Author:
Cowan, P J

Title:
Do coloured filters benefit reading beyond placebo?

Do coloured filters benefit reading beyond placebo?

(Philippa) Jane Cowan

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of PhD in the Faculty of Science, Experimental Psychology. September 2018.

45,282 words

Motivation for thesis

Reading text was always a problem for me. Letters vibrated and had bright halos around them. My head felt fuzzy and comprehension was difficult. These symptoms of visual stress were apparently removed by coloured filters. Coloured filters enabled me to read and I recommended them to people with visual stress when I taught. Then I became interested in the science; I was surprised to learn that there is a divide in opinion over the effectiveness of coloured filters and whether associated reading benefits are a placebo effect.

Abstract

Individually chosen coloured filters (overlays or lenses) are claimed to benefit reading for visual stress sufferers, but this could be attributed to a placebo effect. Previous research lacks both suitable measures of reading benefit and methods of controlling for placebo effects, and the effect of contrast on visual stress is relatively unexplored.

A ‘cloze’ task was developed to measure the reading benefit of filters. Additional experiments tested whether chosen coloured filters improved reading performance, visual search or single word reading of filter users; placebo effects were controlled for using change blindness and Stroop Interference. The effect of leading instructions on performance of the reading task used to diagnose visual stress (WRRT) was investigated, and a visual stress questionnaire was administered. The effects of contrast on preference ratings for people who reported visual stress were measured. Lastly, the effect of reducing text contrast on filter users’ reading performance was quantified.

The cloze task demonstrated no reading benefit of a chosen coloured filter. Leading instructions were found to improve WRRT reading speeds with a chosen coloured filter but not the cloze task. Participants who reported visual stress had improved cloze reading performance with any filter and reported that they found text easier to read on blue or grey paper rather than white. However, filter users did not improve their reading speed with a reduced text contrast setting.

This thesis presents no new evidence to support the view that chosen coloured filters are needed to reduce visual stress beyond placebo. However, effects of placebo were observed with a chosen filter on the WRRT, casting doubt on this method for the diagnosis of visual stress. Some data reported here indicate that reduced contrast may reduce visual stress.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE:........................

Acknowledgements

I'd like to thank my supervisors for their guidance, support, inspiration and humour. I feel privileged to have worked with them and thank them for taking me on. I'd like to extend gratitude to Arnold Wilkins for his help with Experiment 4 and support with the research early on. I'd like to thank Ute Leonards and Peter Rogers for their pastoral support and guidance, and Dave Turk for his support with recruitment. I've made some great friends at the department who have all been generous with their time and provided me with security as I've bounced ideas with them. Amongst them are Will Chapman, Ella Gale, Gaurav Malhorta, Brian Sullivan, Lizzie Smith, Robyn Wootton, Adele Wang, Jim Lumsden, Rob Udale and Bobby Stuijfzand – Thank you for your encouragement and ideas. Finally, I'd like to thank my housemate, Sarah Collins, my sister and brother in law, Ellie and Andy Lytton and my mum, Anthea Cowan, for their unwavering support at the end.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Visual Stress and Chosen Coloured Filters for Reading

Visual stress, sometimes referred to as 'Irlen syndrome', 'Scotopic Sensitivity Syndrome' (Irlen, 1991a; Robinson & Foreman, 1999a), 'Meares-Irlen syndrome' (Evans, 2005) and 'MISViS' (Kruk, Sumbler & Willows, 2008) refers to perceptual distortions that are reportedly alleviated by a chosen coloured filter. Coloured filters are coloured, plastic overlays that are placed over text for reading, or coloured, plastic lenses that are worn as glasses or contact lenses for general use. Wilkins, Huang and Cao (2004) described visual stress to be 'the inability to see comfortably without distortion and discomfort... and to read for long periods. Reading symptoms involves perceptual distortions in text that distract and prevent the words from being easily decoded such as blurring, moving or disappearing letters and a dazzling or shimmering white page (Irlen, 1991a; Wilkins, 2003; Wilkins, Allen, Monger & Gilchrist, 2016). Physical symptoms involve headaches, eyestrain, photophobia, reduced span of focus, dizziness, fatigue and limited attention (Irlen, 1991a; Wilkins, 2003). Although visual stress has been associated with migraine and epilepsy (Irlen, 1991a; Wilkins, 2003), this thesis is only concerned with the effects of chosen coloured filters on reading and so the term 'visual stress' will refer to perceptual distortions when viewing and reading text, not when perceiving the general environment.

Visual stress is not categorised as a disability by any large regulatory body (Griffiths et al., 2016), for example in the Diagnostic Statistic Manual of Mental Disorders (American Psychiatric Association, 2013). This lack of recognition may be due to an absence of validated diagnosis methods that rely upon a subjective response to a pre-selected filter. There is also an absence of convincing research that supports the assertion that chosen coloured filters are needed to reduce visual stress (section 1.9). Systematic reviews concluded that there is insufficient evidence to advocate the use of chosen coloured filters for a reading intervention and highlight weaknesses in methodology and statistics, variable techniques and inconsistent findings (Albon et al., 2008; Galuschka, Ise, Krick & Schulte-Körne, 2014; Griffiths, Taylor, Henderson & Barrett, 2016; American Academy of Paediatrics, 2009). A common limitation of the studies was that they were often not suitably blinded: blinding is particularly important because participants would be likely to be motivated to perform better with a colour of filter they had chosen to reduce their visual stress (Wilkins, 2003, p41).

Coloured filters are widely used in practice, administered in opticians and disability services across the UK, and endorsed by the college of Optometrists in the UK and six out of eight UK dyslexia charities (Henderson, Taylor, Barrett, Griffiths, 2014). Some local authorities have even subsidised the cost of filters (Albon et al., 2008; Ritchie, Della Sala & McIntosh, 2011). Coloured lenses are expensive, reportedly costing up to £400 (Albon et al., 2008). Continued use despite insufficient evidence may reflect a placebo response to a chosen coloured filter (McIntosh & Ritchie, 2012). A better understanding of the role of placebo in the use of coloured filters would inform the development of more appropriate, cost-effective reading interventions.

1.2 Background

Chosen coloured filters were first documented to reduce perceptual distortions and benefit reading in the context of Irlen syndrome (Irlen, 1983; Irlen, 1991a). Adults with unknown reading difficulties reported that their perceptual and physical symptoms when viewing and reading text were reduced when reading with a chosen colour of overlay (Irlen, 1991a). A chosen coloured lens, often a different hue to the coloured overlay, appeared to benefit these subjects when perceiving the general environment. Irlen (1991a; 1994; 2010) proposed that the brain becomes 'over-active' when attempting to process certain light frequencies, and that the removal of these offending light frequencies reduces the over-activity and therefore removed a barrier to reading development. Irlen syndrome is treated and diagnosed with the Irlen method using Irlen filters that are administered by privately trained educationalists. The Irlen method is marketed by the Irlen Institute in 170 Irlen Clinics and 46 countries worldwide (Irlen Institute, 2017a).

The Irlen Institute (2017c) claims that neuroimaging studies provide support for the existence of Irlen syndrome and the effectiveness of the Irlen method. However, these claims are either misrepresented or based on unpublished research. For example, one peer-reviewed, neuroimaging study was claimed to provide evidence of the neurobiological basis of Irlen syndrome (Irlen Institute, 2017b) but was a post-hoc analysis of a comparison between one participant with Irlen syndrome and nine subjects who had already completed an fMRI reading study (Chouinard et al., 2012) and so can only be treated as preliminary. One unpublished studies was described to demonstrate effectiveness of the Irlen method with SPECT scans showing an increased activity in the brain's 'emotional and visual processing centers' (Irlen Institute, 2017c). The accompanying image (Figure 1.1) has been widely used (e.g. Irlen, 2010, p101) to demonstrate the effectiveness of Irlen filters, yet its over-simplicity, and lack of detail regarding what would constitute an 'overactive brain' raises questions regarding its authenticity.

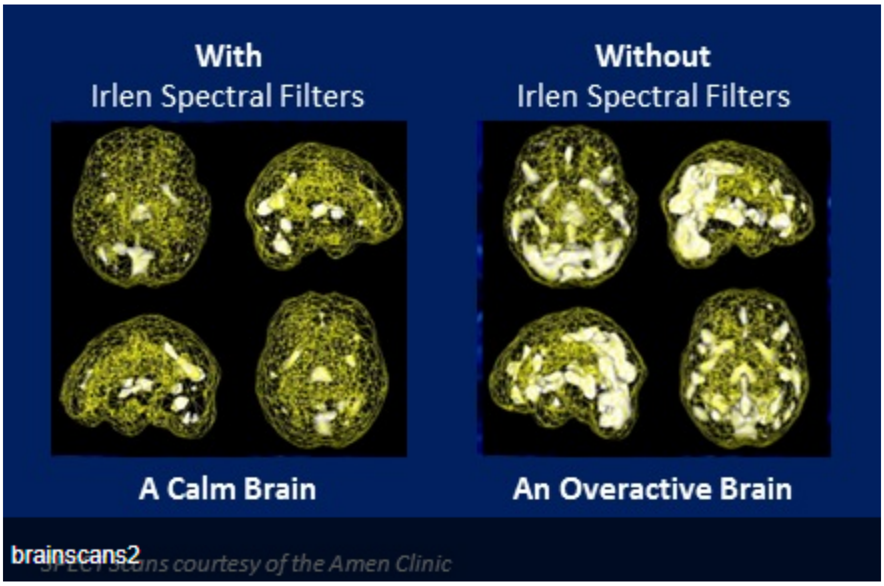


Figure 1.1. Image used by the Irlen Institute to demonstrate effectiveness of filters

Wilkins (2003) criticised the Irlen method for its exaggerated claims, subjective diagnostic criteria and lack of transparency: there was no public record of the colour specification process by which the colours of filters were chosen. Wilkins (1995) created the term 'visual stress' to describe symptoms of Irlen syndrome. He developed a new treatment and diagnosis method, which will be collectively referred to as the Intuitive method. Having reported a potential link between visual stress and visual discomfort and without evidence from blinded, placebo-controlled trials, Wilkins was "convinced… that the treatment could be very beneficial" (Wilkins, 1995, p. xvi). The intuitive method is marketed in UK opticians magazines and used by some UK opticians and regulated by the 'society of coloured lens providers' (Evans & Allen, 2016).

1.3 A simple view of reading

Coloured filters are claimed to benefit reading, apparently according to the 'simple view of reading' (Gough & Tunmer, 1986). The simple view of reading separates reading into two distinct processes; firstly, decoding the written word, and secondly, comprehending the overall meaning (Gough & Tunmer, 1986; Hoover & Gough, 1990). By alleviating visual stress symptoms, filters are claimed to improve the ability to perceive words in text prior to the decoding stage (Irlen, 1991a; Wilkins, 2003), thereby adding in an extra stage to the simple view of reading.

Irlen (2011) stipulated that the reading benefits attributed to the reduced perceptual distortions from chosen filters can only be observed when participants have mastered the basic word decoding skills and don't have other reading difficulties. In support of this claim, some studies demonstrated that populations without reading difficulties improved speed of decoding text and comprehension with a chosen filter (Robinson & Conway, 2000; Noble, Orton, Irlen & Robinson, 2004). Other studies (section 1.9.2) have used a word reading test that was designed to be appropriate for people with reading difficulties by not measuring comprehension and containing only high frequency, known word (Wilkins, Jeanes, Pumfrey & Laskier, 1996). However, it is unclear whether the results of these studies were due to a placebo effect attributed to using a chosen coloured filter.

1.4 How is colour processed in the visual system?

The initial stage of colour processing is believed to use a Trichromatic and Opponent colour system (Buchsbaum & Gottschalk, 1983). On the retina, three types of photoreceptors, long (L), medium (M) and short (S) cones, are maximally triggered by long (red) medium (green) and short (blue) wavelengths. The Opponent system is believed to begin when these three signals are transmitted by retinal ganglion cells to neurons in the lateral geniculate nucleus of the thalamus (LGN) (Gegenfurtner, 2003; Solomon & Lennie, 2007). Here, the signal is recombined by Magnocellular (M), determining luminance, and Parvocellular (P) and Koniocellular (K) cells, determining chromaticity. P cells compare L and M activation to detect the ratio of red and green light and K cells compare S and combined L and M activation to detect the ratio of blue and yellow light (Gegenfurtner, 2003; Solomon & Lennie, 2007).

The physiological processes involved in the primary visual cortex are more complex. P cells and K cells innervate cortical cells in layer 4 of the V1 and the upper layers of V1 in the primary visual cortex (Gegenfurtner, 2003). Other colour selective regions are triggered by a narrow range of wavelengths in the V2 and the posterior inferior temporal (PIT) cortex (Xiao, Wang and Felleman, 2003). At later stages, the perception of colour is argued to be cognitively penetrable (Macpherson, 2012). However, the complete processes involved in experiencing colour perception (hue, lightness and saturation) are not yet known.

1.5 Theories behind visual stress

The underlying neurological mechanism of visual stress and benefit from coloured filters remains unknown but the theory of visual stress and the magnocellular hypothesis have been proposed.

1.5.1 The theory of visual stress

The theory of visual stress posed that visual stress was part of a wider phenomenon known as 'visual discomfort' (Wilkins, 1995). Visual discomfort was proposed to be due to an inappropriate firing of neurons in the visual cortex associated with the hyper-excitability theory of migraine (Wilkins et al., 1984; Wilkins, 1995). Otherwise known as 'pattern glare', visual discomfort described perceptual distortions, fatigue, dizziness and headaches when people with migraine or epilepsy viewed patterned stimuli of particular specifications or were exposed to flickering lights (Wilkins et al., 1984). The 'pattern glare test' was created to trigger and measure the severity of visual discomfort by recording the number of reported symptoms when viewing a stimulus designed to trigger visual discomfort; a square wave grating with spatial frequency of three cycles/degree and duty cycle of 50% (Wilkins et al., 1984; Wilkins, 1995). The pattern glare test elicited abnormal BOLD responses in people who suffer from migraine (Huang et al., 2011), and induced seizures in photosensitive epilepsy (Wilkins et al, 1999, Parra, Lopes da Silva, Stroink, & Kalitzin, 2007).

The theory of visual discomfort was extended to incorporate visual stress when the number of reported symptoms when viewing the pattern glare test correlated with markers of visual stress (Wilkins, 1995), namely the number of self-reported Irlen symptoms (Wilkins & Nimmo-Smith, 1987) and perceived effectiveness of the chosen filter warranting an Irlen diagnosis (Wilkins & Neary, 1991). These correlations of subjective measures were carried out in non-blinded conditions and may be the result of a placebo effect rather than visual stress symptoms.

1.5.2 The magnocellular hypothesis

The magnocellular hypothesis (Stein, 2001; Ray et al., 2005), was developed from the visual magnocellular theory of dyslexia (section 1.8.1). Stein (2001) argued that magnocellular deficits were only relevant to some participants with dyslexia who experience visual perceptual distortions which he labelled 'visual dyslexia'. Distortions were suggested to arise by an impeded magnocellular pathway causing delayed signals to the visuomotor centres that control eye movements and binocular stability. Only blue or yellow chosen filters were claimed to rebalance magnocellular system and reduce distortions (Ray, Fowler & Stein, 2005). The hypothesis appears to refer to individuals with visual stress because the perceptual distortions experienced by dyslexic readers were described to involve blurry and moving text (Stein, 2001), and these perceptual symptoms characterise visual stress (Wilkins, 1995; Wilkins, 2003). The theory has also been referenced to explain the effect of Irlen lenses on visual stress (Noble et al., 2004; Whiteley & Smith, 2001).

There is some evidence that suggests that magnocellular deficit and associated visual-attention problems in participants with visual stress and dyslexia. Iles, Walsh and Richardson (2000) reported that a subset of dyslexic participants who had demonstrated a deficit in motion coherence thresholds also demonstrated a deficit in serial visual search. Fisher, Chekaluk and Irwin, (2015) reported that dyslexic individuals who had high visual stress symptomology showed a significantly poorer performance on a driving task than dyslexic individuals who did not have visual stress, and that self-reported measures of dyslexia and visual stress were well correlated. However, there is uncertainty over whether these tasks tap into magnocellular processes (Skottun, 2000).

1.6 A Placebo effect?

A 'placebo effect' describes improved symptoms or outcome derived from an inert treatment (Beecher, 1955) known as a 'placebo'. Placebos don't cause placebo effects because they are inert (Gotsche, 1994). Rather, a placebo effect is a physiological response attributed to the meaning associated with the placebo treatment (Moerman & Jonas, 2002). The placebo effect has been demonstrated to be a biological response whereby neurotransmitters are released (endorphins and dopamine) and symptoms are improved (Dodd, Dean, Vian & Berk, 2017).

The placebo effect has only been studied in a narrow range of conditions and so the theoretical underpinnings are not well understood. The main theories are the conditioning hypothesis and the expectation hypothesis (Haour, 2005). The conditioning hypothesis is not relevant to this thesis because it involves exposure to an active treatment prior to exposure to a placebo. The expectancy hypothesis posits that the placebo effect occurs due to a pre-existing belief about the treatment, which could be harboured by explicit instruction, suggestion or social cues (Dodd et al, 2017). This thesis will use the expectancy hypothesis to refine the definition of the placebo effect: *the placebo effect is the effect of expectation of success of a treatment.*

1.7 Diagnosis and treatment filter methods

Despite uncertainty over whether coloured filters do benefit reading beyond placebo (section 1.9), there is a factional dispute regarding which method of diagnosis and treatment is most effective, particularly between advocates of the Irlen method and Intuitive method. This has meant that findings suggesting no true efficacy of chosen coloured filters beyond placebo have been contested to be due to the use of the 'wrong' filter method. For example, Ritchie et al.'s (2011) findings did not support efficacy of chosen coloured filters beyond placebo and were criticised for using the Irlen method (Wilkins et al, 2016). Equally, the Irlen Institute has maintained that Irlen syndrome and visual stress is only alleviated by Irlen filters and the Irlen method (Irlen, 1991; Irlen, 2010; Irlen, 2015, Irlen Institute, 2017a).

There is no evidence to suggest that one method of diagnosis and treatment is more effective than the other. Diagnosis and treatment of visual stress now follow a full eye test to rule out undiagnosed ocular motor disorders that can be confused with visual stress, as reported by Blaskey et al. (1990) and Scheiman et al. (1990). Regardless of the method, diagnosis of visual stress relies upon an assessment of the benefit of a chosen treatment; a colour of filter that had been chosen by its apparent effectiveness in alleviating symptoms of visual stress. Since chosen treatments often promote a placebo response (Thompson, Ritenbaugh & Nichter, 2009) it would be surprising if there were no placebo effect in the diagnosis of visual stress.

1.7.1 Treatment

Treatment methods require a subjective preference of a colour of filter (overlays or lenses) from a selection of colour choices. The Irlen and Intuitive method of overlay selection are similar. Different pairs of coloured overlays, selected from 10 options, are placed side by side over some high contrast text stimuli, and the preferred overlay, the one that is reported to make text appear most comfortable, stable and clear, is identified and retained for a comparison with the next colour while non-preferred overlays are eliminated. Further preferred overlays are added if they are thought to benefit.

The process of acquiring lenses is from a wider range of colour choice and so is deemed to be a more accurate and effective treatment than overlays (Irlen, 1991a; Wilkins, 2003). The Irlen and Intuitive methods of lens selection are quite different. The Irlen method lasts a few hours and compares the appearance of text and the environment through two pairs of lenses at a time; the preferred lens is fine-tuned by combination with further preferred lenses (Irlen, 2010). Intuitive lenses are obtained in a 20-30 minute process using the ‘Intuitive Colorimeter’; this machine independently adjusts hue, saturation and luminance (Wilkins, Nimmo-Smith & Jansons, 1992). The colorimeter illuminates a page of text with coloured light that has been calibrated to give the appearance of coloured lenses being worn. There are twelve chromaticities of approximately equal luminance and saturation and approximately equally spaced hues. Light changes gradually through the different hues in a set order (Wilkins et al., 1992). A preferred hue is chosen, sometimes by forced choice, and adjusted according to its subjective effect of reducing visual stress symptoms. The saturation and luminance is then independently adjusted (Wilkins et al., 1992; Wilkins, 1997; 2001).

The Irlen method presents filters in any order and combines any preferred filters together. The Intuitive method presents filters in a specific order to avoids presenting complementary colours in close proximity to one another and combines only neighbouring hues of overlay to prevent complementary hues from being combined (Wilkins, 2003). This specified order may lead to order effects, as reported by Kriss and Evans (2005), although Veszeli and Shepherd (2019) reported no evidence of order effects.

The Intuitive method can be more easily replicated than the Irlen method due to the accessibility of the chromaticities used to calibrate the colorimeter (Wilkins & Sihra, 2001) (Appendix XV). Additionally, the Irlen method of colour sampling has been criticised for its unsystematic approach to acquiring the most effective hue of filter (Wilkins, 2003; Evans & Allen, 2016): Intuitive filters sample the colour space more systematically due to their comparatively reduced variability in luminance and saturation and approximately evenly distributed hues (Wilkins, 2003). However, the chromaticities of Intuitive filters were calculated using the CIE[1] 1976 Lu'v' (CIE LUV) chromaticity diagram, which is not recommended to accurately quantify small and medium colour differences (Witt,

[1] The 'Commission Internationale de l'eclairage' (CIE) is the International Commission on Illumination.

2007). Therefore, even the Intuitive method does not constitute a fair comparison of hue between filters because, as the colour space is not perceptually uniform, the methodology can only result in approximately similar saturation between filters. Appendix XV displays the chromaticities and range of saturations of two saturation settings of the colorimeter.

The nature of the treatment methods appear to contradict the notion that a specific hue of filter is the only filter component that reduces visual stress. The effective (chosen) colours of overlay and lens are usually unrelated (Irlen, 1991a; Wilkins, 2003; Lightstone et al., 1999) and explanations maintaining that a specific hue is required are not convincing: Irlen referred to a difference between a partial and full modification of light in the field of view when using an overlay or lens (Irlen, 1991a); Wilkins et al. (2016) referred to processes of colour adaptation when wearing the lenses. Lightstone et al. (1999) suggested that individuals may have differences of susceptibility to distortion in the visual field that require different colours. The Intuitive Colorimeter manual (Wilkins, 1997; 2001) and overlay methods (Irlen, 1991b; Wilkins, 2003) recommend forced choice if participants are undecided about their most preferred hue. This suggests that there is sometimes no observable difference in symptoms when observing one hue over another. The adjustment of saturation and luminance on the colorimeter (Wilkins et al., 1992) and the inclusion of a grey overlay (Irlen, 1991a; Wilkins, 2003) also implies that factors other than hue may be of relevance.

There are two further lens choosing methods that involve selection from fewer alternatives: Chromagen and DRT (Dyslexia Research trust) filters. Chromagen lenses have a different chosen colour of lens for each eye out of eight available choices of hue and were initially used to enhance colour perception for colour vision deficient patients (Harries, Hall, Ray & Stein, 2015). DRT filters contain two options of coloured lenses: blue and yellow. Chromagen lenses are produced by the Harries Foundation and sold in opticians and DRT lenses are produced by the Dyslexia research trust and sold in educational establishments. These filter methods are justified by their own theoretical hypotheses: Hall, Ray, Harries & Stein (2013) proposed that visual stress is the result of delayed transmission of visual information to the visual cortex in one eye relative to the other. DRT lenses are proposed to benefit reading according to the magnocellular hypothesis (section 1.5.2).

