthalmic and Physiological Optics*, 22, 68-77.
- Eperjesi, F., Fowler, C. W. & Evans, B. J. W. (2004). The effects of coloured light filter overlays on reading rates in age-related macular degeneration. *Acta Ophthalmologica Scandinavica*, 82, 695-700.
- Evans, B. J. W. & Joseph, F. (2002) The effect of coloured filters on the rate of reading in an adult student population. *Ophthalmic and Physiological Optics*, 22 (6), 535-545.
- Evans, B. J. W. & Stevenson, S. J. (2008). The Pattern Glare Test: a review and determination of normative values. *Ophthalmic and Physiological Optics*, 28 (4), 295-309.
- Evans, B. J. W., Patel, R. & Wilkins, A. J. (2002). Optometric function in visually sensitive migraine before and after treatment with tinted spectacles. *Ophthalmic and Physiological Optics*, 22 (2), 130-142.
- Fawcett, A. J., (2000). Mono-ocular occlusion for treatment of dyslexia. *Lancet*, 356 (9224), 89-90.
- Fenwick, E. & Byford, S. (2005). A guide to cost-effectiveness acceptability curves. *British Journal of Psychiatry*, 187 (2), 106-108.
- Fergusson, D. M. & Lynskey, M. T. (1997). Early reading difficulties and later conduct problems. *Journal of Child Psychology and Psychiatry*, 38 (8) 899-907.
- Folstein, M. F., Folstein, S. E., McHugh, P.R. (1975). "Mini-Mental" State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Folstein, S. E. & Rosen-Sheidley, B. (2001). Genetics of autism: complex aetiology for a heterogeneous disorder. *Nature Reviews Genetics*, 2 (12), 943-955.
- Frith, C., & Frith, U. (1996). A biological marker for dyslexia. *Nature*, 382, 19-20.
- Frith, U. (2006). *Autism: Explaining the Enigma – Second Edition*. Oxford: Blackwell.
- Gabrieli, J. D. E. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science*, 325 (5938), 280-283.
- Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F. & Geschwind, N. (1985). Developmental dyslexia: four consecutive patients with cortical anomalies. *Annals of Neurology*, 18 (2), 222-233.
- Garzia, R. P., Richman, J. E., Nicholson, S. B. and Gaines, C. S. (1990). A new visual-verbal saccade test: the development eye movement test (DEM). *Journal of the American Optometric Association*, 61, 124-135.
- Gillberg, C. (1991). The treatment of epilepsy in autism. *Journal of Autism and Developmental Disorders*, 21 (1), 61-77.
- Gottlieb, R. & Wallace, L. (2001) Syntonic Phototherapy. *Journal of Behavioural Optometry*, 12 (2), 31-38.
- Gray, W. S. (1963). *Gray Oral Reading Test (First Edition)*. Indianapolis: Bobbs-Merrill.
- Griffin, J. R. & Walton, H. N. (1981). *Dyslexia determination test (DDT)*. Los Angeles: I-MED.
- Hannell, G., Gole, G. & Dibden, S. (1991) Reading improvement with tinted lenses: a report of two cases. *Journal of Research in Reading*, 14 (1), 56-71.
- Hatcher, P. J & Hulme, C. (1999). Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction: Evidence from a longitudinal intervention study. *Journal of Experimental Child Psychology*, 72 (2), 130-153.
- Hatcher, P. J., Hulme, C. & Ellis, A. W. (1994) Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, 65 (1), 41-57.

- Hatcher, P. J., Hulme, C. & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry*, 45 (2), 338-358.
- Hawke, J. L., Wadsworth, S. J. & DeFries, J. C. (2006). Genetic influences on reading difficulties in boys and girls: the Colorado twin study. *Dyslexia*, 12 (1), 21-29.
- Heath, S. M., Bishop, D. V. M., Hogben, J. H., Roach, N. W. (2006) Psychophysical indices of perceptual functioning in dyslexia: a psychometric analysis. *Cognitive Neuropsychology*, 23 (6), 905-929.