1.7.2 Diagnosis

Diagnosis relies on an assessment of the effect of the chosen filter on reading by comparison with and without the chosen filter. The Irlen method relies upon subjective reports of improved symptoms when viewing text and reading with the chosen coloured filter (Irlen, 1987, Irlen, 1991b). The Intuitive method measures improved reading speed with the chosen coloured filter using the Wilkins Rate of Reading Test (WRRT) (Wilkins, Jeanes, Pumfrey & Laskier, 1996). Although the WRRT appears to provide a more specific diagnostic criterion, it is still conducted in non-blinded, uncontrolled conditions and there is a lack of evidence to support that it diagnoses visual stress (section 1.7.3).

Subjects may be susceptible to a placebo response because they are emotionally and financially invested to find a solution to their reading problems (Howell & Stanley, 1988). The personalised consultation prior to diagnosis may enhance these personal beliefs (Dodd et a., 2017; Kaptchuk et al., 2008; Paradis & Sutkin, 2017). Additionally, questionnaires and tasks conducted prior to choosing a filter may prime subjects by referring to symptoms that are later expected to be alleviated by the chosen filter: the Irlen method conducts a reading symptom questionnaire and records the experienced symptoms and rated difficulties of visual tasks (Irlen, 1987, Irlen, 1991b); the Intuitive method elicits symptoms using the pattern glare test (section 1.7.4).

There is a commercial drive to assert that coloured filters benefit reading, and this was spread via worldwide media prior to peer-reviewed, placebo-controlled evidence (Podell, 1990). The Irlen institute makes unsubstantiated claims about the effectiveness of their filters without published evidence (Cotton & Evans, 1990; McIntosh & Ritchie, 2012). For example, Irlen filters were claimed to benefit 12-15% of the general population and 47% of people with learning difficulties (Irlen Institute, 2012). In 'The Optician' magazine, Cerium Visual Technologies refuted a systematic review that had highlighted the lack of research supporting effectiveness of chosen coloured filters beyond placebo (Griffiths et al., 2016) by drawing attention to the 'thousands of patients' who had been helped (Fitzmaurice, 2016). This type of commercialisation of the benefit of chosen coloured filters without reference to the potential placebo effect may enhance a placebo response to the benefit of coloured filters for reading.

1.7.3 The WRRT

Wilkins et al. (1996) developed the WRRT to rapidly assess visual stress and to measure the reading benefit of coloured filters to children. The task involves reading a passage of text containing lines of 15 randomly ordered, high frequency words for one minute with and without the pre-chosen filter (Wilkins et al., 1996). To isolate the perceptual effect of the filter the text was tightly spaced, which according to Wilkins' (1995) theory of visual stress (section 1.5.1) should quickly trigger the symptoms of visual stress. The test was suitable for developing readers because it only contained 15 high-frequency, known words (Wilkins et al., 1996).

The improvement in reading speed of accurately read words with a chosen coloured overlay (WRRT reading speed increase) is taken as a measure of visual stress severity, and as evidence for the reading benefit of the filter (Wilkins et al., 1996; Wilkins, Lewis, Smith, Rowland & Tweedie, 2001; Wilkins et al., 2016). A criterion of 5% reading speed improvement with a chosen coloured filter on the WRRT is often used to diagnose visual stress (Monger et al., 2015), although a more conservative criterion of 15% has recently been suggested (Wilkins et al, 2016). Prevalence estimates of visual stress in the population appear to depend upon the criterion (Table 1.1) and have been suggested to reflect that visual stress exists on a continuum (Wilkins et al., 2016).

Table 1.1

Reported criteria and prevalence of improved reading speed with a chosen filter for adults (below), and children (next page).

Criterion	Prevalence	Sources	Participant ages (years)
>5%	34-58%	Evans and Joseph (2002); Henderson, Tsoga & Snowling (2013)	18-44 19-30
>8%	46 %	Henderson et al. (2013)	
>10%	35 %	Henderson et al. (2013)	
>25%	2%	Evans and Joseph (2002)	

Criterion	Prevalence	Sources	Ages (years)
>5%	22-43%	Kriss and Evans (2005);	7-12
		Singleton and Henderson (2007);	7-17
		Wilkins et al. (1996)	8-11
		Wilkins et al. (2001)	6-8
>8%	18-28%	Kriss and Evans (2005);	
		Singleton and Henderson (2007)	
>10%	13-22%	Kriss and Evans (2005);	
		Singleton and Henderson (2007)	
>25%	5%	Wilkins et al. (2001)	

WRRT reading speed increases were claimed to measure visual stress severity and a reading benefit of filters because they correlated with longer periods of voluntary use of the chosen filter (Wilkins et al, 1996; Jeanes et al, 1997; Wilkins and Lewis, 1999) and colour re-test consistency (Wilkins et al, 2001). However, these effects may be moderated by placebo effects, and there is no convincing evidence to support the claim that WRRT reading speed increases directly correlate with measures of visual stress severity or reading benefit. Furthermore, the re-test reliability of the WRRT, necessary for any use of a diagnosis method or intervention, is uncertain.

1.7.3.1 A measure of visual stress?

Increased WRRT reading speeds were not found to correlate with visual stress severity when measured by the number of reported symptoms of visual stress when reading (Lightstone, Lightstone & Wilkins, 1999; Hollis & Allen, 2006; Henderson et al., 2013). It has also been found that some individuals with no reading symptoms of visual stress demonstrate dramatically improved reading speeds on the WRRT with a chosen coloured filter (Wilkins, 2003, p18).

Increased WRRT reading speeds correlated with visual stress severity when measured by the number of reported perceptual symptoms when viewing the pattern glare test (Hollis and Allen, 2006), but perceptual symptoms were recorded immediately prior to conducting the WRRT, and so increased reading speeds may be an induced placebo response.

1.7.3.2 A benefit to reading?

A relationship between WRRT reading speed increases and a measure of reading benefit has not been published. Wilkins (2002) cited unpublished data indicating that increases in WRRT reading speed predict a benefit to natural reading. A relationship between reading speed increases on the WRRT and a silent reading sentence task was found for children with reading difficulties, but not for unselected children. The lack of significance in the latter was claimed to be due to variation in reading speed attributed to the task requiring comprehension, but if this was the case, such variability would likely extend to the children with reading difficulties. An alternative explanation could be that the children with reading difficulties had a stronger placebo response. Placebo effects are thought to be most likely to occur when they contribute towards achieving a personal goal (Geers, Weiland, Kosbab, Landry & Helfer, 2005) and so subjects with reading problems may be more motivated to read better than those without.

Henderson et al. (2013) found coloured filters improved performance on the WRRT, but had no significant effect on reading rate or comprehension for meaningful text. Wilkins et al. (2016) suggested that this discrepency could be attributed to increased noise in the data from the comprehension condition: uncontrolled variability associated with memory and visual search processes.

1.7.3.3 Re-test reliability

Both reading speed (*r* = .83) (Wilkins et al., 1996) and reading speed improvement (*r* = .72) (Wilkins et al., 2001) were reported to highly correlate when measured eight months apart. However, this only describe the behaviour of the 40% of participants who continued to use their overlays and believed that they benefitted their reading. Henderson et al. (2013) reported the re-test reliability of WRRT reading speeds over one month was high for dyslexic (*r* = .94) and control (*r* = .89) participant groups, but the re-test reliability of reading speed improvement was low (*r* = .40 and *r* = .38 respectively).

1.7.4 The pattern glare test

The Intuitive method sometimes conducts the pattern glare test to contribute towards a diagnosis of visual stress (Monger, Wilkins & Allen, 2015). This is because Hollis and Allen, (2006) had reported that the number of symptoms on this test correlated with WRRT reading speed increases and recommended that the pattern glare test measured visual stress. However, Hollis and Allen (2006) also reported that WRRT reading speed increases did not correlate with self-reported reading symptoms and concluded that self-reported reading symptoms did not measure visual stress. This is concerning because such reading symptoms are used for a filter test referral and have been documented to play a central role in a visual stress diagnosis (Evans et al., 2016; Irlen, 2010). Considering the lack of convincing evidence supporting the use of the WRRT to diagnose visual stress, an alternative interpretation could be that improved reading speeds on the WRRT and the pattern glare test correlate with one another but do not diagnose visual stress.

1.8 Dyslexia

The Diagnostic Statistical Manual of Mental Disorders (DSM-V code 315.00) describes dyslexia to involve poor word recognition, decoding and spelling abilities, which cannot be accounted for by mental age, visual acuity or inadequate schooling (American Psychiatric Association, 2013). A single coherent theoretical framework does not account for all behavioural traits of dyslexia (Elliott & Gibb, 2008; Ramus & Ahissar, 2012). There are three main theories: the phonological, cerebellar and magnocellular theory (Ramus et al., 2003). In general, dyslexia involves impaired or delayed phonic acquisition (the association of graphemes to phonemes or letters to sounds), and this is sometimes accompanied by visual, auditory and motor difficulties (Ramus et al., 2003).

1.8.1 Phonological theory

The phonological theory proposes that the cognitive basis of dyslexia is a fundamental weakness in phonological awareness and the processing of phonological representations (Vellutino, 1979; Snowling, 2001). The development of phonological skills involves the storing and retrieval of grapheme-phoneme associations. Individuals with dyslexia consistently perform poorly on tasks such as phonological awareness, rapid naming and verbal short-term memory (Ramus, 2003; Ramus & Ahissar, 2012), and variance in literacy skills has been most accurately explained by phonological variables (Saksida et al., 2016). The importance of phonics instruction for children's development of reading and writing skills was emphasised at the 2009 Rose Review and is now the most intensively investigated reading intervention. Phonics instruction is cited as the only reading developmental approach to have shown a statistical improvement on reading performance (Galuschka et al, 2014).

1.8.2 Cerebellar theory

The cerebellar theory postulates that phonic acquisition is affected by a dysfunction in the cerebellum. The cerebellum was activated in automatic motor tasks (Jenkins, Brooks, Nixon, Frackowiak & Passingham, 1994). Impaired motor skills and automisation were posed to delay phonic acquisition by affecting development of handwriting, articulation and the consolidation of the grapheme-phoneme correspondence (Nicholas, Fawcett & Dean, 2001). There is little evidence to support these ideas (Ramus, 2003).

Children with dyslexia were shown to perform significantly poorer on tasks associated with the cerebellum than children without dyslexia. These involved balance (Nicolson and Fawcett, 1990), estimation of time (Nicolson, Fawcett and Dean, 1995) and a battery of motor tasks involving coordination and muscle tone (Fawcett, Nicolson & Dean, 1996; Fawcett & Nicolson, 1999). Brain imaging studies also revealed anatomical, metabolic and activation differences in the cerebellum of dyslexics when compared to controls (Rae et al., 1998; Brown et al., 2001; Leonard et al., 2001). Later studies have not repeated these findings and confirmed that motor difficulties are restricted to a subset of people with dyslexia (Ramus, 2003; Danelli et al., 2017).

1.8.1 Magnocellular theory

An impaired magnocellular system was posed to explain dyslexia. The basis of this hypothesis depended upon the smaller and disordered magnocellular layers of the LGN in dyslexic students documented in post-mortem studies of dyslexic subjects (Galaburda et al., 1985; Livingstone, Rosen, Drislane & Galaburda, 1991).

Measures of contrast sensitivity and coherent motion were posed to provide evidence for the hypothesis. Contrast sensitivity, measured by the contrast at which sinusoidal gratings could be detected, was poor at low spatial frequencies and high at high spatial frequencies for dyslexic subjects when compared to controls (Lovegrove, Martin, Blackwood & Badcock, 1980). Coherent motion sensitivity, measured by the width of moving dots needed for detection, was found to be poorer in dyslexic subjects than controls (Cornelissen et al., 1994), and particularly poor at higher dot densities (Talcott, Hansen, Elikem & Stein, 2000). Training in coherent motion tasks was reported to improve word reading performance (Cornelissen et al., 1998a) and letter position encoding (Cornelissen et al., 1998b).

These studies were argued not to support the magnocellular hypothesis since the tasks are not known to trigger the magnocellular system specifically over the parvocellular system (Skottun 2000; Skottun & Skoyles, 2007) and were found to have little correlation to each other (Goodbourn et al., 2012). Also, visual and auditory impairments in people with dyslexia have been found not to be confined to tasks involving the magnocellular pathway (Amitay, Ben-Yehudah, Banai & Ahissar, 2002).

1.8.1 Visual stress and dyslexia

In general, dyslexia refers to impeded reading performance outcomes and visual stress refers to perceptual symptoms when reading. It was suggested that people with visual stress are more likely to develop dyslexia: unpleasant visual stress symptoms when viewing text could negatively affect the process and enjoyment of learning to read (Irlen, 1991a; Wilkins, 2003). However, a relationship between the two conditions has not been reliably established and the causal contribution of visual stress to dyslexia is unknown.

Conflicting reports exist regarding the ratio of people with dyslexia who have visual stress. This may be partly due to an absence of a reliable and validated method of diagnosing visual stress when reading. When diagnosed by the pattern glare test, visual stress was diagnosed in an equal proportion of dyslexic and control participants (Saksida et al. 2016). However, when diagnosed with increased WRRT reading speeds with a chosen filter, significantly more children (Kriss & Evans, 2005; Singleton & Henderson, 2007) and undergraduates (Singleton & Trotter, 2005; Henderson et al., 2013) with dyslexia were diagnosed with visual stress than controls participants without dyslexia.

The different outcome in these studies may be due, in part, to the method by which participants were categorised to be dyslexic: with the exception of Saksida et al.'s (2016) study, participants were pre-diagnosed with dyslexia by an educational psychologist but the diagnostic criteria is not fully described. This is necessary because methods of diagnosing dyslexia have changed in response to the various hypotheses over the years (Le Jan et al., 2011).

These results may also only reflect an enhanced placebo effect to coloured filters of people with dyslexia because these studies were non-blinded and not sufficiently controlled: participants aware of which condition contained their chosen coloured filter. Additionally, the WRRT has not been demonstrated to represent reading of naturalistic text. This could be explained by an increased likelihood of exposure to coloured filters in dyslexia diagnoses and therefore increased expectation in relation to their effects (Ritchie, Della Sala & McIntosh, 2012; Henderson et al., 2013).

If underlying reading difficulties (other than visual stress) are present, such as delayed phonic acquisition associated with dyslexia, chosen coloured filters are claimed not to lead to reading performance or developmental improvement without explicit intervention for those reading difficulties (Blaskey et al., 1990; Irlen, 1991a; Wilkins, 2003; Wilkins et al., 2016). This has implications regarding studies involving the effect of chosen coloured filters on reading because even if symptoms of visual stress are removed, reading performance may not improve due to the underlying reading difficulties.

1.9 Do chosen coloured filters benefit reading beyond placebo?

In this section, studies that investigate whether chosen coloured filters benefit reading beyond placebo are categorised according to which of the two main methods (Irlen or Intuitive method[2]) were used to acquire the chosen filter. Within these sections, studies are also categorised according to whether they were blinded and whether the participants that were used were children or adults.

Non-blinded studies are likely to be vulnerable to placebo effects. For example, it is likely that such studies would suffer from demand characteristics: having chosen a filter to benefit their reading for treatment, participants may be aware of an experiment's aim to assess the reading benefit of their chosen filter. There may be a novelty effect associated with the use of previously chosen filters; this appears to be supported by the high percentage of participants who stopped using their filter after choosing it in their diagnosis; 48 - 59% children stopped using their chosen filter after 2 – 12 months of use (Wilkins et al., 1996; Jeanes et al., 1997; Wilkins et al., 2001; Ritchie et al., 2012).

Unlike adults, children are developing readers. Consequently, visual stress reduction is likely to affect the reading performance of the two populations differently. Accordingly, measures used to demonstrate a benefit to reading of chosen coloured filters are different for children and adults; reading development measures can demonstrate a reading benefit to children but are not appropriate for adults with developed reading skills. In addition, Wilkins' (2003) has claimed that the high percentages of children who stop using their filter is due to symptoms of visual stress not being permanent.

[2] Irlen and Intuitive methods of acquiring the chosen filter are described in section 1.7.

1.9.1 Irlen filters and the evidence

1.9.1.1 Non-blinded studies with children

Tyrrell, Holland, Dennis and Wilkins (1995) compared audio-recorded oral reading speeds of 46 children, aged 12-16 and diagnosed with Irlen syndrome, with a chosen and clear (control) coloured filter over a 15-minute reading period. Texts were individually chosen and reading speeds were calculated using the number of orally read syllables. There were significantly less reported perceptual distortions when the chosen coloured filter was used. Reading speeds were also found to be significantly more with the chosen coloured filter after 10 minutes, rather than five minutes, of reading. Tyrrell et al (1995) concluded that, over a ten minute time period, a chosen coloured filter had reduced visual stress symptoms and increased reading speed. However, only three one-minute intervals were used to calculate a mean reading speed of the first and last five minutes of the reading and so did not represent the full first and last five minutes. Tyrrell et al. (1995) suggested that placebo effects were unlikely because the difference in reading speeds only emerged after 10-15 minutes of reading. However, there is no evidence to suggest that placebo effects are time-dependent.

Robinson & Conway (1994) reported significantly improved reading rate and comprehension, but not accuracy, in 29 children who had been diagnosed with Irlen syndrome and had used chosen coloured filters for four months, when compared to 31 children who had received a negative diagnosis of Irlen syndrome and had not been using chosen coloured filters. Noble et al. (2004) reported that 31 grade three children diagnosed with Irlen syndrome made significant gains in reading development measures (reading rate, accuracy, fluency and comprehension) after three months of using chosen filters, compared to 40 grade three children who were diagnosed with Irlen syndrome but did not use the chosen filters. The accelerated reading development was reported to tail off after the first three months, which was attributed to participants meeting their age expected reading ability. The group who had not been using their chosen filters then used their filters for three months and demonstrated significantly improved reading achievement in all measures.

1.9.1.2 Non-blinded studies with adults

Blaskey et al. (1990) compared measures of word recognition, reading comprehension, rate of reading random letters and reading accuracy in 30 children and adults (aged 9-51). They found no measure of improvement by the chosen over given filter. Robinson and Conway (2000) reported significantly improved reading accuracy, comprehension and speed in 15 adults, aged 18-62, after three months of using their chosen coloured filter when compared to nine adults who had been diagnosed with Irlen syndrome but read without them and nine adults who had received a negative diagnosis. As had been reported in children (Noble et al., 2004) these accelerated reading performance measures tailed off after the first three months. The group who had not been using their chosen filters then used their filters for three months and demonstrated significantly improved reading achievement in reading accuracy and comprehension, but not reading speed.

In a neuroimaging study, Kim, Seo, Ha and Kim (2015) reported that 15 subjects, aged 13-42, who had been diagnosed with Irlen syndrome, demonstrated increased activation in the left, middle and superior temporal gyri during silent sentence reading with their chosen filter. Since these areas have been associated with sentence comprehension, this finding was suggested to support effectiveness of filters for improving reading (Kim et al., 2015; Irlen-Institute, 2017c). However, there was no mention of pseudorandomising the order of the filter-no filter condition, so this effect could simply be an order effect.

1.9.1.3 Blinded intervention studies with children

Placebo controlled trials have compared oral reading errors, reading speed, comprehension and reading accuracy of a chosen and a disguised placebo lens that were not distinguishably different (Robinson & Foreman, 1999a; Robinson & Foreman, 1999b). Children, aged 9-13, diagnosed with Irlen syndrome showed equally improved oral reading errors (Robinson & Foreman, 1999a) and reading speeds (Robinson & Foreman, 1999b) when using the chosen and placebo lens, but significantly improved reading comprehension and accuracy with their chosen filter (Robinson & Foreman, 1999b).

Ritchie et al. (2011) found no effect of a chosen coloured filter when 75 primary school children did not know that their chosen colour was the effective treatment. Participants in this experiment were led to believe that the colour choosing process was incomplete after they had selected their chosen colour. Apart from three participants who had previously used coloured filters, WRRT reading speeds did not improve with the chosen coloured filter relative to an assigned coloured or colourless filter (Ritchie et al., 2011). Wilkins et al. (2016) suggested that this negative finding may have been due to the fact that the participants did not have visual stress because the Irlen method had been used. However, Wilkins and colleagues accepted voluntary sustained use of the filter as a measure of their effectiveness or diagnosis (Jeanes et al., 1997, Wilkins et al., 1996; Wilkins et al., 2016), and in Ritchie et al. (2012) follow-up study, 18 participants who had habitually used their chosen coloured filters for one year still showed no significant improvement in WRRT reading speed or reading development when using that filter, when compared to a control group. Despite the small sample size, the authors claimed that there was adequate power to detect effect sizes of the type implied by Irlen (2010) and Nobel et al. (2004) and concluded that the chosen coloured filters did not benefit reading. McIntosh and Ritchie (2012) went on to suggest that previous studies showing boosted reading performance were placebo effects. Ritchie et al.'s (2011;2012) studies appear to demonstrate that reading speed does not increase on the WRRT with a chosen coloured filter if neutral instructions are applied.

1.9.2 Intuitive system and the evidence

The evidence used to diagnose visual stress, report prevalence and conclude that coloured filters benefit reading has largely involved the use of the WRRT in non-blinded conditions. This is illustrated by the different conclusions and approaches of two systematic reviews. One systematic review claimed that Intuitive filters 'alleviate symptoms or improve performance in people who suffer from visual stress' (Evans and Allen, 2016, p216) and considered the WRRT to be 'outcome appropriate' citing its use in 12 out of the 13 studies. Another systematic review excluded these studies on the basis that the WRRT had limited external and construct validity (Griffiths et al., 2016). Griffiths et al. (2016) concluded there was not enough evidence to support the assertion that coloured filters benefit reading.

1.9.2.1 Non-blinded studies with children

Non-blinded studies reporting WRRT reading speed increases with a chosen filter (e.g. Wilkins et al., 1996; Jeanes et al., 1997; Wilkins et al., 2001; Kriss & Evans, 2005; Singleton & Henderson, 2007; Allen, Gilchrist, & Hollis, 2008) have been interpreted as evidence that chosen coloured filters alleviate symptoms of visual stress and benefit reading (e.g. Wilkins et al., 2016; Evans and Allen, 2016).

Some attempts have been made to rule out the possibility that increased WRRT reading speeds are a placebo effect by attempting to generate a placebo response to an allocated control filter. Leading instructions, implying that the control filters were effective reading interventions, were applied to a grey filter for 26 children, aged 6-15 years (Wilkins & Lewis, 1999), and to a yellow filter for four adults (aged 18-40 years) and 29 children (aged 7-14 years) (Bouldoukian, Wilkins & Evans, 2002).

Although these experiments reported that WRRT reading speeds were significantly quicker for the chosen rather than control filters, the level of induced expectation between the filter conditions was unlikely to be equal because the act of selecting a chosen treatment can bring about placebo effects, as has been demonstrated by reducing pain perception (Rose, Geers, Rasinski & Fowler, 2012). Wilkins, Sihra and Myers (2005) reported increased WRRT reading speeds when five subjects, aged 11-17 years, read on a chosen hue relative to a selection of assigned hues. The authors argued that the subjects were unlikely to have remembered their chosen hue because in a second study, 22 participants were not able to accurately replicate an assigned hue on the colorimeter used to select the chosen hue. However, the ability to reproduce an assigned colour does not equate to the ability to remember a chosen one and so these results may have also been due to a placebo effect.

1.9.2.2 Non-blinded studies with adults

Evans & Joseph (2002) repeated Wilkins and Lewis' (1999) experiment for 113 university students by applying leading instructions to allocated yellow, control, filters. Although WRRT reading speeds were significantly quicker for the chosen filter, the act of selecting a chosen treatment may have brought about larger placebo effects than those induced by the leading instructions for the control filter.

1.9.2.3 Blinded intervention studies with children

Placebo controlled trials have compared symptoms and reading speeds of a chosen and a disguised placebo hue of lens that were not distinguishably different. The control hue was calculated to be six just noticeable differences from the chosen hue; theoretically enough to reduce perceptual benefit but still be recognised as the chosen hue (Wilkins et al., 1994). Wilkins et al (1994) reported that school children (aged 11-12) made fewer diary entries of symptoms of headache and eye-strain on days when their chosen lens was worn, but the analysis was criticized for its high attrition rate, with only 53% of the sample completing the experiment (Griffiths et al, 2016). In contradiction, a similarly designed between-subjects experiment found that children, aged 7 -11 years, reported equally reduced symptoms by questionnaire with the chosen and disguised placebo condition over a month (Mitchell, Mansfield & Rautenbach, 2008). Relative to no filter, both studies reported equally improved reading speed and comprehension, recorded using the Neal Analysis of Reading when the chosen and placebo lenses were worn (Wilkins et al., 1994; Mitchell et al., 2008). Mitchell et al. (2008) interpreted their results to be a placebo effect.

1.9.3 Chromagen and DRT systems and the evidence

There are only a limited number of peer-reviewed papers which have researched the effectiveness of the Chromagen and DRT systems. These have only been undertaken with children in non-blinded conditions and so have not been suitably placebo controlled.

Without a masked control filter, a chosen yellow or blue DRT filters was reported to increase performance in motion sensitivity, convergence, accommodation (Ray et al., 2005) and reading ability (Ray et al., 2005; Hall et al., 2013; Harries et al., 2015). Harris and MacRow-Hill's (1999) reported that children improved their WRRT reading speeds with a chosen Chromagen lens relative to a blue or clear lens. However, this was contradicted by Cardona et al.'s (2010) study, which reported that adolescents (14-17 years) showed no improved WRRT reading speeds with Chromagen lenses when compared to clear lenses.

In a comparison between the two systems, there was no significant difference found between reading performance and symptom measures of 30 children (7 – 10 years) (Hall et al., 2013; Harris et al., 2015). However, these studies did not address the effectiveness of the coloured filters when compared to baseline.

1.10 Visual Search

Visual search performance has been used to explore the magnocellular theory of dyslexia (e.g. Iles et al., 2000), but there is limited research applying visual search performance specifically to visual stress and the effectiveness of pre-selected coloured filters. A chosen coloured overlay significantly improved letter search performance when compared to a clear overlay (Tyrrell et al., 1995) and grey overlay (Newman Wright, Wilkins, & Zoukos, 2007), but the latter study reported significantly less improvement on re-test. In contrast, no significant improvement was found for number search performance (Allen et al., 2008). These experiments hold the same limitations as reading experiments as they do not fully account for placebo, with participants aware of the condition involving their chosen coloured filter.

1.11 Is a specific hue necessary?

The specificity of hue is implied to be the effective filter component that reduces visual stress. Irlen's (1991a) referred to difficulty processing certain 'light frequencies' and Wilkins directly emphasised the importance of a specific hue of filter in reducing visual stress (Wilkins, 2003; Wilkins et al., 2005; Smith & Wilkins, 2007): A CIE LUV colour difference margin of .065 from the chosen filter was deemed to be effective (Wilkins et al., 1994). However, the lack of evidence and apparently poor re-test reliability of colour choice appears to contradict this notion.

1.11.1 A lack of evidence

The evidence reviewed in the blinded intervention studies of the Irlen (section 1.9.1.3) and Intuitive method (section 1.9.2.3) do not support the claim that hue specificity is critically important in reducing visual stress and benefitting reading. Non-blinded studies with the WRRT (Wilkins et al., 2005; Smith & Wilkins, 2007) arguably cannot be used to draw conclusions about overall reading benefit. Regardless, this evidence does not appear to support a specific colour to be needed. Wilkins et al (2005) measured reading speed improvements of five subjects who reported visual stress on a selection of hues before and after they used their chosen filter. Participants used their chosen filter between the testing, so it is unlikely that the memory of their chosen colour did not affect their colour choice. Despite this, one participant chose a different hue (blue) to that which had been previously chosen (green). In another study, Smith & Wilkins (2007) concluded that 10 colour choices were needed to find the effective hue because a chosen filter from five colour choices did not retrieve improved reading on the WRRT, but this was not a fair comparison because the filters were different sizes. The margin of six just-noticeable differences (6JND) from a chosen hue was reported to render a chosen tint ineffective (Wilkins et al., 1994), but was based upon participants' subjective responses, after having been instructed to find one specific hue to reduce their visual stress symptoms. The intended behaviour, that only one hue would be helpful had been communicated prior to the task, and so this margin may be based upon a placebo effect.

1.11.2 Poor re-test reliability of colour choice

Colour choices of overlays are claimed to have high test-re-test reliability (Evans & Allen, 2016), but when chosen colours of filters have been recorded, poor repeatability in colour choice appears to be evident. Suttle, Barbur & Conway, 2017) reported that 11 out of 21 adults and children chose a completely different coloured overlay 2-57 days apart, and only seven of these chose the same colour (Suttle et al., 2017). This was estimated to equates to an average test-retest colour space differences of more than six JNDs (Elliot and Wood, 2017). Jeanes et al.'s (1997) claimed that children chose 'similar colours' three months apart, but this study was later analysed to described colour differences equating to a colour difference of pink-to-blue or rose-to-yellow (Elliot and Wood, 2017). Coloured filter testing is recommended to be re-tested at least every two years (Wilkins, 2003; Irlen, 2010) implying that the choice of colour is not robust over time.

1.11.3 Factors affecting colour choice

Participants' chosen coloured filter have rarely been specified in the literature, especially regarding studies involving adults and lens colour. Studies involving children and overlays appear to reflect a wide distribution over the choices available. Wilkins et al. (2001) reported that the most frequently chosen colours of overlay were 'rose' and 'aqua' and these were chosen by 10% of the sample (Wilkins et al. 2001). In contrast, Veszeli and Shepherd (2019) reported that the most frequently chosen colours, chosen to improve the appearance of the text, were yellow and pink and these were chosen by 38% of the sample. Kriss and Evans (2005) reported that ‘mint green’ and ‘grey’ were most frequently chosen. However, since these were the last two option available and chosen by 57% of the sample, Kriss and Evans (2005) suggested that participants in this study may have been tempted to choose a remaining option before the overlay procedure was completed.

Colour choice of filters may be driven by factors other than visual stress symptoms. For example, it is possible that factors affecting colour preference may affect the colour of filter that is chosen: there is evidence to suggest that stereotypical gender-based colour preferences are associated with filter colour choice. Conway, Evans, Evans and Suttle (2016) reported that more males chose stereotypical male colours such as blues and greens, and more females chose stereotypical female colours, such as pinks and purples. However, the effect was small (p = .04) and only observed when the three age groups of participants were combined into a larger participant group of 238 participants (aged 7-65). Also, the effect was only observed for chosen lenses, not overlays: the authors suggested that lens colour choice may have been more influenced by gender-based colour preferences than overlay colour choice due to participants being influenced by how they would appear when worn.

1.12 The role of contrast

Coloured filters may reduce visual stress by reducing the contrast of the underlying text. Measurements of text contrast and contrast reduction resulting from filters have rarely been specified in the literature. Using Weber contrast (1), Wilkins et al. (2001) stated that overlays reduce the contrast of text by 'about 2%' under directional lighting and 'generally less than 5%' under diffuse illumination. However, given that the luminance and saturation of Intuitive filters are not constant, text contrast must also vary between filter choices; these variations have not been specified.

Reduced text contrast was initially thought to reduce perceptual distortions in text[3] for poor readers (Meares, 1980), but this idea was succeeded, and apparently overshadowed, by Irlen's (1983) assertions that a specific colour of filter was necessary. Wilkins has stated that the reduction in contrast from filters may be the important component for 'a few patients' (Wilkins, 2002), but asserts that this is rare, and the specific colour of filter is crucial component (Wilkins, 2003).

The effect of reduced contrast on visual stress was dismissed when neither assigned nor chosen grey overlays increased reading speeds on the WRRT (Jeanes et al., 1997; Wilkins et al., 2001). However, this overlooked the uncertainty over whether the WRRT reading speed increases represented a reading benefit that generalises over all aspects of reading performance (*e.g.* comprehension, ability to sustain attention) or a placebo effect. Additionally, Simmers, Bex, Smith and Wilkins (2001) found no difference in contrast and motion sensitivity performance in children, aged 12-14, who had been diagnosed with visual stress (by their preference to read with a chosen coloured filter) and control subjects. However,

[3] These perceptual distortions later became known as symptoms of visual stress

these tasks measured the detection of contrast rather than the contrast threshold that was needed for improved performance.

Symptomatic evidence indicates that visual stress is both triggered and reduced by increasing and reducing the contrast of stimuli in the visual field. Allen, Evans and Wilkins, (2012) stated that medium spatial frequencies patterns at high contrasts trigger visual stress and this is in line with the stimuli used: the 'pattern glare test' (Figure 1.2) and the Irlen equivalent[4] (Irlen, 1991; Wilkins, 2003). Interestingly, the pattern glare test was rated to be more comfortable when it was presented in low contrast conditions for people with visual stress (Fernandez & Wilkins, 2008). High contrast environments (including glare) have also been reported to trigger visual stress (Irlen, 1991a). Reduced symptoms of visual stress were reported when text was observed in a text window (Wilkins & Nimmo-Smith, 1984) which may be because the RMS contrast of the page had been reduced by covering most of the text.

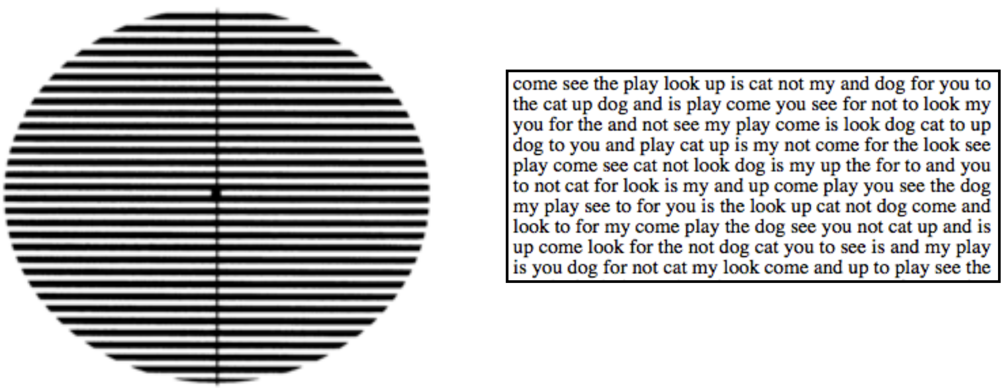


Figure 1.2. Stimuli used to trigger visual stress. Left: the pattern glare test stimuli (Hollis & Allen, 2006). Right: the WRRT (Wilkins et al., 1996). Both stimuli are presented with higher contrast than normal text. Not actual size.

[4] A page of Dutch, bolded text (Irlen, 1991a)

Some studies have reported that participants with visual stress equally improved their performance with chosen and placebo filters relative to no filter, which may inadvertently support the hypothesis that reducing text contrast reduces visual stress. Improved performance measures when using both chosen and placebo filters have included improved silent reading for comprehension (Wilkins et al., 1994), reduced visual stress symptoms (Mitchell et al., 2008) reduced large micro-fluctuations during steady-state accommodation (Simmers, Gray & Wilkins, 2001) and improved reading development (Robinson & Foreman, 1999a; Mitchell et al., 2008). Mitchell et al. (2008) concluded their results to reflect a placebo effect but another interpretation is that they are an effect of reduced contrast because the chosen and placebo lenses reduced the text contrast by similar amounts.

Some research has suggested that that some dyslexic children performed tasks better when the contrast of text was slightly reduced and these findings may be due to participants having visual stress. Badcock and Lovegrove (1981) reported reduced visible persistence and improved comprehension scores when children read on medium grey paper. Giddings and Carmean (1989) reported that the comprehension scores of dyslexic children was 10% higher on medium-grey paper than on white paper. Wlliams & LeCluyse, (1990) demonstrated that visual search and reading speed performance of 'disabled readers' was significantly improved when text contrast was reduced by blurring. However Hogben, Pratt, Dedman and Clark (1996) did not replicate Wlliams & LeCluyse's (1990) finding.

O'Brien, Mansfield and Legge (2000) reported that reduced text contrast did not impact reading speed in dyslexic children differently to typical readers and concluded that the effect did not extend to reading performance. However, each trial contained a stimulus of one sentence (300 characters) which did not appear to represent the contrast level of a page of text. It is also plausible that each trial was not viewed long enough to allow visual stress to develop and that the participants in this study did not have visual stress.

Williams, May, Solman & Zhou (1995) demonstrated that poor readers performed better on a search task when the contrast was reduced. The task was to search for the letter 'z' within 18 lines of six angular (e.g. E, W, I) letters. The normal condition had a black text background and white letters, and the experimental condition reduced the contrast be reducing the luminance of the letters to grey letters. Williams et al. (1995) suggest that magnocellular defecits could be alleviated by reducing the contrast of text

1.13 Thesis aims

The first aim of this thesis was to establish methods that can be used to investigate the effects of chosen coloured filters on reading beyond placebo in adults. The second aim was to determine if there was an effective component of chosen coloured filters by investigating the roles of hue, placebo and contrast. The focus of this thesis was on reading performance of adult participants because they are more likely to have developed reading, less likely to be affected by placebo than children and are easier to recruit.

Chapter 2: General methods

2.1 Photometer readings

Luminance (cd/m^2) and chromaticity (CIE, 1931, Xxy) of stimuli were measured using a Konica Monolta CS-100A photometer with 9° field of view and 1° circle measurement area. Prior to taking measurements, the aperture was focussed on a central black square of 20x20 pixels that was displayed on the screen. Each measurement of the photometer was taken after the trigger was held for the length of time it took the luminance value in the view finder to change twice.

2.2 Screen specifications

Screen stimuli were displayed on a Lacie electron 19 blue IV monitor in a laboratory with no other light sources. The screen was 40.4 x 30.3cm with a resolution of 1024x768 and a refresh rate of 100hz. The stimuli were produced using Matlab and the Psychtoolbox extensions (Brainard, 1997; Brainard, Pelli & Robson, 2002) on a 13 inch MacBook Pro. Unless otherwise stated, text was presented with a font size of 12 pt.

2.2.1 Screen brightness and contrast settings

The range in screen luminance was measured at different screen brightness and contrast settings in order to obtain the most appropriate settings for replicating text stimuli in in office lit conditions (Figure 2.1). The photometer was secured to a tripod at a 90˚ angle and 25cm from the screen with close-up lens attached. The area captured by the photometer had a diameter of 55 pixels located at the centre of the screen. The meter's response speed was set to slow.

At each brightness and contrast setting, the maximum and minimum luminance was measured and the difference was calculated. As expected, the range in screen luminance increased as the screen brightness and screen contrast settings increased. The screen contrast was set to 100 to optimise the level of contrast that could be achieved on screen: this would help achieve the level of contrast required for replicating text. The screen brightness setting was set to 60.7: at its maximum, this setting outputted the background luminance (119.625 cd/m^2) that had been determined to represent the white page of text in office- lit conditions (section 2.3).

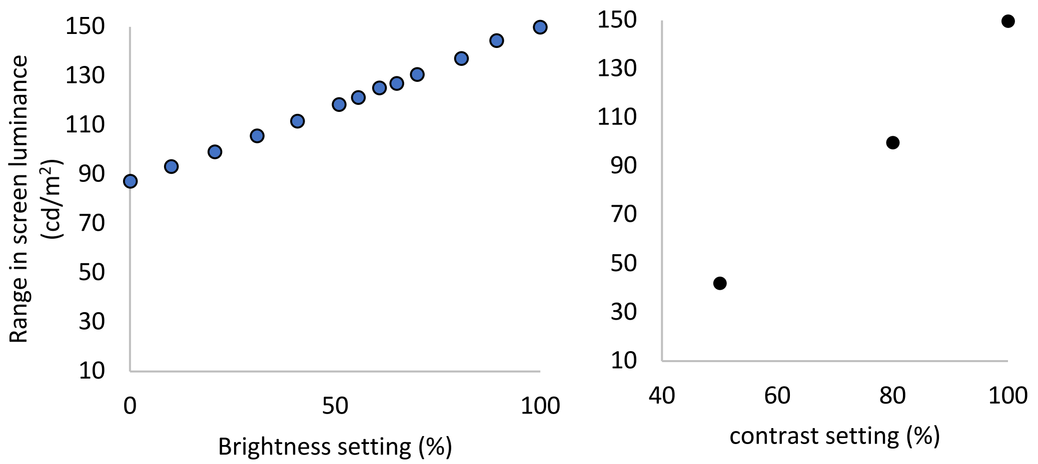


Figure 2.1 Range in screen luminance, calculated by the difference between the maximum and minimum screen luminance at different screen brightness and contrast settings. Left: The contrast setting was held constant at 100 and the brightness setting was manipulated. Right: The brightness setting was held constant at 100 and the contrast setting was manipulated.

2.2.2 Variable screen luminance

The screen luminance was variable over time and area: stable luminance occurred after a 40 minute 'warm up' period and luminance was higher towards its centre. These features are typical of CRTs and recommendations to switch on the monitor 45 minutes before use and to present stimuli symmetrically in the centre of the screen (Metha, Vingrys & Badcock, 1993) were followed.

The variable screen luminance was not considered to be problematic because the average luminance measured across the screen (Figure 2.2) equalled the average luminance that was calculated by Matlab (to within 0.1 cd/m^2) (section 2.3). Therefore, after gamma correction, Matlab calculations of luminance and contrast would accurately represent the luminance and contrast of presented stimuli (to within 0.1 cd/m^2). Additionally, the variance had some ecological validity because reflected luminance of paper in office conditions also varies across the page.

A final consideration was that the luminance of the central area of the screen changed when a different border was used. To account for this, all measurements and calibrations were conducted with the border that would be used for each experiment.

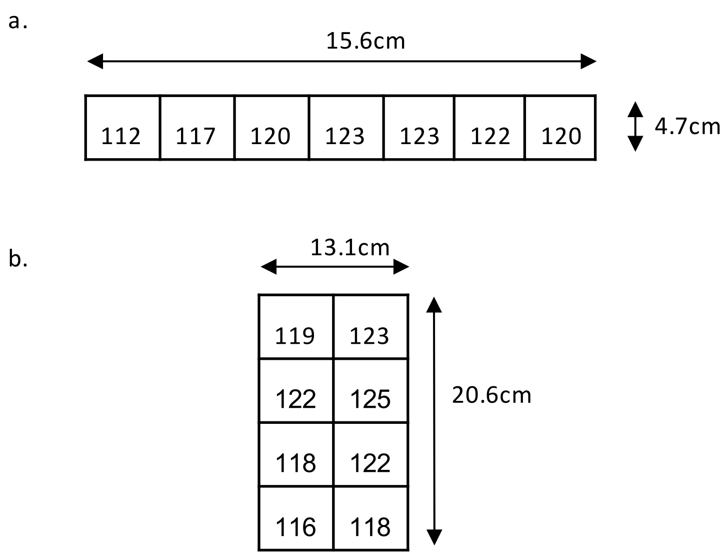


Figure 2.2. Variation of screen luminance (cd/m^2) across the central area of the screen that text stimuli would be displayed. Measurements were taken after a 45 minute ‘warming up’ period. Each luminance is the average of 10 photometer measurements which varied by < 0.1%. a: Experiment 1: average screen luminance was 119.625 cd/m^2. b: Experiments 3-6,8: average screen luminance was 120.375.

2.3 Calibrations

2.3.1 Adjacent pixel nonlinearity

Calibration for adjacent pixel nonlinearity may have been necessary for the displaying of text on a CRT, and so was explored. Adjacent pixel nonlinearity occurs when displayed images on a CRT have high horizontal spatial frequency and contrast (Klein, Hu and Carney, 1996). The electron beam moves along the raster from left to right, changing its voltage intensity according to the coded luminance designated to each pixel. If the changes from low to high luminance are rapid, the voltage required for higher luminance pixels can be attenuated, causing reduced screen luminance.

To demonstrate adjacent pixel non-linearity, the luminance of a black-white alternating, striped image was measured when the alternations were along (horizontal) and perpendicular (vertical) to the raster, and when stripe width was 5, 10 and 20 pixels respectively. Each measurement was taken at a distance so that the same image (six black and white lines) was in the collecting area. Vertical gratings were found to be 2.9 cd/m^2 darker than horizontal gratings when stripes were five pixels wide but not at other stripe widths. This indicated that adjacent pixel non-linearity occurred when stripe width was five pixels, but not at higher stripe widths (see Table 2.1).

Table 2.1.

Luminance of horizontal and vertical gratings when presented on screen at different stripe widths and distances

Stripe width (pixels)	Distance (mm)	horizontal grating luminance (cd/m^2)	vertical grating luminance (cd/m^2)
5	250	62.9	60.0
10	500	62.9	62.9
20	1000	62.8	62.8

Note. Adjacent pixel non-linearity occurred when stripe width was five pixels. Distance of measurement varied according to stripe width so that the same image was measured in the viewfinder.

To determine if adjacent pixel non-linearity occurred when text was displayed, the luminance of various text sizes was measured. Sizes were pt. size 12, 24 and 48, and each measurement was taken at a distance so that the same image was in the viewfinder. The screen was not gamma corrected, to maximize the text contrast. Text luminance did not differ at different text sizes (see Table

2.2); this demonstrated that adjacent pixel non-linearity was not operating when text was displayed on screen.

Table 2.2

Luminance of different text sizes when presented on screen

Text Size (pt. size)	Distance (mm)	Text luminance (cd/m^2)
12	250	111.83
24	500	111.33
48	1000	111.54

Note. Adjacent pixel non-linearity did not occur when text was pt. size 12, the size used for the stimuli.

2.3.2 Gamma correction

'Gamma' expresses the relationship between the input voltage and the light output (screen luminance) of a monitor. The input voltage allocated to each pixel on the screen is determined by its coded luminance values on a scale of 0-255 (8-bit). On a CRT that has not been Gamma corrected, the Gamma is non-linear. Gamma correction calibrates the screen by applying a look-up table (LUT) that manipulates its output to one that is linearly related to the voltage. This calibration is required to accurately determine, using Matlab, the contrast and luminance of stimuli displayed on the screen.