- Helveston, E. M. (1990) Scotopic sensitivity syndrome. *Archives of Ophthalmology*, 108 (9), 1232-1233.
- Helveston, E. M. (2001). Tinted lenses. *Perspectives: the International Dyslexic Association*, 27 (3) 12-13.
- Henning W. (1936) *The Fundamentals of Chrome-orthoptics*. California: Actino Laboratories Inc.
- Hoeft, F., Hernandez, A., McMillon, G., Taylor-Hill, H., Martindale, J. L., Meyler, A., Keller, T. A., Siok, W. T., Deutsch, G. K., Just, M. A., Whitfield-Gabrieli, S. & Gabrieli, J. D. E. (2006). Neural basis of dyslexia: a comparison between dyslexic and nondyslexic children equated for reading ability. *Journal of Neuroscience*, 26 (42), 10700-10708.
- Hollis, J. & Allen, P. M. (2006). Screening for Meares-Irlen sensitivity in adults: can assessment methods predict changes in reading speed? *Ophthalmic and Physiological Optics*, 26 (6), 566-571.
- Howell, E. & Stanley, G. (1988) Colour and learning disability. *Clinical and Experimental Optometry*, 71 (2), 66-71.
- Huang, J., Cooper, T. G., Satana, B., Kaufman, D. I. & Cao, Y. (2003). Visual distortion provoked by a stimulus in migraine associated with hyperneuronal activity. *Headache*, 43 (6), 664-671.
- Hughes, H. C., Nozawa, G. & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, 8, 197-230.
- Hulme, C. & Snowling, M. J. (2009) *Developmental disorders of Language Learning and Cognition*. Oxford: Wiley-Blackwell.
- Hulme, C., Hatcher, P. J., Nation, K., Brown, A., Adams, J. & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology*, 82 (1), 2-28.
- Hulme, C., Snowling, M. J., Caravolas, M. & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: a comment on Castles and Coltheart. *Scientific Studies of Reading*, 9 (4), 351-365.
- Humphrey, N. (2002). Teacher and pupil ratings of self-esteem in developmental dyslexia. *British Journal of Special Education*, 29 (1), 29-36.
- Humphreys, P., Kaufmann, W. E., Galaburda, A. M. (1990). Developmental dyslexia in women: neuropathological findings in three patients. *Annals of Neurology*, 28 (6), 727-738.
- Hutzler, F., Kronbichler, M., Jacobs, A. M. & Wimmer, H. (2006). Perhaps correlational but not causal: no effect of dyslexic readers' magnocellular system on their eye movements during reading. *Neuropsychologia*, 44 (4), 637-648.
- Iovino, I., Fletcher, J. M., Breitmeyer, B. G. & Foorman, B. R. (1998). Colored overlays for visual perceptual deficits in children with reading disability and attention deficit/hyperactivity disorder: are they differentially effective? *Journal of Clinical & Experimental Neuropsychology*, 20 (6), 791-806.
- Irlen Institute (2010) *Website of the Irlen Institute*. Available at <http://www.irlen.com/> (accessed 8<sup>th</sup> August 2010).
- Irlen Scotland (2009) *Website of the Irlen Institute (Scotland)*. Available at <http://www.irlenscotland.co.uk/> (accessed 8<sup>th</sup> August 2010).
- Irlen, H. (1983) Successful treatment of learning difficulties. *The Annual Convention of the American Psychological Association*. California, Anaheim.
- Irlen, H. (1987). *Irlen Differential Perceptual Schedule*. Long Beach: Perceptual Development Corporation.
- Irlen, H. (1991) *Reading by the Colors: Overcoming Dyslexia and Other Reading Disabilities Through the Irlen Method*. New York: Avery Publishing Group.
- Irlen, H. (2010) *The Irlen Revolution: A Guide to Changing Your Perception and Your Life*. New York: Square One Publishers.
- Jeanes, R., Busby, A., Martin, J. Lewis, E., Stevenson, N., Pointon, D. & Wilkins, A. (1997). Prolonged use of coloured overlays for classroom reading. *British Journal of Psychology*, 88 (4), 531-548.