Gamma correction was carried out immediately before each experiment to ensure that the screen was calibrated. The red, green and blue gun were calibrated separately before experiments that involved colour manipulation (Experiments 3-6) and collectively prior to monochrome experiments (Experiments 1, 2). In each case, the non-linear gamma curve (gamma) was determined by measuring the screen luminance of 11 equally distributed, coded luminance values between 0 – 255. The inverse of the gamma was then applied to calculate the gamma correcting look-up table. Figure 2.3 displays the Gamma and corrected gamma curve that was used on the screen used to display stimuli for the experiments in this thesis.

The gamma was not corrected for Experiment 8 (chapter 6), but instead used to calculate the displayed luminance of the three conditions. This was because gamma correction reduced text contrast on screen to a level which could be associated with overlay use. Since the aim of Experiment 8 was to investigate if reduced text contrast reduced visual stress, it was important to have a high contrast text condition.

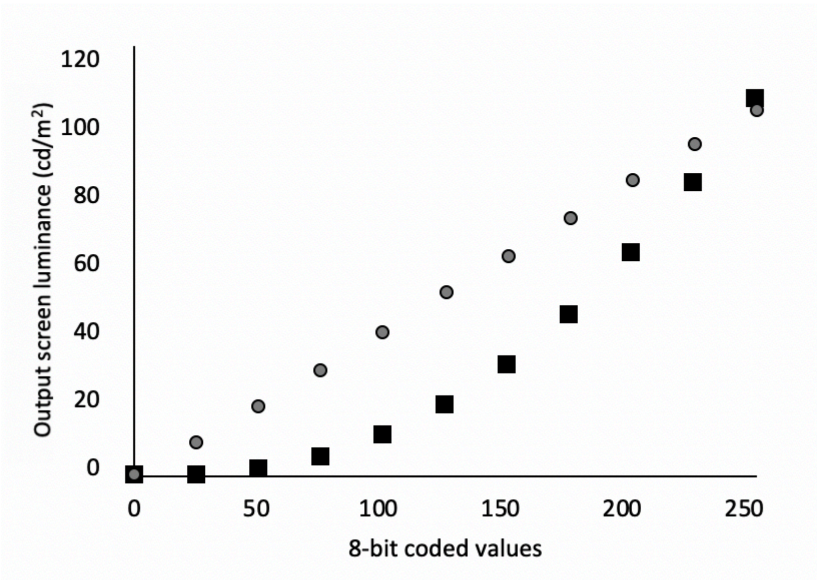


Figure 2.3. Gamma ■ and corrected gamma ● curve of the screen monitor. When gamma corrected, the screen's maximum and minimum luminance was 119.625 and 0.34 cd/m^2; when not corrected, these values were 120.375 and 0.110 cd/m^2 .

2.4 Defining contrast

The measurement of contrast is important for this research because coloured overlays may reduce visual stress by reducing the *text contrast* (the contrast between the text and text background). With the exception of Experiment 7, all experimental stimuli displayed text on a computer monitor with manipulated minimum (text) luminance and maximum (background) luminance. For Experiments 1, 2 and 8, the text contrast and overall luminance was the main manipulation. For Experiments 3-6, the chosen coloured filter, presented as a coloured text background on the screen, was the main manipulation, but contrast also varied in these experiments because the saturations and luminance of the coloured text backgrounds were not held constant; saturation, luminance and contrast is specified in Appendix XII and Appendix XIV for Experiment 3; Appendix XVII and Appendix XVIII for Experiment 4; and Appendix XX and Appendix XXI for Experiment 6.

2.4.1 Measures of contrast

The formula that optimally specifies the contrast of a stimuli depends upon its features (Peli, 1990). Weber contrast (1) is the minimum luminance subtracted from the maximum luminance value, divided by the maximum luminance. Weber contrast is defined by

$$\frac{I-I_b}{I_b} \tag{1}$$

where I is the luminance of a target in a stimulus and I_b is the background luminance. Weber contrast was used to configure that Intuitive overlays reduce text contrast by 2-5% (Wilkins et al., 2001). However, Weber contrast most accurately defines contrast of stimuli containing one small target that does not affect the average luminance of the stimulus, for example, one target letter (Peli, 1990) rather than a page of text.

Michelson contrast (2) is the minimum luminance value subtracted from the maximum luminance value, normalised by the sum of those values. Michelson contrast is defined by

$$\frac{I_{max}-I_{min}}{I_{max}+I_{min}} \qquad (2)$$

where I_{max} and I_{min} are the maximum and minimum luminance of the image. Michelson contrast is used to determine the contrast of periodic patterns with two luminance values over a wider area (Peli, 1990). Michelson contrast has frequently been used to determine the contrast of printed text by substituting the maximum and minimum luminance for the luminance of the text background and printed letter ink (Legge, Parish, Luebker & Wurm, 1990; Knoblauch, Arditi & Szlyk, 1991).

Root mean Squared (RMS) contrast (3) is the standard deviation of the luminance in an image relative to a normalised mean luminance. RMS contrast is defined by

$$\sqrt[2]{\frac{1}{n-1}\sum_{1}^{n}(x_i-\bar{x})^2} \qquad (3)$$

where x_i is the luminance of each pixel and $\bar{x}$ is the mean luminance. RMS contrast is used as a measure of contrast in images where there is a distribution of luminance values (Peli, 1990). RMS contrast was found to be the most reliable indicator of the visibility of natural images (Bex and Makous, 2002) and has been used to determining the contrast of text on screen (Levien, 2003).

2.4.2 Acquiring the luminance properties of printed 'normal' text

It was important that the stimuli used in Experiments 1-6 and 8 represented printed text because experiments have found effects of improved reading performance by using coloured filters with printed text (Wilkins, 2003). The baseline (no filter) conditions in Experiments 1, 2, 3, 6 and 8 were replicated from printed *normal text*, a page of printed text typically found in a book in office-lit conditions. The printed normal text stimulus was single-spaced, 12 point, Times New Roman font, printed by a black and white Cannon C5045i printer using a normal print setting onto a page of 80g/m2 Office Depot paper.

Michelson contrast (2) of this printed text could not be accurately calculated with the materials available due to the difficulty in acquiring the minimum luminance of the text; the luminance of the printed ink. This was because the diameter of the collecting area (the area captured by the photometer when measuring luminance), even with a close up lens of highest available magnification (14.5mm), was too large to be covered by the largest area afforded by the printed letters ink (letter ink with was less than 1mm). Approximate Michelson contrast was measured using a printed shape that was large enough to cover the photometer aperture to understand the behaviour of printed text contrast with increasing background luminance (section 2.4.4). However, to obtain more accurate measurements of the contrast of normal text, the luminance properties of normal text were measured for replicated on screen (section 2.4.5) upon which the contrast was measured using Matlab (section 2.4.6).

Luminance values of the printed text stimulus were acquired in various office conditions where students were reading. The page was positioned on a stand 60˚ to the horizontal and angled so that the light sources caused minimum shadow or specular reflectance of the text. The photometer was secured to a tripod with the view finder positioned 132cm normally from the page. The lighting was a combination of natural and fluorescent lights; the photometer's response speed was set to fast, as stipulated in the photometer's instructions. Photometer readings of *background luminance* (page without text) and *text luminance* (page with text) were recorded in each location. Text luminance measurements were taken with six lines of text in the collecting area. The areas of these measurements were marked on a blank sheet of paper to measure the corresponding background luminance in the same location (Figure 2.4.).

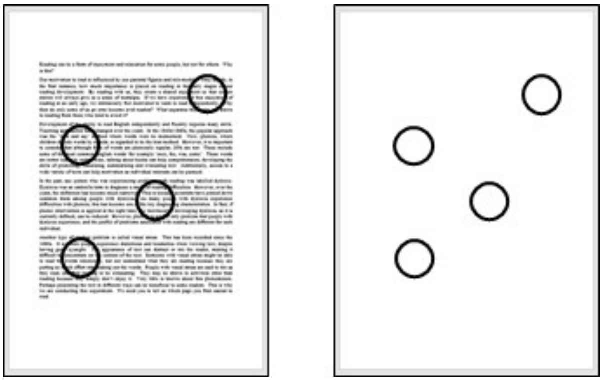


Figure 2.4. Procedure for measuring printed text luminance and background luminance for the purpose of replication on screen: black circles represent the collecting areas of the photometer of a page of text (left) used to measure *text luminance* (letters and spacing) and a page without text (right) used to measure *background luminance*. Collecting areas of text luminance covered six lines of text and were in the same areas of the page without text to measure background luminance.

Although measurements of background luminance were highly variable, ranging from 91.3 – 554 cd/m^2, there was a cluster of recorded values around the highest luminance of the screen background (Figure 2.5). This luminance (119.6 cd/m^2) was chosen to represent the background luminance of the screen. The value to be taken as the corresponding text luminance was unclear because there was no clear linear relationship between these text and background luminance due to varying natural light levels (Figure 2.5).

A linear relationship between background and text luminance of normal text was acquired by repeating the procedure for measuring printed text luminance and background luminance (see Figure 2.4) in controlled, laboratory conditions. A high luminance (15000 lumen), daylight Ecozone bulb was used to reach the background luminance that had been recorded in office locations. The resulting linear relationship between background and text luminance allowed the calculation of the text luminance (97.9cd/m^2) that corresponded to a background luminance of 119.6 cd/m^2. Figure 2.5 displays the measurements taken in the office and controlled conditions to establish the relationship between the background and text luminance of normal text.

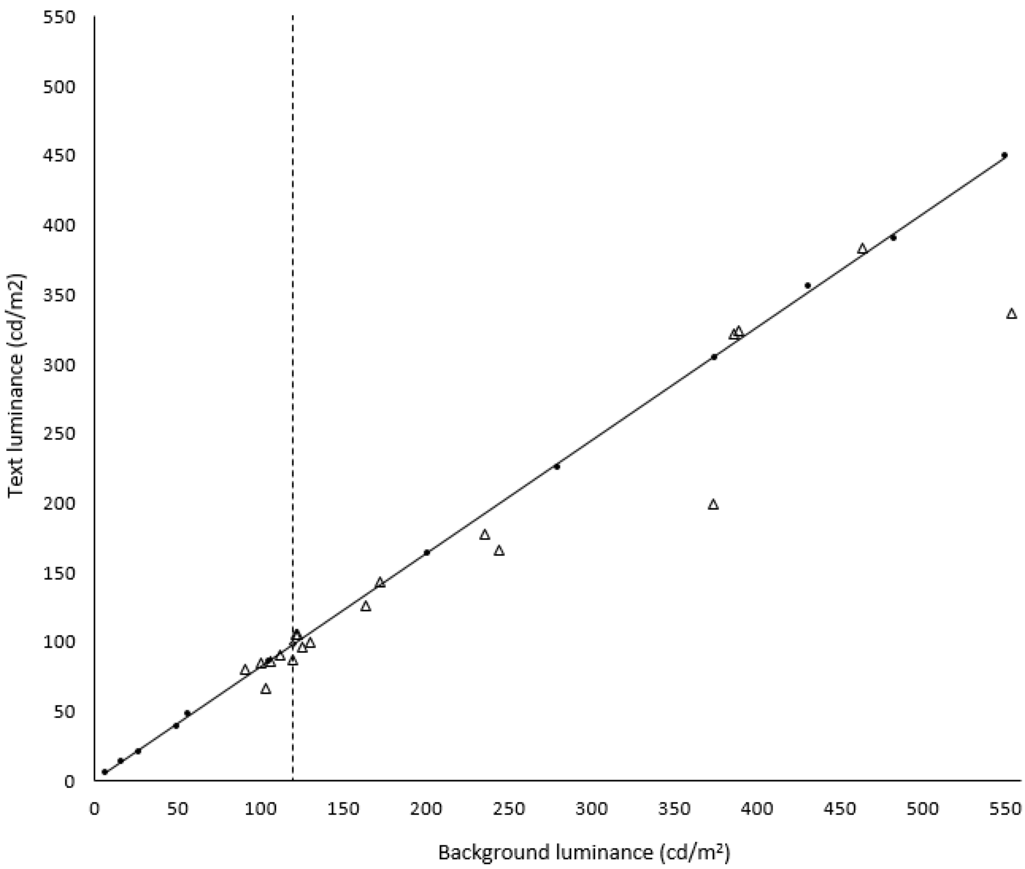


Figure 2.5. The relationship between background and text luminance (letters and spacing) in office △ and controlled ● environments. The vertical dotted line identifies the luminance of the screen (119.6 cd/m^2) which was used as the text background luminance of the baseline conditions of Experiments 2, 3 and 6. The corresponding text luminance was 97.9cd/m^2.

2.4.3 Acquiring the luminance properties of printed normal text with an overlay

The *overlay condition* in Experiments 1 and 2 and 8 was replicated from printed normal text with an overlay. Text and background luminance measurements were acquired using the procedure for measuring printed text luminance and background luminance (see Figure 2.4) from normal text covered by a grey Irlen overlay. Prior to placing the overlay on the text, a high luminance daylight Ecozone bulb was used in controlled, laboratory conditions to acquire the baseline background luminance of 119.6 cd/m^2 so that the same lighting conditions could be applied to the overlay condition. A grey Irlen overlay, matt (not glossy) side up, was superimposed on the text and positioned to minimise specular reflection of the light source. Text luminance and text background measurements were taken using the procedure for measuring printed text luminance and background luminance (see Figure 2.4). Text luminance was calculated to be 42.9cd/m^2 and text background luminance was calculated to be 50.4cd/m2.

2.4.4 Relationship between printed text contrast and background luminance

Since very high levels of background luminance in office conditions (550 cd/m^2) were recorded (Figure 2.5), the relationship between the printed text contrast and background luminance was investigated to ensure that these high levels of luminance did not result in reduced text contrast. The outcome of this enquiry was relevant if filters were found to reduce visual stress by reducing text contrast and would also confirm whether using one luminance profile for the baseline condition (section 2.4.2) was appropriate.

Michelson contrast (2) values of printed text were calculated at different levels of background luminance in controlled, laboratory conditions. A high luminance (15000 lumen), daylight Ecozone bulb was used so that the text background luminance could reach the higher luminance levels recorded in office conditions (Figure 2.5). Michelson contrast (2) values were calculated by substituting the luminance of a black, printed shape, large enough to cover the photometer aperture, for the minimum luminance, and the luminance of the white page for the maximum luminance. The blank page and page with black printed shape were positioned on a stand 60˚ to the horizontal and angled so that the light sources caused minimum shadow or specular reflectance of the text. To obtain luminance measurements, the photometer was secured to a tripod with the view finder positioned normally from the page. The recorded Michelson contrast values were only approximate (Figure 2.6) because the luminance of the black printed shape could not be verified to be the same luminance as the printed letter ink.

At lower levels, as the text background increased, Michelson contrast of the text increased. When text background luminance reached the level that would be used for the screen text background of the baseline conditions of Experiments 2, 3 and 6 (119.6 cd/m^2), the Michelson contrast of the text appeared to have almost reached its maximum, only marginally fluctuating until the highest recorded background luminance. This outcome supported the use of one luminance profile as the base line (no filter) condition.

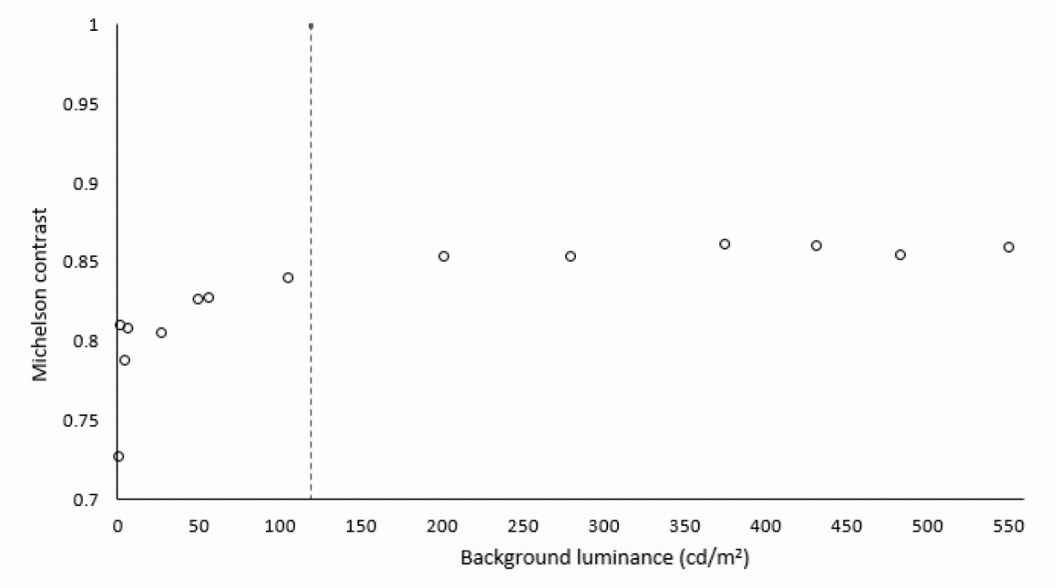


Figure 2.6. The approximate relationship between the text background luminance (page without text) and the Michelson contrast of normal text with no overlay; the minimum luminance was the luminance of a black, printed shape.

2.4.5 Presenting text on screen

Replicating a dark enough text luminance to represent normal (baseline) text (97.9cd/m^2), and the associated text contrast, was hindered by the presence of grey-scales in the screen text; rendering text on a screen is different to printing text on a page. When text is printed on paper, a resolution of 600dpi (dots per inch) is common. This resolution is beyond the observable limit of a human reading printed text at an average distance of 40cm from the page, and as such a font like Times New Roman at 12pt can be recreated in a readable format. However, when displaying text on the screen the pixel density is lower at 63.37ppi (pixels per inch).

To accommodate for this drop in resolution and the unpredictable alignment between the dark areas on a letter glyph and the pixels on a screen, a process called rasterization is applied to the letters so that they appear as smooth and sharp as possible; for example, where a part of a letter to be displayed only takes up half a pixel, that pixel is displayed at 50% luminance and the human visual system is relied upon to smooth it out. Therefore, due to rasterization the actual output of text on a screen will include many grey pixels to form the text.

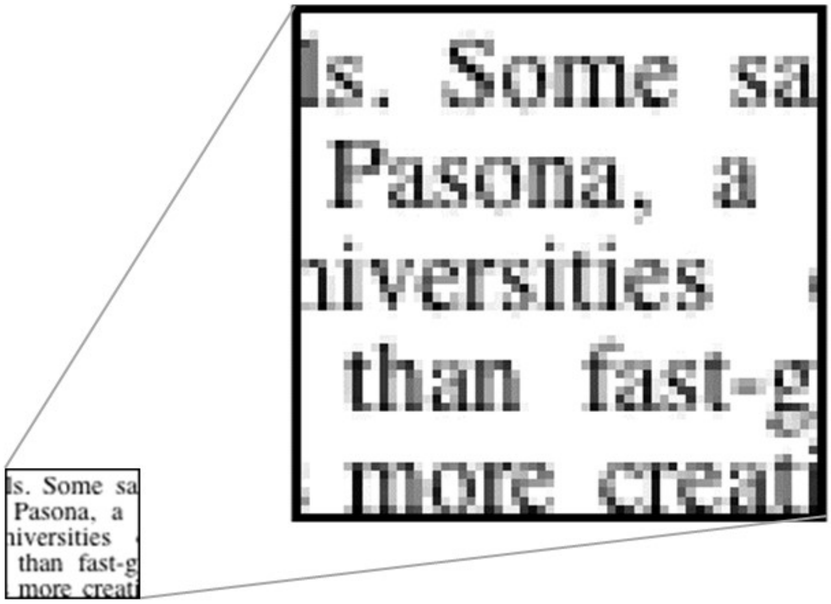


Figure 2.7. A screenshot of text that was presented on screen using 'DrawFormattedText' in Experiment 6. The image on the left is the actual size of the stimuli and the image on the right is magnified (scale: 1:16) to illustrate the greyscales that comprise the letters. This image consists of 60.7% white (RGB = 255) pixels, 6.5% black (RGB = 0) pixels and 32.8% grey scale pixels.

2.4.5.1 Text background and Text luminance of normal text conditions

The luminance properties of normal text were applied to the normal (baseline) text conditions on screen of Experiments 1, 2, 3 6 and 8[5]; in Experiments 3 and 6, these baseline conditions were condition with no filter. The text background luminance of the normal text condition in Experiment 1, 2, 3, 6 and 8 was set to the maximum screen luminance; this was 119.625 cd/m^2 in all experiments except for Experiment 8, which had not been gamma corrected and so had a maximum screen luminance of 120.375 cd/m^2.

In Experiment 1, the normal text condition had a text luminance of 105.772 cd/ m^2; this condition used a relatively high minimum RGB text pixel value of 50 to allow for a 'reduced luminance' condition with RGB text pixel value of 0. The elimination of this 'reduced luminance' condition in Experiment 2 allowed the normal text condition to use the minimum RGB text pixel value of 0, resulting in a darker text luminance of 102.393 cd/ m^2; this value of text luminance was applied to the baseline conditions of Experiment 3 and 6.

Experiments 1-6 used a gamma corrected screen so that the luminance required for replication of filters (Experiment 3 and 6) and screen colour changes (Experiment 3 – 5) could be accurately displayed. However, this came at the expense of achieving a dark enough text luminance to represent normal text. Gamma correction increased text luminance by increasing the luminance of the minimum pixel from 0.110 to 0.34 cd/m^2, and by increasing the variability of grey scales in the lower luminance settings. Gamma correction was not needed in

---

[5] Experiments 4 and 5 did not contain baseline (no filter) conditions. The stimuli of Experiment 7 was on paper, not screen

Experiment 8 and so an uncorrected screen was used to achieve a text luminance of 96.552 cd/m^2. The stimuli of the normal text condition was the closest approximation to the acquired text luminance measurements of normal text (97.9cd/m^2) out of all the experiments that presented text on screen and so produced the text contrast that most closely approximated text contrast of normal text.

For values of luminance and contrast in these experiments, see Appendix XI (Experiment 1 and 2), Appendix XIV (Experiment 3), Appendix XVIII (Experiment 4) Appendix XXI (Experiment 6) and Appendix XXIV (Experiment 8).

2.4.6 Measuring contrast of text once presented on screen

The text stimuli in this thesis were created using a function in Psychtoolbox called 'DrawFormattedText'. The use of this function allowed the manipulation of text spacing so that a Times New Roman, a proportional font, could be presented with a left and right justified edge in Experiment 2-6 and Experiment 8. However, with regard to text luminance, only the *minimum text pixel luminance* (the lowest luminance of pixel in the letters) can be specified; this does not represent the *average text pixel luminance* (the average luminance of the pixels that comprise the letters). Figure 2.7 displays text that was presented on screen with the minimum text pixel RGB set to 0, and the background RGB set to 255, yet only 6.5% of this image comprises pixels of RGB values of 0.

DrawFormattedText generates a luminance profile of grey-scale pixels according to the minimum text pixel luminance and the text background luminance that has been inputted. As the luminance of either of these values increase, so too does the range and average text pixel luminance. For example, when minimum pixel luminance was held constant at 0.34cd/m^2, the luminance of the average text pixel had a positive linear relationship with the luminance of the text background, ranging between 0.34 - 70.58 cd/m^2 for a text background of 0.34- 119.625 cd/m^2. Furthermore, when background luminance is constant and the text luminance is set to increase, the range and average luminance of text pixels increases.

2.4.6.1 RMS or Michelson contrast?

RMS contrast incorporates a calculation of the difference between the luminance of each pixel, including the grey scales in text, with the average text luminance. However, RMS contrast of text on screen changes when the range in text grey scales change, even if the average text pixel luminance remains constant. Therefore. whether or not to use RMS to specify contrast in this thesis depends upon whether these individual grey scales are perceivable when reading.

Contrast sensitivity, (the ability to perceive contrast, and therefore greyscales in text) increases to 5 cycles/degree, after which it falls (Levien, 2003). Levien (2003) recommended a weighted RMS contrast metric with a model of the human visual system's contrast sensitivity function to measure text contrast on screen for the purpose of evaluating perception of display quality. However, for the purpose of reading, the perception of individual text pixels are unlikely to be perceived: eye movement data has revealed that Information is only processed during the fixation period, which, for silent reading, is only 225-250 ms (Rayner, 1998). Therefore, it was assumed that the greyscales in text displayed in these experiments were imperceptible when reading. Accordingly, for conditions that had constant spatial properties of the text (all conditions apart from condition 3 in Experiment 8), Michelson contrast, requiring two luminance values, would appear to be the most appropriate contrast measure for text.

2.4.6.2 Which minimum luminance value?

Whilst Michelson contrast is usually determined by the maximum and minimum luminance, minimum text pixel luminance only accounts for a few percent of the pixels used to form text. Levien (2003) demonstrated that applying Michelson contrast using the minimum text pixel luminance underestimated the antialiased perceived contrast of antialiased (greyscale) text (Figure 2.8).

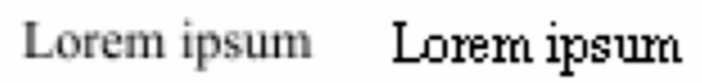


Figure 2.8. A comparison of the perception of antialiased (grey scale) text (left) and text that is not rasterised. Michelson contrast is identical for the two samples, even though the antialiased version clearly has lower perceived contrast (Levien, 2003).

RMS contrast would over specify the contrast of the grey scales, which are assumed not to be perceived when reading. Therefore, this thesis applied the average text pixel luminance, rather than minimum text pixel luminance to the Michelson contrast formula to calculate the perceived text contrast of text on screen. In this thesis, this measure of contrast is referred to as *Michelson text contrast*, and is determined by:

$$\frac{I_{max}-I_{text}}{I_{max}+I_{text}} \tag{4}$$

where I_{max} and I_{text} are the maximum luminance and the average text pixel luminance of the image. The decision to use Michelson text contrast to specify text contrast in this thesis was taken further to completion of the experiments and so had some implications for this thesis.

2.4.6.3 Implications for this thesis

Appendix XI (Experiment 1 and 2) and Appendix XXIV (Experiment 8) display luminance and contrast values of experiments that investigated the effect of luminance and contrast. RMS contrast was useful to determine the properties of the increased linewidth condition in Experiment 8; this condition had the same luminance values as normal text but a higher percentage of pixels allocated to the text background. However, where RMS contrast was used to replicate contrasts on a different background luminance, their values were not equal to other contrast definitions, in particular, Michelson text contrast. The unintended luminance properties of these conditions are described in Experiment 1 (section 3.1.2.2) 2 (section 3.1.5.2),and 8 (section 6.2.1.2).

Appendix XIV (Experiment 3), XVIII (Experiment 4) and XXI (Experiment 6) display luminance and contrast values of the experiments that used chosen filters. The application of the average text pixel luminance, rather than minimum text pixel luminance to determine Michelson contrast would appear to demonstrate that the application of background colour on the screen increased the contrast of many of the stimuli in Experiment 3, and, due to their low luminance, all of the stimuli in Experiments 4 and 6. The implication of this is discussed in relation to the findings of Experiment 6 (section 5.5.4).

2.5 Colour spaces

2.5.1 CIE (1976) LUV colour space

The CIE (1976) *u′ v′* diagram was used to plot chromaticities of filters (Experiment 3, 4, 6) and paper conditions (Experiment 7) in this thesis. However, it is important to note that Euclidean distances on this chromaticity diagram cannot accurately determine saturation and colour differences because the colour space is not perceptually uniform and only approximately represents what the eye perceives. The CIE (1976) colour space was linear transformed from the CIE (1931) colour space to be more perceptually uniform. However, the accuracy of colour differences calculated by the CIE (2000) colour-difference formula alongside the CIE 1976 La’ b’ colour space (Luo, Cui and Rigg, 2001) has out-performed those calculated by the CIE (1976) colour space and is currently recommended by the CIE (Witt, 2007).

Euclidean distances on the CIE Luv (1976) diagram to obtain chromaticities of filters with similar saturation and colour differences was claimed to be sufficiently accurate for the purpose of determining Intuitive filters (Wilkins 2001; Wilkins et al., 1992) and placebo control filters (Wilkins, Patel, Adjamian & Evans, 2002). This method has been used to demonstrate effects of chosen filters for reading (Evans & Allen, 2016). Experiments 4 and 6 adhered to the Intuitive method of acquiring a chosen filter (Wilkins et al., 1992) and Experiment 4 adhered to Wilkins et al.'s (2002) method of acquiring a placebo control filter. As such, chromaticities are presented on the CIE (1976) Luv colour space for consistency with these methods. Colour difference (h_{uv}) and saturation (S_{uv}) were specified using the formulae that relate to the CIE (1976) colour space. Accordingly, colour difference (5), was specified as:

$$arctan\left(\frac{u}{v}\right) \tag{5}$$

and saturation (6), was specified as:

$$\frac{\sqrt{(u^2 + v^2)}}{L} \tag{6}$$

where *L* is luminance.

2.5.2 MacLeod-Boynton (1979) cone-excitation colour space

Chromaticities of filter were also presented on the MacLeod-Boynton (1979) cone-excitation diagram (MacLeod & Boynton, 1979) to provide a comparison with the CIE (1976) LUV colour diagram and the potential for investigation into whether colour choice of filter could be attributed to any low-level colour processing. The abscissa (horizontal *r* axis) of the MacLeod-Boynton diagram represents the ratio of L to M cone activation with constant S cone activation and luminance, with luminance assumed to be the sum of the L and M cone activation; The ordinate (vertical *b* axis) represents S cone activation with constant L and M cone activation, or luminance (MacLeod & Boynton, 1979). Shepherd (1997, 1999) applied a logarithm of *b* to the ordinate of the MacLeod-Boynton diagram to better approximate the diagram to the way the visual system processes colour differences; they reported that colours around neutral are approximately equally perceptually spaced in this diagram following this transformation. Since this log transform was used to present the cone-excitation of colour choices of filters previously (Veszeli and Shepherd, 2019) MacLeod-Boynton, *r, log10(b),* diagrams were used to present cone-excitation of filter choices in this thesis.

2.6 Colour names

Perceptual categorisation of colour is subjective and affected by previous experience of similar colours (Nosofsky, 1987; Nosofsky & Palmeri, 1997). Due to the absence of standardised colour names, colour names were applied to the 12 chromaticities of the colorimeter (Appendix XV). Colour names were based on those used by Conway et al. (2016) to describe chromaticities of Intuitive overlays, those used by Wilkins (1994) to describe approximate chromaticities of the colorimeter and those used by Al-Rasheed (2015) to describe hue preference.

2.7 Viewing distance and viewing angle

Participants were instructed to sit at a distance which allowed them to read comfortably and to then keep their viewing distance constant. For each participant, the chair was adjusted so that the top edge of the computer screen was just below eye level. Viewing distances were recorded and ranged from 400-550mm, resulting in viewing angles of 16 - 22˚; this range does not affect reading speed and accuracy (Gould and Grischkowsky, 1986).

2.8 Ethics

All experiments were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Prior to each task in each experiment, participants were given written and verbal descriptions of the tasks, were told that they could stop the experiment at any time and were given a consent form to sign. A debrief sheet was given at the end of each experiment, explaining the purpose of the experiment, and participants were asked to sign the final part of the consent form.

2.9 Participants

Due to the absence of a validated measure of visual stress, for Experiments 3,4 5 and 8, participants who had used coloured filters for reading for more than six months (filter users) were recruited: their continued use of filters was assumed to verify that they suffered from visual stress. Filter users in Experiments 3-5 used filters acquired from Intuitive and Irlen methods, and filter users in Experiment 8 included participants who used coloured screen filters.

These participants were difficult to recruit because opticians and disability services could not share their lists of filter users due to issues with confidentiality, and participants with reading problems do not always notice adverts and signs that involve text. Various methods were employed to advertise a website containing the research information (Appendix II). Flyers were handed out at public events and displayed across notice boards in disability services and opticians. Animated adverts were displayed across the university, on social media and on the big screen at @ Bristol: the science museum in Bristol. Letters were written to universities and 6th form colleges. Lastly, a local radio broadcast was carried out.

2.10 Analyses

Statistical significance was accepted at the level adapted using the Holm-Bonferroni multiple comparison correction method (Holm, 1979) for all *t*-tests and pairwise comparisons. For all tables, * = significant according to Holm-Bonferroni's test for correction. All bar graphs contain error bars that are 95% confidence intervals.

Chapter 3: Developing methods

The objective of the work described in this chapter was to develop methodologies for subsequent studies that would determine if chosen coloured filters benefited reading beyond placebo for people with visual stress (Chapters 2-4). The overarching objective of these experiments was to develop a sensitive measure of reading performance to this end. Experiment 3 began the investigation into the effect of coloured filters by involving the development of a placebo-controlled design that exploiting the phenomenon of 'change-blindness', a technique used in the experiments described in Chapter 4.

3.1 Experiment 1 and 2: developing a measure of reading performance

3.1.1 Introduction

3.1.1.1 A measure of reading performance

Reading is complex and consists of many processes, but little is known about how these processes interrelate (Rayner & Reichle, 2010). Reading speed is influenced by task instructions (Carver, 1990) and the purpose of the task may affect the balance of processes required to complete it. For example, performance on a reading task that requires searching for information might be primarily determined by scanning processes, whereas a task that requires remembering information might require more use of working memory. The purpose of natural reading is to efficiently understand information derived from text. However, the level of comprehension in natural reading is highly variable and difficult to measure.

Previous studies have measured children's reading development with separate markers of reading comprehension, speed and accuracy from the Neale Analysis of Reading Ability, the Formal Reading Inventory, the Gray Oral Reading Test (GORT) and the Woodcock reading test (Robinson & Conway, 1994; Robinson & Conway, 2000; Noble et al., 2004; Blaskey et al., 1990). However, these measures are not appropriate for this thesis, which concerns the reading speed of adults with fully developed reading.

An adult reading performance measure should simultaneously measure reading speed and understanding to ensure that reading speed improvements are not the result of an offset between reading speed and understanding. Dillon (1992) and Wu and Yuan (2003) measured reading performance by using the time taken to read a page of text and the number of correctly answered, multiple choice, comprehension questions. However, this reading performance measure was not appropriate to demonstrate if filters reduce visual stress because the reading duration may not have been long enough for visual stress to develop. Additionally, although this measure gave a general idea of reading performance, it was not sensitive enough to demonstrate subtler improvements in reading speed or understanding that may have been attributed to reduced visual stress. Therefore, to definitively answer whether coloured filters reduce visual stress, a reading performance measure was required with a substantial reading duration and frequent (sensitive) measures of reading speed and understanding.

Single word reading tasks may provide a suitably sensitive reading performance measure than text (natural) reading tasks because key-presses can measure reaction times more regularly: after each word. Some correlations of eye-fixation data from word recognition lexical decision tasks and silent reading tasks have been found to correlate (Schilling, Rayner & Chumbley, 1998). Yet previewing a successive word in natural reading is known to reduce the fixation time of a word (Rayner & Juhasz, 2004). Legge, Hooven, Klitz, Mansfield and Tjan (2002) suggested that text reading was faster than word reading due to the higher predictability of words when positioned within a sentence. But the relationship between the underlying processes required to complete single word and text reading tasks, when text understanding is monitored, is unknown.

Experiments 1 and 2 investigated if lexical decision (single word reading) tasks were related to text reading tasks and therefore could be used to measure reading performance when text understanding is also monitored. Experiment 1 investigated if there was a correlation between the reading speeds in a semantic word level task when words were singularly or collectively presented and silently read. Experiment 2 investigated if there was a correlation between the reading speeds of a silent word reading and text level task. If reading speed correlated between the two tasks in each experiment, then the more sensitive measure could be used to determine if there was a reading benefit of coloured filters. Tasks were longer than five minutes to allow enough time for visual stress symptoms to develop.

3.1.1.2 Does reduced contrast affect reading speed of typical readers?

A second objective of Experiment 1 and 2 was to determine if the reduction in contrast or luminance of a filter affected the reading performance of typical readers. When text stimuli was printed on paper, typical readers were reported to have better reading performance in high contrast text conditions than conditions of reduced contrast. Legge, Rubin and Luebker (1987) reported that, whilst contrast does predict text readability, reading rate was very tolerant to contrast reduction: when the contrast had been reduced by 90%, reading speed had only reduced by half. Although exact contrast values were not reported, high contrast colour combinations, including black on white text, yielded faster reading rates (Tinker and Paterson, 1931), and higher scores of legibility (Poffenberger, 1925; Preston, Schwankl & Tinker, 1932) than lower contrast combinations.

All tasks were conducted with different measures of text and background luminance (luminance conditions). The design was a within-subjects experiment with two factors: task type and luminance setting. Tasks were single word reading task and successive word reading task. The dependent variable was reading speed. Figure 3.1 displays sequence and timings of Experiment 1 and 2.

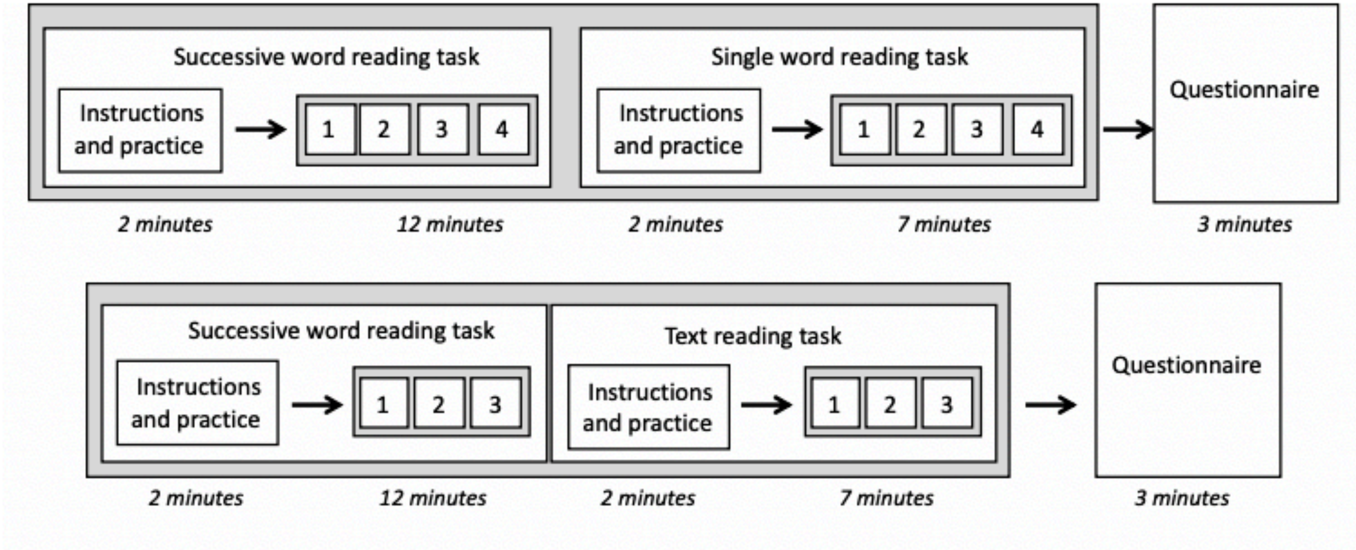


Figure 3.1. Sequence and timings of Experiment 1 and 2. **a**. In Experiment 1, tasks were successive word reading and single word reading. **b.** In Experiment 2, tasks were successive word reading and text reading. Numbers represent luminance conditions 1-4: Normal (baseline) condition, overlay condition, reduced contrast condition and reduced luminance condition[6]. Grey borders illustrate counterbalancing of luminance conditions and tasks. Timings are in italic font.

[6] The fourth condition of Experiment 1 could not be achieved in Experiment 2 due to the higher contrast of times new roman font, see section 3.1.5.2.

3.1.2 Experiment 1: Method

3.1.2.1 Participants

24 undergraduate Psychology students (22-39 years old, 17 female) with normal or corrected to normal vision were recruited by advert in exchange for payment. Participants had had an eye-test within the last two years, and all had English as their first language except one participant who was bilingual. No participants in the sample reported that they had visual stress and two participants reported that they had dyslexia.

3.1.2.2 Stimuli

The *single word reading task* was a semantic categorisation task that measured the time taken to categorise whether a single noun was the name of an animal (animal word) or not (Forster & Shein, 1996). The *successive word reading task* measured the time taken to locate an animal word within a block of text containing non-animal words. The animal and non-animal words of both tasks were pseudorandomly selected from animal and noun word banks (Appendix III) that had been assessed to contain words that were easily recognised to belong within each category. Each word was only used once throughout the course of the experiment.

Figure 3.2 displays the stimuli used for the single and successive word reading task. One trial of the single word reading task required a singular presented word to be categorised according to whether it was an animal word or not. An 'M' keypress recorded the decision that the word was an animal word and a 'Z' keypress recorded the decision that the word was not an animal. Both keypresses recorded reaction times. Correctly categorised words activated a blank screen for 0.5 seconds and incorrectly categorised words returned a short beep and 1 second pause. Within each luminance condition, there were three trials displaying an animal word and three trials displaying a non-animal word. The presentation order of animal and non-animal words was counterbalanced across participants.

One trial of the successive word reading task required each word of the presented text to be read in order, the 'Z' key to be pressed on reaching the end of each line, the 'M' key to be pressed on reaching the animal word and the located animal word to be typed in. The 'Z' keypresses returned a clicking sound to remind participants to read each word in order and provide a measure of line reading speed for analysis. If six 'Z' keypresses were pressed in one trial, a screen was displayed showing the location of the unfound animal word and an extra trial was added; this was to encourage participants to complete the task correctly. Accidental keypresses returned a short beep. The text contained six lines that had been pseudorandomly chosen from a pre-arranged line bank, each containing ten words and 64 characters. Word lengths in each line were formatted to match that of regular text. The animal word was not located in the first or last word position of a line and so was positioned three times in each of the 8 available word positions and four times on each of the six lines in the complete session of 24 trials. The orders of these locations were different in each session.

Both tasks used a non-proportional 'Courier' font to aid the formatting of the successive word reading task in which animal words could easily replace a randomly selected noun of equal word length, resulting in a straight left and right margin. Since courier font is a relatively low contrast font, the luminance properties of printed normal text for the baseline condition were better approximated by setting the font to bold with a tighter line-spacing of 0.9; this resulted in an equivalent text luminance to non-bolded, single spaced Times-New-Roman font (used for all other text stimuli in this thesis). The luminance properties of each condition categorised the screen settings throughout so that blank screens also exhibited the set background luminance for that condition. Each new condition began with a screen with the new text and background luminance settings, a timer and the text 'Now the luminance settings will appear like this. Press M or Z to continue'.

Both tasks had four luminance conditions:

1. *Normal (baseline) condition:* luminance properties to match normal text without an overlay
2. *Overlay condition*: luminance properties to match normal text with an overlay
3. *Reduced contrast condition*: luminance to match normal condition and RMS contrast[7] to match the overlay condition
4. *Reduced luminance condition*: luminance to match overlay condition and RMS contrast to match the normal condition

The text and background luminance of the normal and overlay condition were replicated from printed normal text and printed normal text with an overlay in office conditions (See sections 2.4.2 and 2.4.3). Conditions 3 and 4 had unintended luminance properties due to the use of RMS to match their contrast to Conditions 2 and 1. ‘Michelson text contrast’, determined to be the most accurate way to define screen contrast post testing (section 2.4.6.2), calculated the text contrast of Condition 3 to be 75.7% lower than condition 2 and the text contrast of Condition 4 to be 23.5% higher than Condition 1. Luminance and contrast measurements of the four conditions are displayed in Appendix XI.

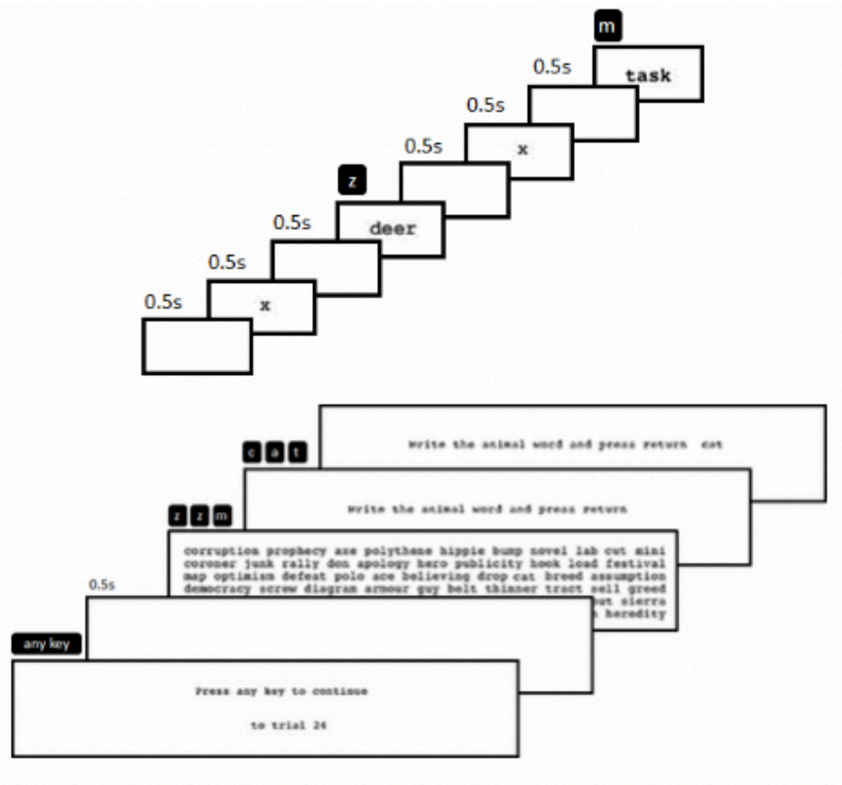


Figure 3.2. Stimuli of word reading tasks with times and correct keypresses (Experiment 1). Top: Single word reading task. Bottom: Successive word reading task.

3.1.2.3 Procedure

The sequence of the experiment is displayed in Figure 3.1. Participants completed the single and successive word reading task in counter-balanced order with a short break between them. With the exception of the second participant[8], the order of luminance conditions was counterbalanced so that each participant completed the conditions in a different order.

Prior to completing each task, participants were verbally instructed and prompted to read the written instructions in the programme. Instructions for both tasks specified that each word should be read and categorised according to whether it was an animal word. The single word reading task required 'Z' and 'M' keypresses to determine if the presented word was an animal or non-animal word. The successive word reading task specified that each block of text should be read from top left to bottom right, the 'Z' key should be pressed on reaching the end of each line, the 'M' key should be pressed on reaching the animal word and then the animal word should be typed. The 'Z' and 'M' key presses were instructed to be pressed with the left and right index fingers, respectively.