- Johannes, S., Kussmaul, C. L., Münte, T. F. & Mangun, G. R. (1996) Developmental dyslexia: passive visual stimulation provides no evidence for a magnocellular processing defect. *Neuropsychologia*, 34 (11), 1123-1127.
- Jordan, T. R., Sherman, S. M. & Tonkin, R. P. (2007) Constructing isoluminant stimuli for word recognition research: A precautionary study. *Behavior Research Methods*, 39 (3), 494-501.
- Kaplan, R. (1983) Changes in form visual fields in reading disabled children produced by syntonic (colored light) stimulation. *International Journal of Biosocial Research*, 5, 20-33.
- Kirk, J. & Reid, G. (2001) An examination of the relationship between dyslexia and offending in young people and the implications for the training system. *Dyslexia*, 7, 77-84.

- Koch, M., Uyttenboogaart, M., Polman, S. & De Keyser, J. (2008). Seizures in multiple sclerosis. *Epilepsia*, 49 (6), 948-953.
- Kriss, I & Evans, B. J. W. (2005). The relationship between dyslexia and Meares-Irlen syndrome. *Journal of Research in Reading*, 28 (3), 350-364.
- Kyd, L. J. C., Sutherland, G. F. M. & McGettrick, P. M. (1992). A preliminary appraisal of the Irlen screening process for scotopic sensitivity syndrome and the effects of Irlen coloured overlays on reading. *British Orthoptic Journal*, 49, 25-30.
- LaPray M. & Ross, R. (1972) *Teaching Children to Become Independent Readers*. New York: Center for Applied Research in Education.
- Lieberman, J. (1986) The effect of syntonic (colored light) stimulation on certain visual and cognitive functions. *Journal of Optometric Vision Development*, 17, 4-15.
- Livingstone, M. S., Rosen, G. D., Drislane, F. W. & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 88, 7943-7947.
- Lovegrove, W., Martin, F. & Slaghuis, W. (1986). A theoretical and experimental case for a visual deficit in specific reading disability. *Cognitive Neuropsychology*, 3, 225-267.
- Lu, Z. L., Lesmes, L. A. & Sperling, G. (1999) The mechanism of isoluminant chromatic motion perception. *Proceedings of the National Academy of Sciences*, 96 (14), 8289-8294.
- Ludlow, A. K., Wilkins, A. J. & Heaton, P. (2006). The effect of coloured overlays on reading ability in children with autism. *Journal of Autism and Developmental Disorders*, 36 (4), 507-516.
- Maclachlan, A., Yale, S. & Wilkins, A. J. (1993). Open trial of subjective precision tinting: A follow-up of 55 patients. *Ophthalmic and Physiological Optics*, 13, 175-175.
- Meares, O. (1980) Figure/background, brightness/contrast and reading disabilities. *Visible Language*, 14, 13-29.
- Menacker, S. J., Breton, M. E., Breton, M. L., Radcliffe, J. & Gole, G. A. (1993) Do tinted lenses improve the reading performance of dyslexic children? A cohort study. *Archives of Ophthalmology*, 111, 213-218.
- Mulleners, W. M., Chronicle, E. P., Palmer, J. E., Koehler, P. J., Vredeveld, J. W. (2001) Suppression of perception in migraine: evidence for reduced inhibition in the visual cortex. *Neurology*, 56 (2), 178-183.
- Muter, V., Hulme, C., Snowling, M. J. & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology*, 40 (5), 665-680.
- Nabuzoka, D. & Smith, P. K. (1993). Sociometric status and social behaviour of children with and without learning difficulties. *Journal of Child Psychology and Psychiatry*, 34 (8), 1435-1448.
- Neale, M. D. (1989). *Neale Analysis of Reading Ability: Revised British Edition*. Windsor: NFER-Nelson.
- Newman Wright, B., Wilkins, A. J. & Zoukos, Y. (2007). Spectral filters can improve reading and visual search in patients with multiple sclerosis. *Journal of Neurology*, 254 (12), 1729-1735.
- Nicolson, R. I. & Fawcett, A. J. (1999). Developmental dyslexia: the role of the cerebellum. *Dyslexia*, 5 (3), 155-177.
- Noble, J., Orton, M. Irlen, S. & Robinson, G. (2004) A controlled field study of the use of coloured overlays on reading achievement. *Australian Journal of Learning Disabilities*, 9 (2), 14-22.
- Northway, N. (2003) Predicting the continued use of overlays in school children – a comparison of the Developmental Eye Movement test and the Rate of Reading test. *Ophthalmic and Physiological Optics*, 23, 457-464.
- Northway, N. (2003). Predicting the continued use of overlays in school children - a comparison of the Developmental Eye Movement test and the Rate of Reading test. *Ophthalmic and Physiological Optics*, 23 (5) 457-464.
- Northway, N., Manahilov, V. & Simpson, W. (2009) Coloured filters improve exclusion of perceptual noise in visually symptomatic dyslexics. *Journal of Research in Reading*, Article in Press.
- Ouvrier, R. A., Goldsmith, R. F., Ouvrier, S. & Williams, I. C. (1993). The value of the Mini-Mental State Examination in childhood: a preliminary study. *Journal of Child Neurology*, 8, 143-148.
- Palmer, S. E. (1999) *Vision Science: Photons to Phenomenology*. Cambridge, Mass.: MIT Press.
- Pammer, K. & Lovegrove, W. (2001). The influence of color on transient system activity: implications for dyslexia research. *Perception and Psychophysics*, 63 (3) 490-500.
- Paracchini, S., Scerri, T. & Monaco, A. P. (2007). The genetic lexicon of dyslexia. *Annual Review of Genomics and Human Genetics*, 8, 57-79.
- Pennington, B. F. & Olson, R. K. (2005). Genetics of dyslexia. In Snowling, M. J. & Hulme, C. (Eds.) *The science of reading: a handbook* (pp. 453-472). Oxford: Blackwell.
- Rainbow Readers (2010). *Website of Rainbow Readers*. Available at <http://www.rainbowreaders.co.uk/> (accessed 9<sup>th</sup> August 2010).
- Ramus, F. (2001). Talk of two theories. *Nature*, 412 (6845), 393-395.

- Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, 13 (2) 212-218.
- Ramus, F. (2004). Neurobiology of dyslexia: a reinterpretation of the data. *Trends in Neurosciences*, 27 (12), 720-726.
- Ramus, F. & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology*, 61 (1), 129-141.
- Ray, N. J., Fowler, S. & Stein, J. F. (2005). Yellow filters can improve magnocellular function: motion sensitivity, convergence, accommodation, and reading. *Annals of the New York Academy of Sciences*, 1039, 283-293.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K., Juhasz, B. J. & Pollatsek, A. (2005). Eye movements during reading. In Snowling, M. J. & Hulme, C. (Eds.) *The science of reading: a handbook* (pp. 79-98). Oxford: Blackwell.
- Riddick, B., Sterling, C., Farmer, M. & Morgan, S. (1999). Self-Esteem and Anxiety in the Educational Histories of Adult Dyslexic Students. *Dyslexia*, 5, 227-248.
- Robinson, G. L. & Whiting, P. R. (2003). The interpretation of emotion from facial expression for children with visual processing problems. *Australasian Journal of Special Education*, 27 (2), 50-67.
- Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T.E., Meltzer, H., & Carroll, J. (2004) Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association*, 291 (16), 2007-2012.
- Samuelsson, S., Herkner, B. & Lundberg, I. (2003). Reading and writing difficulties among prison inmates: a matter of experiential factors rather than dyslexic problems. *Scientific Studies of Reading*, 7 (1), 53-73.
- Scheiman, M. (2004). Coloured lenses to improve reading comfort and performance: are underlying vision problems being missed? *Journal of Optometric Vision Development*, 35, 37-41.