Both tasks included four practice trials in the luminance condition that was presented first. The experimenter sat with the participant to ensure that the task was completed properly. There were six trials of each task in each luminance condition. On completion, participants were asked two questions to elicit whether they had dyslexia or visual stress.

1. Have you ever been given extra time in exams? If so, what for?

[8] This participant mistakenly received the order of the third and fourth luminance conditions the wrong way around.

2. Do you have any vision problems which you feel may have affected this task?

3.1.3 Experiment 1: Analysis

3.1.3.1 Data preparation

Reading speeds of the successive word reading task were calculated from 'Z' keypress reaction times, the time taken to read each line of ten words, of which there were 1487 data points. Twenty-eight of these data points corresponded to 10 (out of 144) error trials where an incorrect word had been typed. A further 84 data points corresponded to 14 error trials where the animal word had not been found. These 112 error data points were not removed from the analysis because error trials were not thought to affect the rate of reading the previous lines of text. However, their removal did not affect the analysis. For the single word reading task, there were 576 reaction times in total, but 14 were incorrectly categorized words and so were removed from the analysis. For both tasks, outlying data points were excluded from the analysis. Criteria for exclusion were values greater than 2.5 standard deviations from each participant's mean, or absolute reading speeds that were either less than 60 words/minute or more than 600 words/minute. The latter criterion was to eliminate accidental or forgotten 'Z' keypresses. See Table 3.1 for details on outliers.

Data were categorised according to the two task types and four luminance conditions (see Figure 3.3).

Table 3.1

Categories of outliers in word search task and single word decision task (Experiment 1)

	Word search task	Single word decision task
Total data points	1487	562
<60 or >600 words/min	65	0
Z-score > ±2.5	43	12
Total outliers	89	12
Remaining data points	1398	550

Note. For the word search task, data points were reading speeds retrieved from 'Z' key-press reaction times. For the single word decision task, data points were reading speeds retrieved from correctly categorized words.

3.1.3.2 Did luminance affect reading speed?

To determine if a change in luminance elicited statistically significant differences in reaction speed for these tasks and participants, a repeated measure ANOVA was conducted on each task. There were no outliers within any group. All groups, except luminance condition 2 in the successive word reading task (Shapiro-Wilk, p = 0.044), were normally distributed. ANOVAs of randomly generated numbers have been shown to be robust against violations of normally distributed data in a one-way ANOVA (Schminder, Ziegler, Danay, Beyer & Bühner, 2010). Therefore, it was assumed that this violation would not affect the outcome of the analysis.

Mauchly's test was used to assess the assumption of sphericity and this was met on both the successive word reading task, $\chi^2(5) = 4.223$, $p = .518$ and single word reading task, $\chi^2(5) = 1.358$, $p = .929$. Change in luminance did not elicit statistically significant changes in reaction speed for the successive word reading task $F(3,69) = 2.109$, $p = .107$, $\eta_p^2 = 0.84$, or single word reading task $F(3,69) = 1.781$, $p = .159$, $\eta_p^2 = 0.72$ (Figure 3.3).

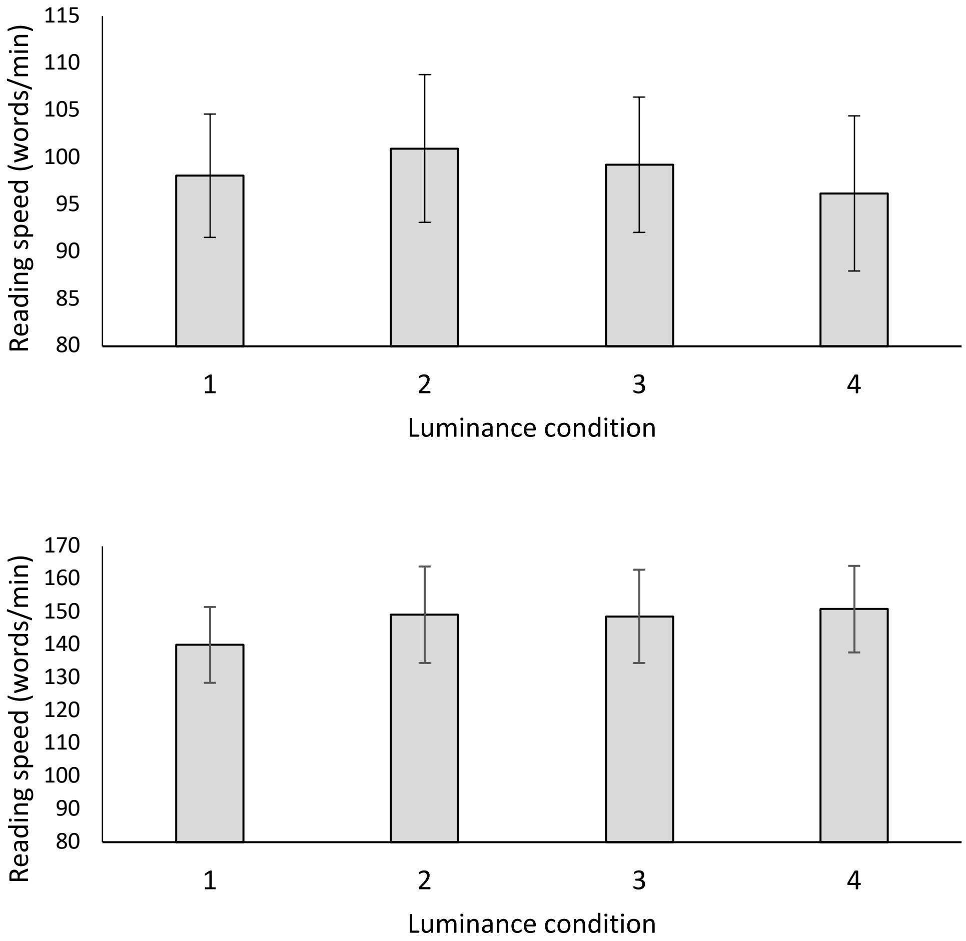


Figure 3.3. Reading speeds of word reading tasks across luminance conditions 1-4: Normal (baseline) condition, overlay condition, reduced contrast condition and reduced luminance condition (Experiment 1). Top: Single word reading task. Bottom: Successive word reading task.

3.1.3.3 Were the tasks related?

To determine if the data obtained from both tasks were related, the data for each task type were collapsed across luminance conditions and a Pearson correlation was administered. Reaction times on the two tasks were not significantly correlated, n = 24, r = 0.27, p = .899.

3.1.3.4 Order effects

Order effects for each task were investigated to determine if the data of the second participant, who had completed the third and fourth luminance condition in the wrong order, should be removed. This analysis would also serve to determine if the sensitivity of the tasks could be improved by reducing any practice effects. Each condition was ranked according to the order that it had been completed. This resulted in four groups specifying the data which had been completed in the first, second, third and fourth conditions. A repeated measures ANOVA was conducted to assess if the order of conditions had affected reading performance in both tasks (Figure 3.4).

For the single word reading task, the assumption of sphericity was met, as assessed by Mauchly's test of sphericity, $\chi^2(5) = 8.528$, p = .130. Differences in the order in which the conditions had been completed did not have a statistically significant effect on reading speed, $F(3,69) = .959$, p = .417, η_p^2 = .040. For the successive word reading task, the assumption of sphericity was met, as assessed by Mauchly's test of sphericity, $\chi^2(5) = 5.217$, p = .390. Change in order elicited a statistically significant effect in reading speed, $F(3,69) = 3.051$, p = .034, η_p^2 = .117. Pairwise comparisons revealed that the first condition elicited significantly slower reading speeds than the other three conditions, whilst the other three conditions were not significantly different from one another, see Table 1.2. Although the second participant had completed the third and fourth luminance conditions in the wrong order, there were no order effects between these conditions so the data from this participant were retained.

Table 3.2.

Pairwise comparisons between the order in which the conditions had been completed in the successive word reading task (Experiment 1)

Order of conditions		*t*(23)	Adjusted target alpha	*p*	d
1	4	-2.705	0.05	0.013*	0.552
1	2	-2.574	0.025	0.017*	0.525
1	3	-2.140	0.0167	0.043	0.437
3	4	-0.416	0.0125	0.681	0.085
2	4	-0.336	0.01	0.740	0.069
2	3	0.104	0.0083	0.918	0.021

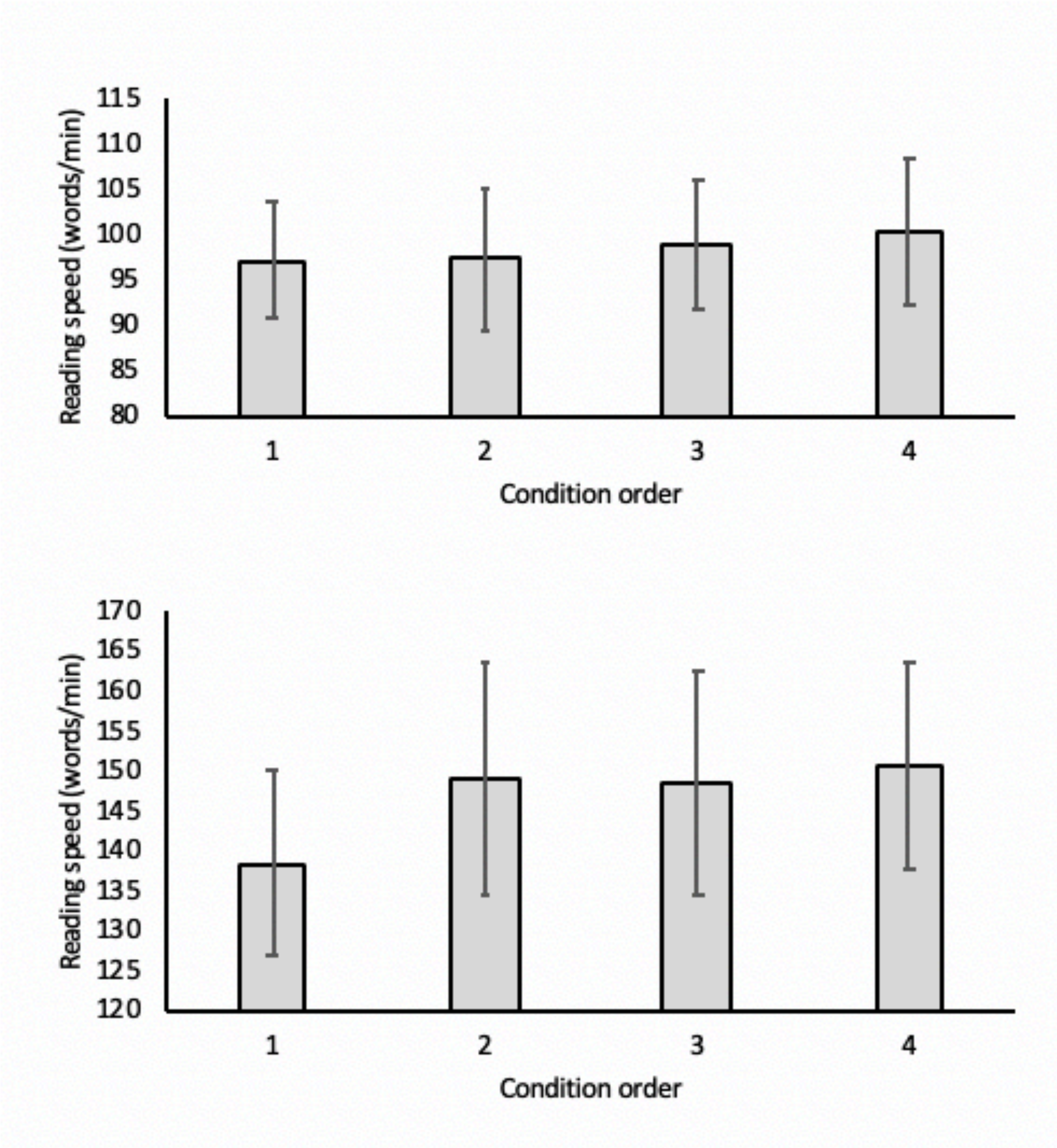


Figure 3.4. Reading speeds of word reading tasks when categorised according to the order that the luminance conditions had been completed (Experiment 1). Top: Single word reading task. Bottom: Successive word reading task.

3.1.4 Experiment 1: Discussion

Reduced contrast or luminance did not affect reading speed of typical readers using the tasks and range of values described here. However, the focus of the experiment was whether the tasks were related. No significant correlation was found between average reading speeds, across all luminance conditions, obtained from the single and successive word reading task. Additionally, the single word reading task was found to be significantly slower than the successive word reading task.

The absence of a correlation between the two tasks implies that there is no relationship between the reading processes required for single and successive word reading for typical adult readers. It is plausible to suggest that categorical information from successive rather than singular words could be processed in a parallel fashion, thereby using attentional resources more efficiently. Therefore, although the single word reading task measured reaction times more regularly than the successive word reading task, reading requires sequential processing of written words. It would appear that single word reading tasks cannot provide a suitably natural reading performance measure.

An order (practice) effect was found for the successive word reading task. Significantly slower reading speeds were recorded in the condition that was completed first in comparison with those completed second, third or fourth. Practice effects had been accounted for by counterbalancing the order in which the conditions were completed across participants. However, this order effect adds to the overall variance and makes the task less sensitive. Experiment 2 aims to reduce this practice effect by increasing the number of trials (see Experiment 2).

3.1.5 Experiment 2: Method

3.1.5.1 Participants

24 undergraduate Psychology students (19-55 years old, 12 female) with normal or corrected to normal vision were recruited by advert in exchange for payment. Participants had had an eye-test within the last two years, and all had English as their first language. No participants in the sample reported that they had visual stress or dyslexia.

3.1.5.2 Stimuli

The word search task was based on the word search task described in section 3.1.2.3, with the exception that the animal word was hidden within 24, rather than six lines of non-animal words. No target appeared where it could easily be spotted, for example, at the beginning or end of paragraphs or lines, and positions within a session were pseudorandomised in locations that were evenly distributed across the page.

The text reading task (Dillon, 1992; Wu & Yuan, 2003) was presented in the same format as the word search task. This required participants to read three non-fiction newspaper stories and press a key, which recorded their reaction time, to indicate they had finished. Their comprehension was measured by three multiple choice test questions. Article reading times were used in the analysis if two out of three multiple choice answers were correct.

Unlike Experiment 1, the text in Experiment 2 was formatted to appear like text in a book. Each page had three indentations to give the appearance of three paragraphs. Text was a proportional font, Times-New-Roman, single spaced and configured to be left and right justified with 10-13 words on each line. The word-length profile of each line was calculated to match that of normal text. This was based on the frequency distribution of high frequency words that had been extracted from the CELEX database (Baayen, Piepenbrock & van Rijn, 1993).

The new layout of text required luminance properties of each condition to be re-calculated. There were three luminance conditions[9]:

1. *Normal (baseline) condition:* luminance properties to match normal text without an overlay
2. *Overlay condition*: luminance properties to match normal text with an overlay
3. *Reduced contrast condition*: luminance to match normal condition and RMS contrast to match the overlay condition

The text and background luminance of the normal and overlay condition were replicated from printed normal text and printed normal text with an overlay in office conditions (See sections 2.4.2 and 2.4.3). Condition 3 had unintended luminance properties, due to the use of RMS to determine the contrast of Condition 2. 'Michelson text contrast', determined to be the most accurate way to define screen contrast post testing (section 2.4.6.2), calculated the text contrast of Condition 3 to be 93.4% lower than condition 2. Luminance and contrast measurements of the three conditions are displayed in Appendix XI.

[9] The fourth condition of Experiment 1 (reduced luminance and normal text contrast) could not be achieved with times new roman font due to its higher contrast.

3.1.5.3 Procedure

The sequence of Experiment 2 is displayed in Figure 3.1. Participants completed the Word Search and Comprehension tasks in counter-balanced order. The order of conditions in each task was counterbalanced so that each order was completed the same number of times. Prior to completing each task, participants were given verbal and written instructions about the task. The experimenter sat with the participant during a practice trial in each luminance condition to ensure that the task was completed properly.

3.1.6 Experiment 2: Analysis

3.1.6.1 Data preparation

For the successive word reading task, the first and last reaction times from 'Z' key presses were removed from each trial, resulting in 4516 reaction times. For the comprehension task, all trials passed the comprehension threshold criteria (two correct multiple choices out of three), resulting in 216 reaction times.

Table 3.3

Total data points and outliers for successive word reading task and text reading task (Experiment 2)

	Successive word reading task	Text reading task
Total data points	4516	216
Outliers	179 (3.96%)	17 (7.87%)

Note. Data points were reaction times of the time taken to read a line of text for the successive word reading task, or a page of text for text reading tasks. Outlying data points were either < 60 or >600 words/min, or more than 2.5 standard deviations from the participant's mean reading speed.

Participant 12 had five outlying reaction times (out of 9) overall, with three outliers within one luminance condition. When data were categorised according to task and luminance condition, reading speeds of participant 12 were above 3.8 standard deviations from the mean of all the participants for each condition. Participant 12 was therefore judged not to have completed the task properly and their data were removed from the analysis.

Apart from participant 20, none of the remaining participants' data had more than two (out of nine) outlying reaction times overall, and no more than one (out of three) outlying reaction times within one condition. Participant 20 had three outlying reaction times, two of which were in one condition. This meant that when the data were categorised according to task and luminance condition, the average reading speed of one of the conditions was based upon only one value. However, z scores reflected that this reading speed was within the acceptable range, so there was no reason to suspect that participant 20 had not completed the task correctly for that trial. Participant 20's data were therefore retained. In any case, the removal of participant 20 or participant 12's data from the analysis made no difference to the outcome of the experiment. See (Figure 3.5) for grouped remaining data.

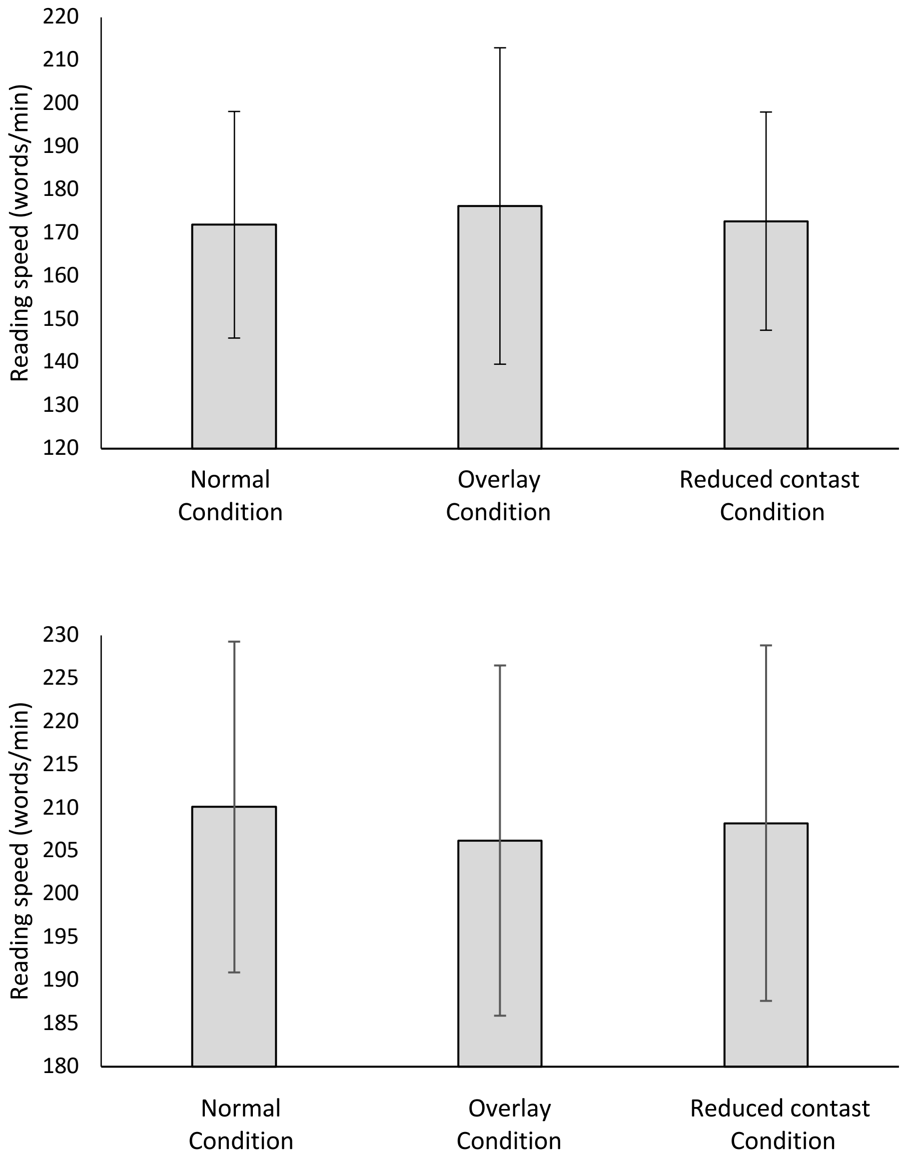


Figure 3.5. Reading speeds of Successive reading task (top) and Text reading task (bottom), (Experiment 2) in Normal, Overlay and Reduced contrast condition.

3.1.6.2 Did luminance affect reading speed?

To determine if a change in luminance elicited statistically significant changes in reaction speed for these tasks and participants, a repeated measure ANOVA was conducted on each task. Four out of six of the groups (displayed in Figure 3.5) were assessed to be normally distributed ($p > .05$). Condition 1 ($p = 0.017$) and 2 ($p < .001$) of the successive word reading task violated the assumption of normality. The analysis continued because ANOVAs have been shown to be robust against violations of normally distributed data (Schminder et al., 2010).

Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for the successive word reading task, $\chi^2(2) = 14.84$, $p = .001$, but not for the text reading task, $\chi^2(2) = 2.199$, $p = .333$. Change in luminance did not elicit statistically significant changes in reaction speed for the successive word reading task $F(1.327, 29.203) = .188$, $p = .829$, $\eta_p^2 = 0.08$, (Greenhouse-Geisser corrected) or text reading $F(2,44) = .082$, $p = .921$, $\eta_p^2 = 0.04$ (sphericity assumed).

3.1.6.3 Were the tasks related?

A Pearson correlation showed that reaction times on the two tasks were not significantly correlated, n = 23, $r = 0.89$, $p = .687$. To compare the reading speeds of both tests, the data were collapsed across the luminance conditions for each task type and a within groups *t*-test was administered. Reading speeds obtained from the successive word reading task were significantly slower ($M = 175.80$, $SD = 66.81$) than those obtained from the text reading task ($M = 209.87$, $SD = 44.97$), $t(22) = 2.188$, $p = .046$, $d = -0.44$.

3.1.6.4 Order effects

To determine if there were any order effects for each task, each condition was ranked according to the order that it had been completed. This resulted in three groups specifying the data which had been completed in the first, second and third conditions. A repeated measures ANOVA was conducted to assess if the order in which the conditions had been completed affected their reading speed. Mauchly's test of sphericity determined that the assumption of sphericity was met for the data of the successive word reading task ($\chi^2(2) = 0.034$, $p = .983$) and the text word reading task ($\chi^2(2) = 3.838$, $p = .147$). The order in which the conditions had been completed was found to elicit a statistically significant effect in both successive word reading task, $F(2,44) = 6.372$, $p < .004$, $\eta_p^2 = .225$, and the text reading task, $F(2,44) = 5.185$, $p = .010$, $\eta_p^2 = .191$. Pairwise comparisons revealed that, for the successive word reading task, the condition which had been administered last was significantly faster than the other two conditions. For the text reading task, the condition which had been administered last was found to be significantly faster than the condition which had been administered second.