- Scott, J., Palmer, S., Paykel, E., Teasdale, J., Hayhurst, H. (2003). Use of cognitive therapy for relapse prevention in chronic depression: cost-effectiveness study. *The British Journal of Psychiatry*, 182 (3), 221-227.
- Scott, L., McWhinnie, H., Taylor, L., Stevenson, N. Irons, P., Lewis, E., Evans, M., Evans, B. J. W. & Wilkins, A. J. (2002). Coloured overlays in schools: orthoptic and optometric findings. *Ophthalmic and Physiological Optics*, 22, 156-165.
- Shapley, R. & Perry, V. H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, 9 (5), 229-235.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M. & Makuch, R. (1992) Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, 326, 145-150.
- Shovman, M. M. & Ahissar, M. (2006). Isolating the impact of visual perception on dyslexics' reading ability. *Vision Research*, 46 (20), 3514-3525.
- Simmers, A. J., Gray, L. S. & Wilkins, A. J. (2001). The influence of tinted lenses upon ocular accommodation. *Vision Research*, 41 (9), 1229-1238.
- Singh, S. & Ernst, E. (2008). *Trick or Treatment? Alternative Medicine on Trial*. London: Bantam Press.
- Singleton, C. & Trotter, S. (2005). Visual stress in adults with and without dyslexia. *Journal of Research in Reading*, 28 (3), 365-378.
- Skottun, B. C. (2000). On the conflicting support for the magnocellular-deficit theory of dyslexia. *Trends in Cognitive Sciences*, 4 (6), 211-212.
- Skottun, B. C. (2000a). The magnocellular deficit theory of dyslexia: the evidence from contrast sensitivity. *Vision Research*, 40 (1), 111-127.
- Skottun, B. C. (2004). On the use of red stimuli to isolate magnocellular responses in psychophysical experiments: A perspective. *Visual Neuroscience*, 21 (1), 63-68.
- Skottun, B. C. (2005). Magnocellular reading and dyslexia. *Vision Research*, 45 (1), 133-134.
- Skottun, B. C. & Parke, L. A. (1998). The possible relationship between visual deficits and dyslexia: examination of a critical assumption. *Journal of Learning Disabilities*, 32, 2-5.
- Skottun, B. C. & Skoyles, J. (2007a). Yellow filters, magnocellular responses, and reading. *International Journal of Neuroscience*, 117 (2), 287-293.
- Skottun, B. C. & Skoyles, J. (2007b). Dyslexia and rapid visual processing: A commentary. *Journal of Clinical and Experimental Neuropsychology*, 30 (6), 666-673.
- Skoyles, J. & Skottun, B. C. (2004). On the prevalence of magnocellular deficits in the visual system of non-dyslexic individuals. *Brain and Language*, 88 (1), 79-82.
- Snowling, M. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatry Review* 3, 4-11.

- Snowling, M. J. (1998). Dyslexia as a phonological deficit: evidence and implications. *Child Psychology and Psychiatry Review*, 3, 4-11.
- Society for Coloured Lens Providers (2010). Website of the Society for Coloured Lens Providers. Available at: <http://www.s4clp.org/> (accessed 10<sup>th</sup> August 2010).
- Sperling, A. J., Lu, Z. L., Manis, F. R. & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience*, 8 (7), 862-863.
- Sperling, A. J., Lu, Z. L., Manis, F. R. & Seidenberg, M. S. (2006). Motion perception deficits and reading impairment. *Psychological Science*, 17, 1047-1053.
- Spitler HR. (1941). *The Syntonice Principle*. Eaton, OH: College of Syntonice Optometry.
- Stankewitz, A. & May, A. (2009). The phenomenon of changes in cortical excitability in migraine is not migraine-specific – A unifying thesis. *Pain*, 145 (1-2), 14-17.
- Stein, J. F. (1989) Unfixed reference, monocular occlusion, and developmental dyslexia - a critique (response). *British Medical Journal*, 73, 319-320.
- Stein, J. F. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia*, 7 (1), 12-36.
- Stein, J. F. & Fowler, M. S. (1985) Effect of monocular occlusion on visuomotor perception and reading in dyslexic children. *The Lancet*, 326 (8446), 69-73.
- Stein, J. F. & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neuroscience*, 20, 147-152.
- Stein, J. F., Richardson, A. F. & Fowler, M. S. (2000b) Monocular occlusion can improve binocular control and reading in dyslexics. *Brain*, 123 (1), 164-170.
- Stein, J. F., Talcott, J. & Walsh, V. (2000a). Controversy about the visual magnocellular deficit in developmental dyslexics. *Trends in Cognitive Sciences*, 4 (6), 209-211.
- Stockman, A., MacLeod, D. I. A. & DePriest, D. D. (1991). The temporal properties of the human short-wave photoreceptors and their associated pathways. *Vision Research*, 31 (2), 189-208.
- Talcott J. B., Hansen, P. C., Assoku, E. L., & Stein, J. F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia*, 38 (7), 935-943.
- Tallal, P. (2003) Language learning disabilities. *Current Directions in Psychological Science*, 12 (6), 206-211.
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., Gabrieli, J. D. E. (2003) Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proceedings of the National Academy of Sciences*, 100 (5), 2860-2865.
- Temple, E., Poldrack, R. A., Salidis, J. Deutsch, G. K., Tallal, P., Merzenich, M. M. & Gabrieli, J. D. E. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *Neuroreport*, 12 (2), 299.
- Torgesen, J. K. (2002). Lessons learned from intervention research in reading: a way to go before we rest. In Stainthorp, R. & Tomlinson, P. (Eds.), *Learning and Teaching Reading* (pp. 89-103). London: British Psychological Society.
- Torgesen, J. K. (2005). Recent discoveries on remedial interventions for children with dyslexia. In Snowling, M. J. & Hulme, C. (Eds.) *The science of reading: a handbook* (pp. 521-537). Oxford: Blackwell.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: immediate and long-term outcomes from two instructional processes. *Journal of Learning Disabilities*, 34, 33-58.
- University of Essex (2010). *Website of the University of Essex (Presentation: 'Examination with the Intuitive Colorimeter and Precision Tints')*. Available at <http://www.essex.ac.uk/psychology/overlays/colorimetry.pdf> (accessed 8th August 2010).
- Vellutino, F. R. & Fletcher, J. M. (2005). Developmental dyslexia. In Snowling, M. J. & Hulme, C. (Eds.) *The science of reading: a handbook* (pp. 362-378). Oxford: Blackwell.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45 (1), 2-40.
- Wadsworth, S. J., Olson, R. K., Pennington, B. F. & DeFries, J. C. (2000). Differential genetic etiology of reading disability as a function of IQ. *Journal of Learning Disabilities*, 33 (2), 192-199.
- Waldie, M. & Wilkins, A. J. (2004). How big does a coloured overlay have to be? *Ophthalmic and Physiological Optics*, 24 (1), 57-60.
- Wauters, A (1999). *Homeopathic Colour Remedies*. Freedom, CA: Crossing Press.
- White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U. & Ramus, F. (2006). The role of sensorimotor impairments in dyslexia: a multiple case study of dyslexic children. *Developmental Science*, 9 (3), 237-269.

- Whiteley, H. E. & Smith, C. D. (2001). The use of tinted lenses to alleviate reading difficulties. *Journal of Research in Reading*, 24 (1), 30-40.
- Wiederholt, J. L. (1986). *Formal Reading Inventory*, Austin: Pro-ed.
- Wiederholt, J.L., & Bryant, B.R. (2001). *Gray Oral Reading Tests – Fourth Edition*. Shoal Creek, Austin, : Pro- Ed.
- Wilkins, A. J. (1994). Overlays for classroom and optometric use. *Ophthalmic & Physiological Optics*, 14, 97–99.
- Wilkins, A. J. (1995). *Visual Stress*. Oxford: Oxford University Press.
- Wilkins, A. J. (2002). Coloured overlays and their effects on reading speed: a review. *Ophthalmic & Physiological Optics*, 22, 448–454.