Table 3.4.

Pairwise comparisons between the order in which the conditions had been completed in both reading tasks of Experiment 2. a: the successive word reading task, and b: the text reading task.

a.

Condition of comparison		t(22)	Adjusted target Alpha	p	d
1	3	-3.456	0.05	0.002*	-0.721
2	3	-2.458	0.025	0.022*	-0.513
1	2	-1.066	0.017	0.298	-0.222

b.

Condition of comparison		*t*(22)	Adjusted target Alpha	p	d
2	3	-3.131	0.05	0.005*	-0.653
1	3	-2.344	0.025	0.028	-0.489
1	2	0.014	0.0167	0.989	0.003

3.1.7 Experiment 2: Discussion

No significant correlation was found between average reading speeds (across all luminance conditions) obtained from the successive word reading task and the text reading task. The absence of a correlation between the two tasks implies that there is no relationship between the reading processes required for successive word reading and text reading for typical adult readers. Reading speeds from the successive word reading task were found to be significantly slower than reading speeds from the text reading task.

A practice effect was found across the duration of the successive word reading task, and between the condition that had been conducted second and third in the text reading task. In general, significantly faster reading speeds were recorded in the condition that was completed last than one or more of the other conditions in both tasks. This indicates that developing the successive word reading task to include more trials did not have the desired effect of eliminating the practice effect.

Neither of these tasks are suitable for measuring the reading benefit of coloured filters. Although the text reading task is a naturalistic reading task, it is not sensitive enough to provide an accurate change in reading speed and accuracy over time. Although the successive word reading task is a sensitive measure, it appears not to tap into the processes required for naturalistic reading. Reduced contrast and luminance did not affect reading speed of typical readers using the tasks and range of values described here.

3.2 Experiment 3: Change blindness tasks

3.2.1 Introduction

Some studies have demonstrated that easily perceptible changes can go unnoticed if the change is gradual (Simons, Franconeri & Reimer, 2000). 'Change blindness', the phenomenon where an easily perceptible change signal goes undetected, was originally demonstrated when attention to the change signal was actively disrupted by inserting a blank screen between the changed images (Rensink, O'Regan, and Clark, 1997). Simons et al. (2000) demonstrated that change blindness was equally likely to occur with a gradually changing signal, even when participants were actively searching for the signal.

To disentangle whether the effects of increased task performance with a chosen coloured filter is a placebo effect or not, Experiment 3 applied a gradually changing transition between the luminance conditions of chosen and no-filter. If participants did not notice the colour change, their expectation of the effect of the chosen filter should be extended to the no-filter condition, thereby creating a placebo-controlled design.

Simon et al. (2000) demonstrated change blindness with a signal that changed over a 12 second interval among participants who were searching for the change. In Experiment 3, the colour change was completed over a longer interval of three minutes and participants were not made aware of the colour change. The colour change was over the whole screen and occurred during completion of tasks so that, according to Gibbs, Davies and Chou (2016), no localized motion or time signals could attract attention to the change.

Tasks were a cloze task and visual search task (section 3.2.2.2), and dependent variables were reading speed (cloze reading task), reaction time gradient and intercept (visual search task). If filter users were found not to notice the colour change and to read significantly quicker with their chosen coloured filter, then this would demonstrate an effect of their chosen filters beyond placebo.

3.2.2 Method

3.2.2.1 Participants

All participants were between the ages of 18-35, self-reported to have had an eye-test within the last two years and had normal, or corrected to normal, vision.

There were two participant groups: ‘Filter users’ (readers who used coloured filters to read) were reimbursed with payment, and ‘control participants’ (psychology undergraduates who did not use coloured filters to read) were reimbursed with a credit towards a course requirement. Filter users were 24 participants who had been administered with a chosen coloured filter by the Irlen method in a disability service or the Intuitive method in an Opticians. Seven participants used Irlen overlays, eight participants used Irlen lenses, five participants used Intuitive overlays and four participants used Intuitive lenses. Filter users were recruited to have used their filter for more than six months.

One participant in each group was bilingual. One filter user self-identified as colour blind, but this participant could tell the difference between the two conditions after they had completed the experiment. Participants in each group were matched for sex and matched, as close as possible, by age. Table 3.5 displays the distribution of their ages and the number of participants who self-identified as dyslexic.

Table 3.5

Number of participants in each participant group, of 24 people, who were dyslexic and the means and standard deviations (S.D) of their ages (Experiment 3).

Participant group	Dyslexia	Age	
		Mean	S.D.
filter users	19	25.67	7.00
Control	1	22.67	3.53

3.2.2.2 Stimuli

Cloze task

Andrews and Hersch's (2000) three-minute cloze task was developed to measure reading performance. The test has a test/re-test reliability estimate of .79 (Andrews & Hersch, 2000), thereby demonstrating its suitability for comparing reading performance between conditions. The task involved silently reading the text for understanding and periodically completing word selection tests that were placed every 50 words (Figure 3.6). The word selection tests contained three word options: one correctly completed the text and the remaining two incorrectly completed the text but were globally or locally relevant (Andrews, pers comm, 2014).

The task was developed to provide a continual measure of reading performance over the course of 15 minutes. To prevent participants from scanning or reading ahead, the critical word options were initially replaced by random letters. For each word selection test, a key press revealed the word options and recorded the time taken to read the previous 50 words, and another key press confirmed the chosen word and recorded the time taken to choose the answer. A right arrow key press then displayed the next page of text.

One of Margaret Atwood's short stories, 'The Dead Hand Who Loved Me' (Atwood, 2014) was edited to the format of the test. There were 6211 words in the whole text. Text was presented to appear as though it was printed in a book (see Figure 3.6).

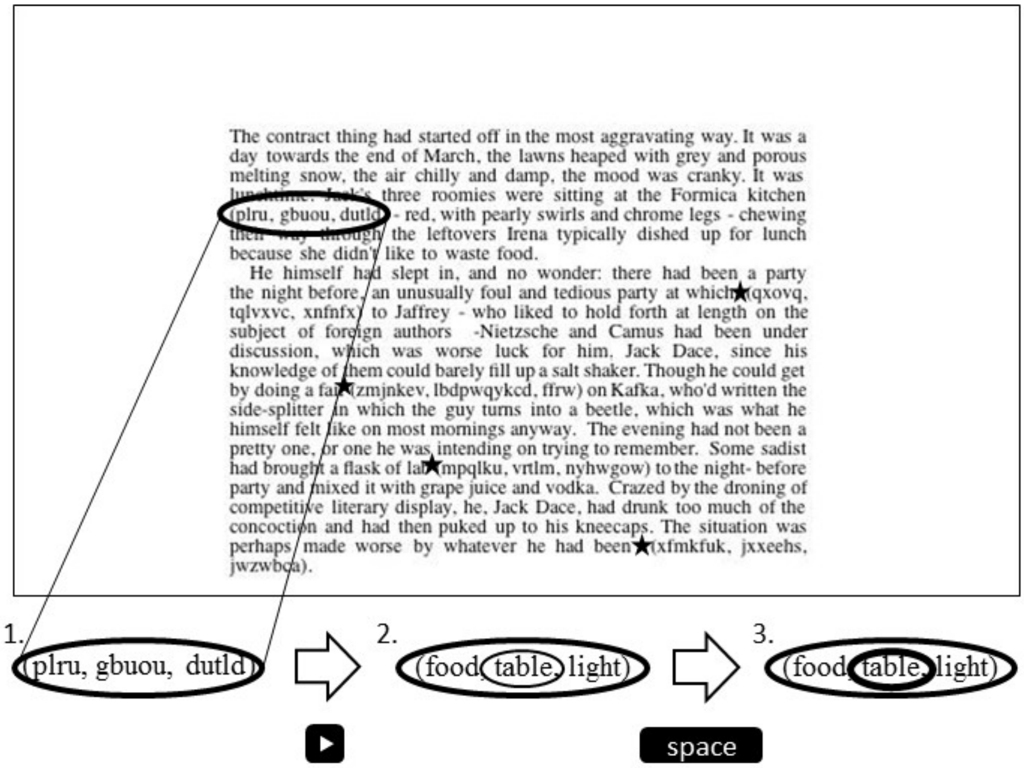


Figure 3.6. Stimulus and key presses of one trial of the cloze task in Experiment 3. Word selection tests were placed every 50 words in the text. 2. A 'right arrow' keypress revealed the word options, displayed a light circle round the first word option and recorded the first time-stamp. Further right arrow key presses circled the following choice words. 3. space A 'space bar' keypress confirmed the choice, displayed a bolded circle round the chosen word, recorded if the chosen word was correct and recorded the second time stamp.

Visual search task

A standard paradigm of forced choice visual search was used. A target stimulus was randomly located amongst distractor stimuli in one of two fields (see Figure 3.7). The target was a 'T' symbol, and the distractors were 'L' symbols.

Symbols were pseudorandomly displayed with 0°, 90°, 180°, 270° orientations. Display size was either 10, 20 or 30 targets. The field in which the target was displayed was pseudorandomly generated. Each trial required the participant to identify the field in which the target appeared. A 'Z' key press indicated that the target was in the left field, and an 'M' keypress indicated that the target was in the right field. A correct response was followed by a blank screen for 0.5 seconds, and feedback for an incorrect response was a short beep and a blank screen for one second. There were 92 trials.

In line with convention, the reaction times for locating a target within different numbers of distractors (display size) are graphed as a function of the display size to obtain a measure of search rate (gradient) and intercept (Wolfe, 2010).

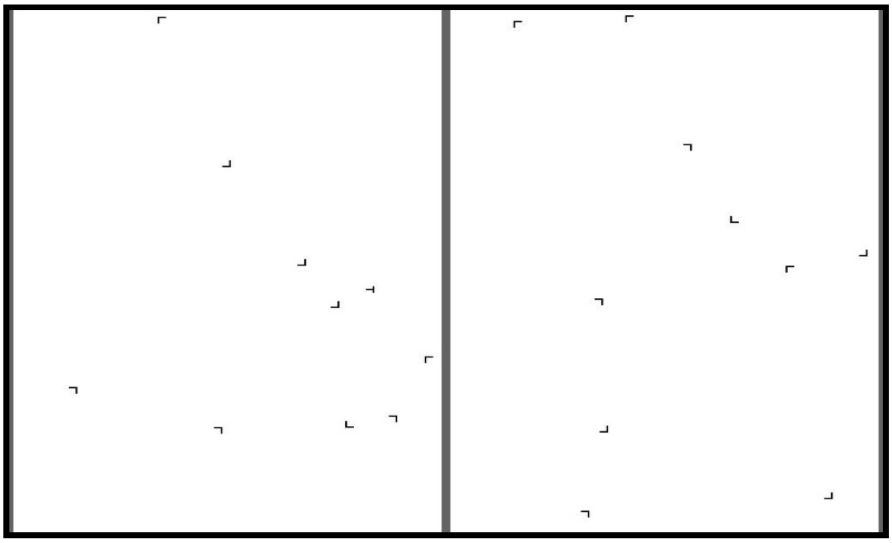


Figure 3.7. Stimulus used for one trial of search task in Experiment 3. For this trial, a 'z' keypress would correctly indicate the location of the target in the left field and record a reaction time.

Luminance conditions and transitions

The background luminance of the no filter (baseline) condition was 119.625 cd/m^2 and Michelson text contrast was 0.258. This was the same baseline condition that had been used in Experiment 2 and calculated in section 2.4.2. The background luminance of the filter condition was the replicated luminance of this baseline condition overlaid by the filter user's own filter. Luminance and contrast measurements of the baseline and filter condition are displayed in Appendix XIV and chromaticities of filters are displayed in Appendix XII.

Participants completed tasks in filter and no filter conditions with a gradual, three minute screen change between. The screen change was achieved by manipulating the colour look-up table so that the hue remained the same. The luminance of the screen background was in five stages of three-minute reading intervals for the cloze task, or 18 visual search trials for the search task. Part 1 transitioned from the filter to no filter condition and returned to the filter condition, and part 2 replicated this pattern in reverse (Figure 3.8).

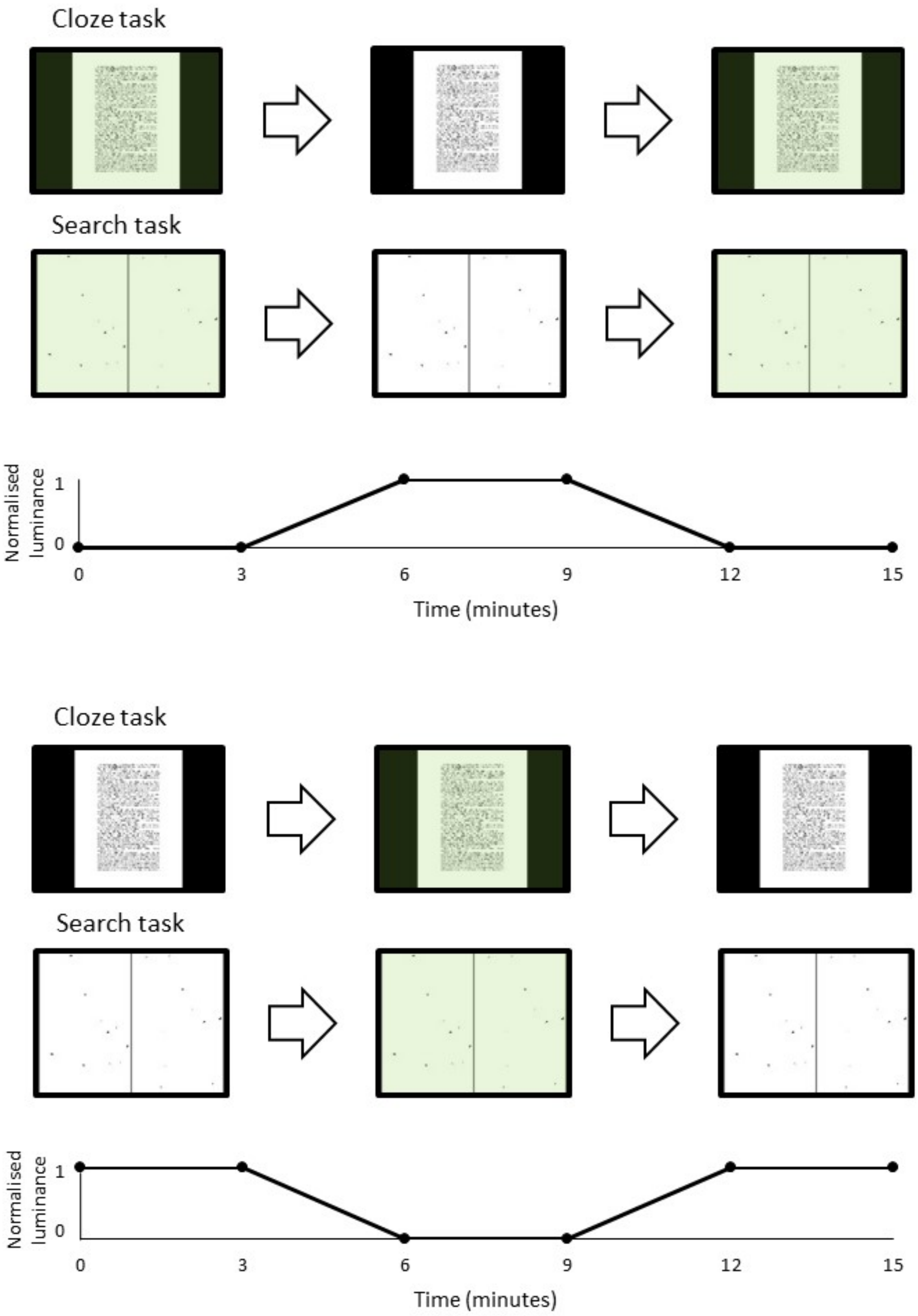


Figure 3.8. Stimuli, luminance conditions and their timings of cloze and visual search tasks (Experiment 3). ▭ and ▭ represent the filter and no filter conditions and ⇨ represent the transitioned from one luminance condition to the other. Graphed luminance is normalised.

3.2.2.3 Procedure

To create a counterbalanced design, the experiment was conducted in two sessions four weeks apart. The filter condition was tested first in the first session and the no-filter condition was conducted first in the second session (Figure 3.9). Participants were not informed about the colour change until they had completed the experiment. They completed the cloze-reading and search task in counter-balanced order. Before commencing the task, participants were asked if they had read the story before; none had. Prior to completing each task, participants were verbally instructed according to the written instructions in the programme. The experimenter sat with the participant during three practice trials to check that the task would be completed properly. To ensure that typical reading speeds were measured, they were not explicitly made aware that responses were timed.

Following the completion of both tasks, participants were given a questionnaire to record any diagnosis of dyslexia by asking if they had had extra time in exams. If they answered yes, they were then asked if this was due to a diagnosis of dyslexia. To elicit whether participants noticed the colour change or not, they were then asked if they found reading easier or faster in some parts than others and their responses were recorded. If participants did not mention the colour change at this stage, they were then asked if they noticed a change in the screen. Those who did not mention a colour change were recorded as having not noticed it.

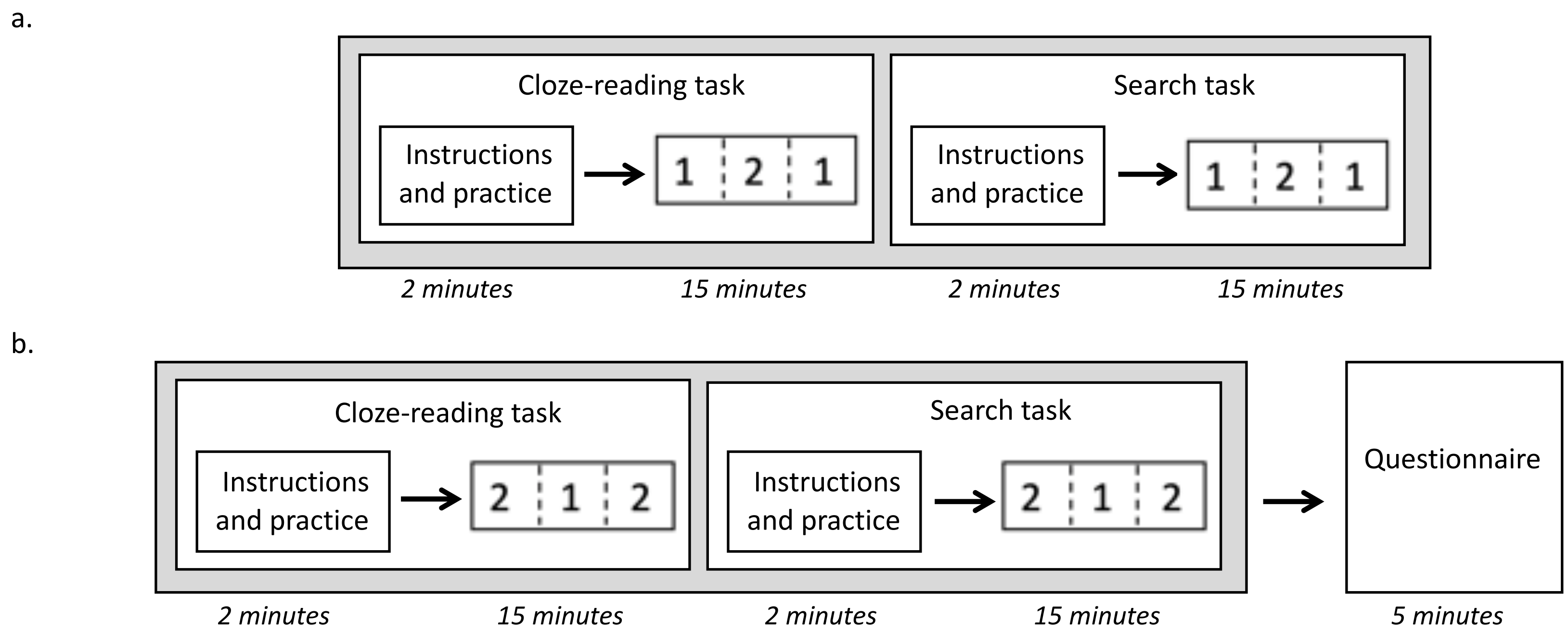


Figure 3.9. Sequence and timings of Experiment 3. **a.** Part 1. Tasks were cloze-reading and visual search. b: Part 2 (conducted less than four weeks later). Tasks were cloze-reading, visual search and questionnaire. Numbers 1 and 2 represent filter and no filter condition. The dashed line represents the gradual transition between the filter and no filter condition. Grey borders illustrates counterbalancing of tasks. Timings are in italic font.

3.2.3 Analysis

3.2.3.1 Chosen Filters

Appendix XIV displays a description of luminance and contrast measurements of the chosen filters and the baseline condition used in this experiment. Appendix XII displays the chromaticities of the filter users' filters in CIE (1976) and MacLeod-Boynton (1979) chromaticity diagrams; Appendix XIII displays the distribution of each filter type[10] within these colour spaces. The chromaticities span the ordinate and abscissa of both colour spaces and there was no clearly observable pattern or clustering. There was a high range of saturations amongst the filters, as demonstrated by chromaticities covering the whole area of each chromaticity diagrams.

The numbers of filter users who chose each colour, along with their average reading times, are displayed in Appendix XIX. The most popular chromaticities were aqua and orange, chosen by 33.3% of the sample. There was no observable pattern in baseline reading speeds or increased reading speeds across the chromaticities. The lack of a distinct pattern within each colour space could be due to the small sample size (n = 24). Chapter 7 continues the analysis by combining data from the three experiments involving chosen filters (Experiment 3, 4 and 6).

[10] The chromaticities of these filters were replicated on screen as the text background

3.2.3.2 Colour change

Only eight people in the experimental group and 11 people in the control group failed to notice the colour change (Table 3.6). A chi-square test for association was conducted to ascertain if the percentage of participants who noticed the colour change was statistically less in the experimental group than the control group. All expected cell frequencies were greater than five, and there was no statistically significant association between the two factors $\chi^2(1) = .784$, $p = .376$.

Table 3.6.

Number of participants who noticed and did not notice the colour change in the experimental and control participant groups (Experiment 3)

Participant group	Total	Noticed	Did not notice
Experimental	24	16	8
Control	24	13	11

The original question was whether filter use increased reading speeds for the experimental group when they did not notice the colour change; if this was found to be the case, it would provide placebo controlled evidence that filters benefit reading beyond placebo. Accordingly, the intention was to discard results from participants who noticed the colour change and perform a mixed ANOVA with factors of filter use and participant group. However, since an unexpectedly high (more than half) proportion of participants noticed the colour change, an alternative analysis was formulated to prevent the analysis from being underpowered.

Whether participants noticed the colour change (change detection) was introduced as another factor in the design to investigate if there was an interaction between filter use and change detection for filter users. If there was a significant placebo effect, then filter users who noticed the colour change should show significantly improved reading rates in the filter condition as compared with the no-filter condition, and this effect of filter use should not be observed either in filter users who did not notice the colour change, or in control participants.