- Wilkins, A. J. (2003). *Reading Through Colour*. Chichester: John Wiley and Sons.
- Wilkins, A. J. & Evans, B. J. W. (2001). Pattern Glare Test. London: i.O.O. Sales Ltd.
- Wilkins, A. J. & Evans, B. J. W. (2009). Visual stress, its treatment with spectral filters, and its relationship to visually induced motion sickness. *Applied Ergonomics*, article in press.
- Wilkins, A. J. & Nimmo-Smith, M. I. (1987). The clarity and comfort of printed text. *Ergonomics*, 30, 1705–1720.
- Wilkins, A. J., Baker, A., Amin, D., Smith, S., Bradford, J., Zaiwalla, Z., Besag, F. M. C., Binnie, C. D. & Fish, D. (1999). Treatment of photosensitive epilepsy using coloured glasses. *Seizure: European Journal of Epilepsy*, 8 (8), pp. 444-449.
- Wilkins, A. J., Binnie, C. D. & Darby, C. E. (1980). Visually-induced seizures. *Progress in Neurobiology*, 15 (2), 85-117.
- Wilkins, A. J., Evans, B. J. W., Brown, J. A., Busby, A. E., Wingfield, A. E., Jeanes, R. & Bald, J. (1994). Double-masked placebo-controlled trial of precision spectral filters in children who use coloured overlays. *Ophthalmic and Physiological Optics*, 14 (4), 365-370.
- Wilkins, A. J., Huang, J. & Cao, Y. (2007). Prevention of visual stress and migraine with precision spectral filters. *Drug Development Research*, 68 (7), 469-475.
- Wilkins, A. J., Nimmo-Smith, I., Tait, A., McManus, C., Della Sala, S., Tilley, A., Arnold, K., Barrie, M. & Scott, S. (1984). A neurological basis for visual discomfort. *Brain*, 107 (4), 989-1017.
- Wilkins, A. J., Patel, R., Adjamian, P. & Evans, B. J. W. (2002). Tinted spectacles and visually sensitive migraine. *Cephalalgia*, 22 (9), 711-719.
- Wilkins, A. J., Sihra, N. & Smith, I. N. (2005). How precise do precision tints have to be and how many are necessary? *Ophthalmic and Physiological Optics*, 25 (3), 269-276.
- Wilkins, A.J. (1995). *Visual Stress*. Oxford: Oxford University Press.
- Wilkins, A.J., Huang, J. & Cao, Y. (2004). Visual stress theory and its application to reading and reading tests. *Journal of Research in Reading*, 27 (2), 152-162.
- 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., Nimmo-Smith, M.I. & Jansons, J. (1992). A colorimeter for the intuitive manipulation of hue and saturation, and its application in the study of perceptual distortion. *Ophthalmic and Physiological Optics*, 12, 381-385.
- Williams, M. & LeCluyse, K. (1990). Perceptual consequences of a transient system deficit in disabled readers. *Journal of the American Optometry Association*, 61, 111-121.
- Williams, M. C., May, J. G., Solman, R. & Zhou, H. (1995). The effects of spatial filtering and contrast reduction on visual search times in good and poor readers. *Vision Research*, 35 (2), 285-291.
- Williams, S. & McGee, R. (1994). Reading attainment and juvenile delinquency. *Journal of Child Psychology and Psychiatry*, 35 (3), 441-459.
- Woodcock, R. (1987). *Woodcock Reading Mastery Test*. Circle Pines: American Guidance Service.
- World Health Organisation (1992). *Manual of the International Statistical Classification of Diseases – Tenth Revision*. Geneva: World Health Organisation.
- Wright, C. (2007) Learning disorders, dyslexia, and vision. *Australian Family Physician*, 36 (10), 843-845.
- Xiao, Y., Wang, Y. & Felleman, D.J. (2003). A spatially organized representation of colour in macaque cortical area V2. *Nature*, 421 (6922), 535–539.
- Zeki, S. (1983). Colour coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, 9 (4), 741-765.
- Ziegler, J., Pech-Georgel, C. & George, F. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12 (5), 732-745.