The data for each task were originally organised into six data-sets according to luminance condition (filter, transition and no filter) and participant group (experimental and normal). However, across all analysis, the data collected during the transition phases approximated to the average of the experimental and normal condition. These datasets were considered not to be relevant or interesting to the purpose of the experiment, and so were removed from the analysis. Therefore, the data of each task was organised into four datasets according to luminance condition (filter and no filter) and participant group (experimental and control). A 2x2 mixed ANOVA with the within-subjects factor of filter use (filter and no filter) and between-subjects factor of whether they noticed the colour change (noticed/did not notice) was conducted for each participant group (experimental/normal) for each task.

3.2.3.3 Cloze task

All data were assessed to represent reading times when the text had been understood sufficiently, meeting the threshold of more than 70% correct trials per participant. Incorrect trials were evenly distributed across the groups (

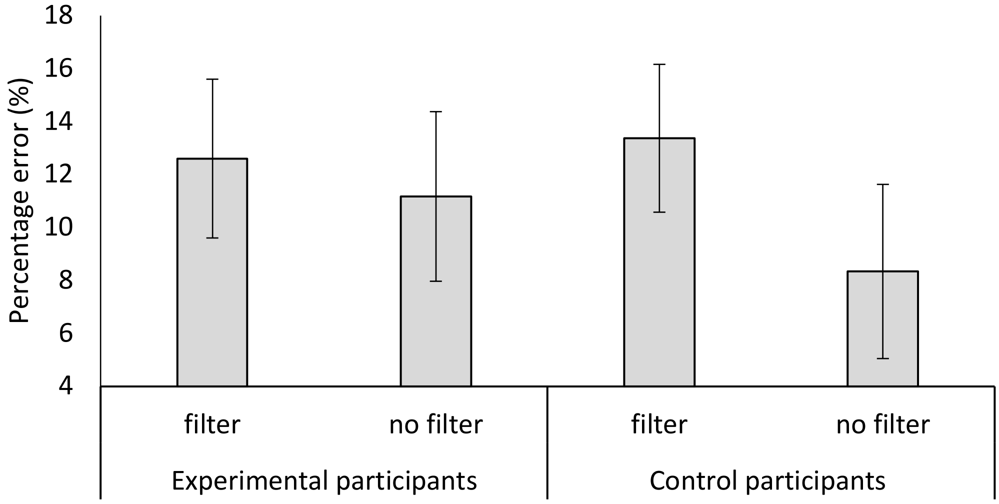


Figure 3.10) and were removed from the analysis.

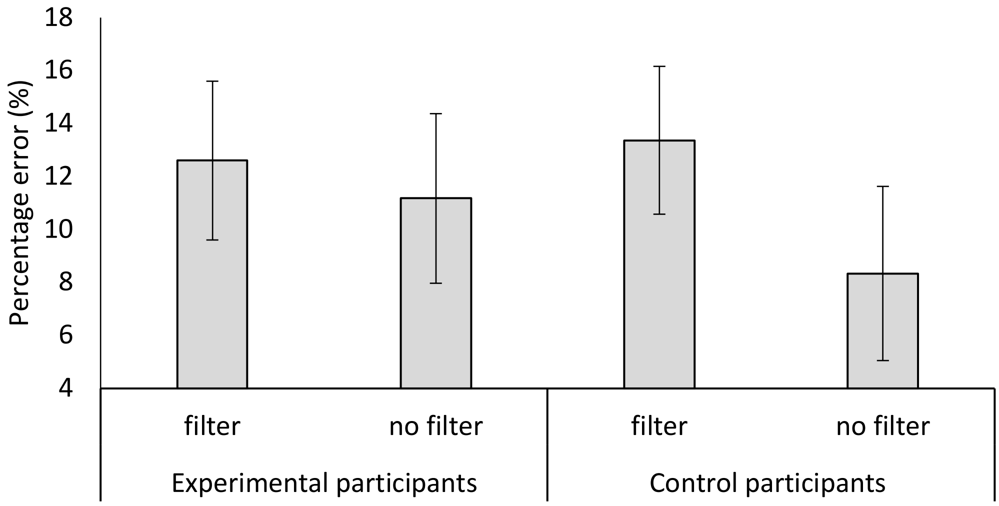


Figure 3.10. Percentage errors of cloze task across filter conditions and participant groups (Experiment 3).

For each participant, reading speeds were highly variable, and so median reading speeds were calculated for each luminance condition, so that extreme values did not skew the results. There were no outlying median reading speeds, as they were all fewer than 2.5 standard deviations from the mean within each participant group and luminance condition.

The combined time taken to read each 50 words and complete the text was used to calculate median reading speeds. The inclusion of the time taken for text completion was considered to most accurately represent reading speed when text had been understood because some trials would involve comprehension processes during the text completion stage. However, the inclusion of this time did not change the outcome of the analysis.

There were some missing data in the second part of the experiment due to two filter users not returning for the second part and some participants (one experimental and five controls) finishing reading the text before the experiment time had finished. Since this was a counterbalanced design, the data were assessed for order effects to determine if the loss of data from these participants (three in each participant group) was likely to impact the outcome of the experiment. The data were categorised according to the order in which the conditions had been completed. Pairwise comparisons revealed that order effects existed over the duration of the task (Figure 3.11 ;Table 3.7). For this reason, the data of these participants were removed from the analysis.

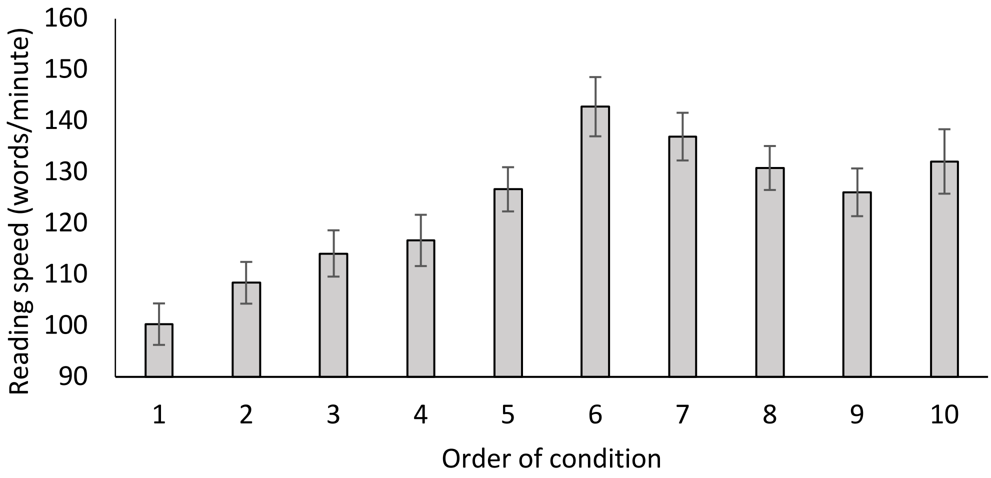


Figure 3.11 Reading speeds of all participants when categorised according to the order in which the conditions had been completed (Experiment 3).

Table 3.7

Paired comparisons between the order conditions (Experiment 3)

Order condition of comparison		*t*(22)	Adjusted target Alpha	*p*	*d*
1	3	-3.485	0.05	0.001*	0.550
3	5	-3.510	0.025	0.001*	0.555
6	8	3.007	0.0167	0.005*	0.475
8	10	-0.288	0.0125	0.776	0.045

Table 3.8 shows the categorisation of the remaining participants according to their participant group and whether they noticed the colour change. The exclusion of these participants did alter the outcome of the analysis and so *p* values are reported accordingly.

Table 3.8

Number of participants who completed the experiment, who noticed and did not notice the colour change in the experimental and control participant groups (Experiment 3)

Participant group	Total	Noticed	Did not notice
Experimental	21	14	7
Normal	19	10	9

Figure 3.12 displays the reading speeds of cloze task categorised according to filter condition, participant group and whether participants noticed the colour change. Normality of the eight data sets, split according to filter type, change detection and participant group, was assessed using Shapiro-Wilk's test (p >.05), identifying one group, filter users who noticed the colour change when the filter was applied, to be not normally distributed (p = .030). Due to the robustness of ANOVAs in the face of departures from normality, it was assumed that this violation would not affect the outcome of the analysis. There was homogeneity of variances with filter (p = .268) and without filter (p = .120) as assessed by Levene's test of equality of variances (p > .05). There was homogeneity of covariances, as assessed by Box's test of equality of covariance matrices (p = .246).

The normal participant group had no statistically significant interaction between the factors of change detection and filter use $F(1,17)$ = .970, p = .338, η_p^2 = .054. But there was a main effect of luminance: reading speeds were significantly faster when the filter had not been used (M = 127.26, SD = 19.27) than when it had (M = 146.84, *SD* = 24.85), $F(1,17)$ = 51.474, p = <0.001, η_p^2 = .752.

The filter using group had a statistically significant interaction between the factors of change detection and filter use, $F(1,19)$ = 6.726, p = .018,

η_p^2 = .261. To explore this interaction further, a test of simple effects was conducted. Those who noticed the colour change read quicker with the filter (M = 114.43, SD = 21.34) than without the filter (M = 108.63, SD = 24.62), but the difference was not significant (*p* = .286). However, those who did not notice the colour read significantly faster when the filter had not been used *(*M = 125.21, SD = 33.32) than when it had (M = 107.27, SD = 24.61), (*p* = .027). With no excluded participants, this difference was not found to be significant (*p* = .079).

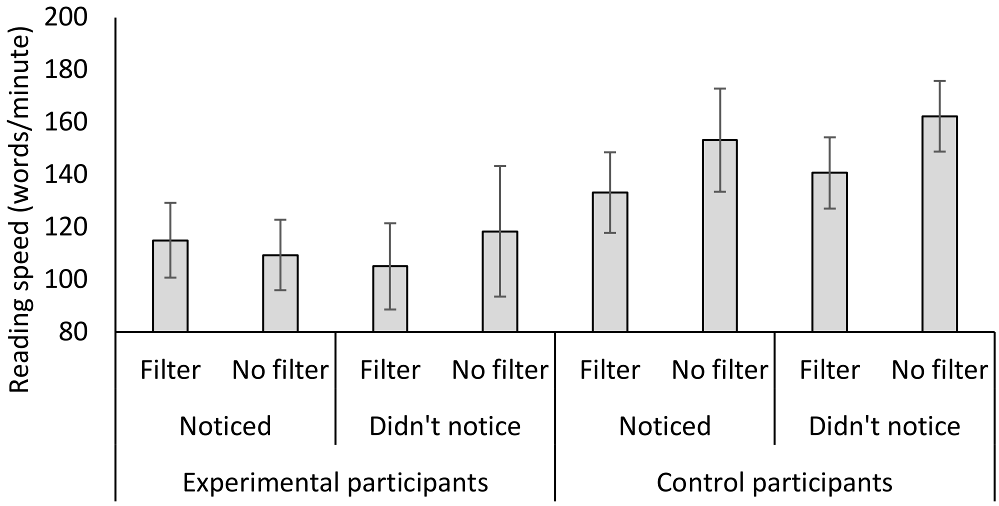


Figure 3.12. Reading speeds of cloze task categorised according to filter condition, participant group and whether participants noticed the colour change (Experiment 3)

To determine if there was a difference between reading speeds of the experimental and control participants, data were collapsed across conditions for the two participant groups and a *t*-test was conducted. When the data had been categorised according to luminance condition and participant group, there had been no outliers. However, when the luminance conditions were collapsed and the data were only categorised according to participant group, one participant had an outlying z-score of -2.67. This was considered to be marginal and so this participant's data point was retained for the analysis. However, removal of this participant from the analysis did not change the analysis outcome. Reading speeds for the normal and filter using group were normally distributed, as assessed by Shapiro-Wilk's test (p = .680, p = .144). There was homogeneity of variances, as assessed by Levene's test for equality of variances (p = .244). The normal group (M = 134.69, S.D. = 20.88) was found to read significantly faster than the experimental group (M = 112.43, S.D = 24.60), t (38) = 3.068, p = .004, g = .997.

3.2.3.4 Visual Search

Incorrect trials, when an incorrect window had been selected, were removed from the analysis. The distribution of these errors was evenly distributed across the whole data set and is displayed in Figure 3.13.

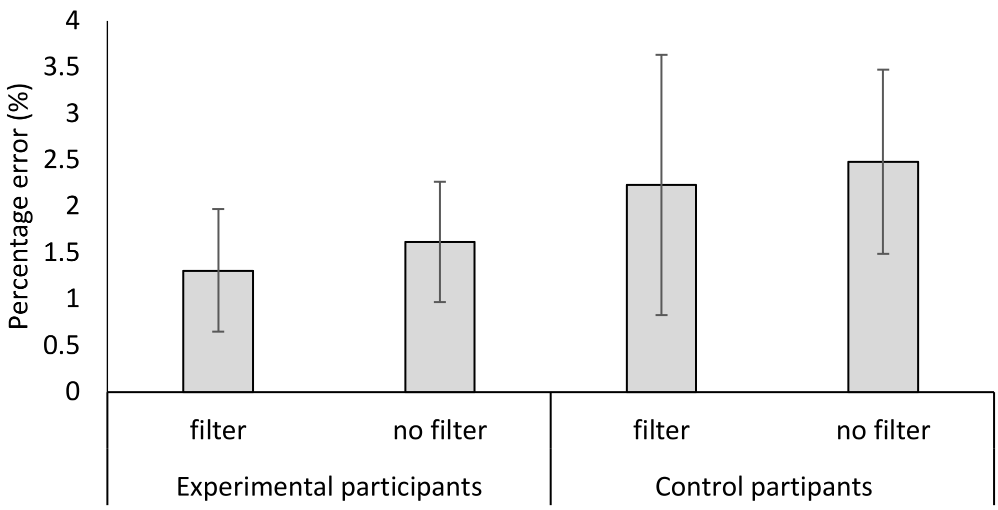


Figure 3.13. Percentage errors of search task across filter conditions and participant groups (Experiment 3)

Median reaction times of 10, 20 and 30 targets in each luminance condition were used to calculate each participant's median gradient and intercept. Data from four participants were removed from the analysis due to incomplete data sets and outlying reaction times, but their removal did not affect the outcome of the analysis. Two filter users had incomplete data sets because they had not returned for the second part of the experiment. Since this was a counterbalanced design, the data collected in the first and second part were assessed to determine if the loss of these data would affect the outcome of the experiment. With the two incomplete data sets removed, a paired *t*-test determined that reaction times collected in the first part (M = 3.074, S.D = 0.855) were significantly higher than reaction times collected in the second part (M = 2.8457, S.D. = 0.792), $t(45) = 3.574$, $p = .001$, $d = .527$. To maintain a counterbalanced design, the data of these participants were removed.

Intercepts and gradients were calculated. Participants 10 and 20 were classified as outlying participants with a median gradient or intercept of more than 2.5 standard deviations from the mean of all the participants within a luminance condition and within each participant group (Table 3.9).

Table 3.9

Outlying participants of visual search task (Experiment 3)

Participant	Participant Group	Dependent Measure	Luminance Condition	Z-score
20	Experimental	Gradient	No-filter	3.794
20	Experimental	Intercept	Filter	3.052
20	Experimental	Intercept	No-filter	2.719
10	Control	Gradient	No-filter	3.994
10	Control	Intercept	No-filter	-3.257

Note. Further to removal of incomplete data sets, outlying participants were identified to be those with more than 2.5 standard deviation from the mean of each data set.

For both gradients and intercepts, eight data sets were categorised, according to filter type, change detection and participant group, were assessed to be normally distributed according to Shapiro-Wilk's test (p > .05). There was homogeneity of variances as assessed by Levene's test of equality of variances (p > .05). and homogeneity of covariances, as assessed by Box's test of equality of covariance matrices ($p > .05$). For both gradients and intercepts, there was no significant interaction found between change detection and filter-use and no main effect of filter-use in the filter using group (Table 3.10) and normal participant group (Table 3.11), ($p > .05$), Figure 3.14.

Table 3.10.

F statistics of mixed ANOVA with factors of filter-use and change detection for filter users (Experiment 3)

Measure	Factor (s)	$F(1,19)$	p	η_p^2
Gradient	Filter use*noticed	0.036	.852	.002
Gradient	Filter use	0.056	.815	.003
Intercept	Filter use*noticed	1.828	.192	.088
Intercept	Filter use	0.000	.983	.000

Table 3.11.

F statistics of mixed ANOVA with factors of filter-use and change detection for control participants (Experiment 3)

Measure	Factor (s)	$F(1,21)$	p	η_p^2
Gradient	Filter use*noticed	6.726	.018	.261
Gradient	Filter use	0.617	.441	.029
Intercept	Filter use*noticed	2.419	.135	.103
Intercept	Filter use	0.558	.463	.026

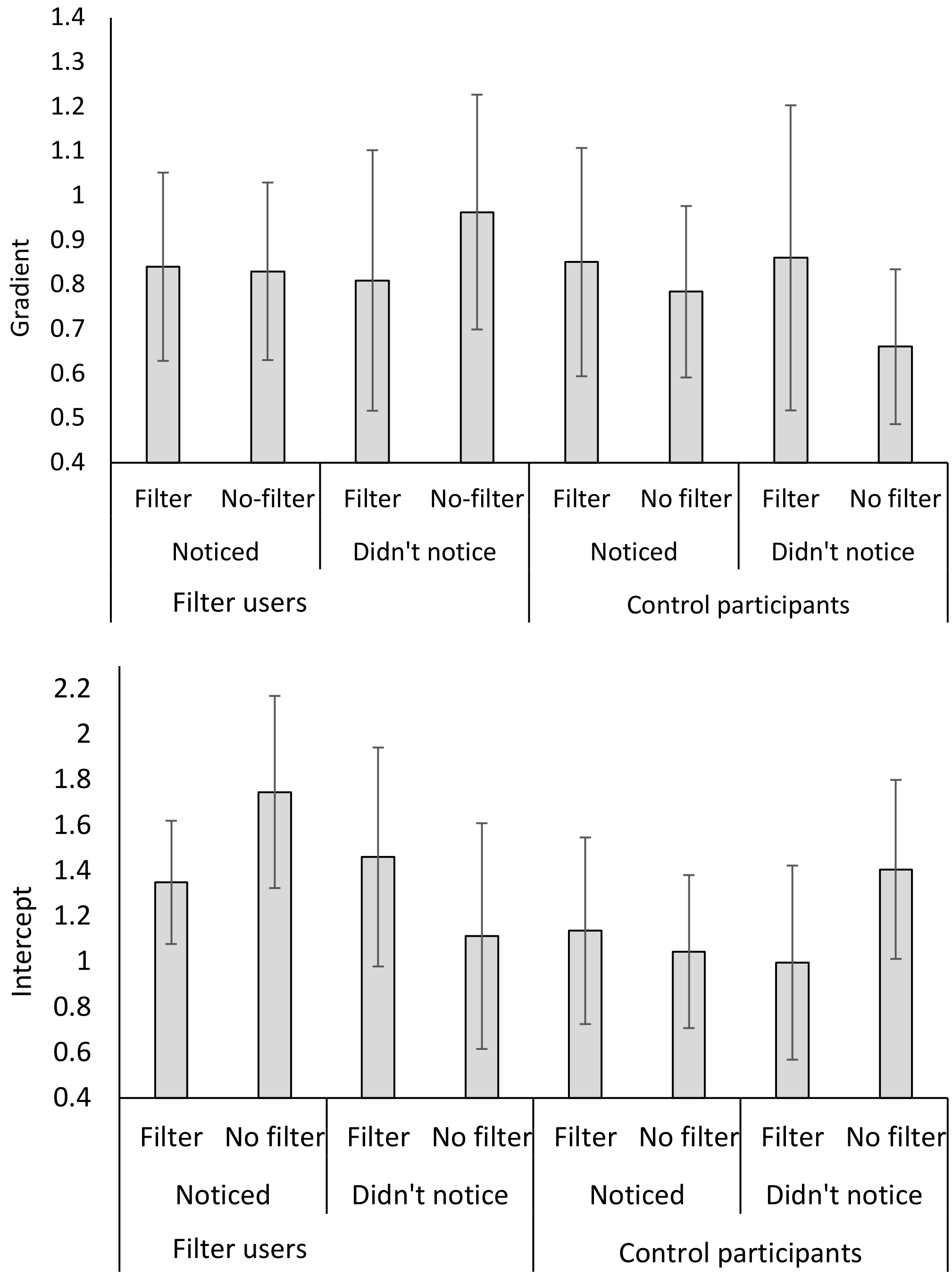


Figure 3.14. Gradients and Intercepts of search task categorised according to filter condition, participant group and whether participants noticed the colour change (Experiment 3)

To determine if there was a difference between the visual search performance measures of the experimental and control participants, data were collapsed into the two participant groups and *t*-tests were conducted. Gradients and intercepts for both control and filter using groups were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). There was homogeneity of variances for both measures as assessed by Levene's test ($p > .05$). Neither gradients or intercepts were found to be significantly different between the experimental and control group Table 3.12.

Table 3.12.

Independent t test between the data of each participant group for visual search gradient and intercept (Experiment 3)

Measure	Participant group	Mean	S.D	$t(44)$	p(2-tailed)	d
Gradient	filter users	0.890	0.211	-0.321	0.75	0.144
	control	0.855	0.258			
Intercept	filter users	1.382	0.442	1.501	0.141	0.457
	control	1.205	0.332			

3.2.4 Discussion

Contrary to expectations, most participants noticed the colour change. filter users were not found to notice the colour change more than control participants, so it appears that filter users were no more sensitive to the presence of a coloured filter than control participants. Whether they noticed the colour change or not, neither participant group read quicker with their chosen filter. This implies that the chosen coloured filter had no beneficial effect for participants who use chosen coloured filters.

Regardless of whether the six incomplete datasets were excluded, control participants read significantly quicker without the filter than with the filter. This suggests that filters impeded reading performance for readers who have not chosen the filters for reading.

Regardless of whether the incomplete datasets were excluded a significant interaction was found between the factors of change detection and filter use in the filter using group. This implies that filter users who noticed and did not notice the change in conditions reacted differently to the use of their filter. When all data were used in the analysis, there was no effect of filter use on either of these experimental groups, but when incomplete datasets were excluded, filter users who did not notice the colour change behaved similarly to the control participants by reading significantly quicker without their chosen filter.

The absence of a significant effect of the seven filter users who noticed the colour change may represent an effect of 'nocebo'. In this case, the negative expectations of reading without the filter had a detrimental effect on reading performance. This negative expectation could be fuelled by these participants' personal investment in the filter, and, being poorer readers than the control participants, as indicated by their slower performance on the reading task and higher proportion of self-reported dyslexia, their intrinsic motivation to benefit from them. An alternative scenario is that the analysis was underpowered since only seven filter users did not notice the colour change.

It is plausible that an effect of coloured filters on reading performance beyond placebo was not observed in this experiment because the reading time was not long enough to develop symptoms of visual stress with no filter. Increasing the number of people who don't notice the colour change, the reading time in each condition and reducing the order effect may bring clarity to these results.

Chapter 4: Is a chosen filter beneficial to reading beyond placebo?

4.1 Experiment 4

4.1.1 Introduction

4.1.1.1 Change blindness design

The first part of Experiment 4 intended to achieve the placebo-controlled design of Experiment 3: 14 out of 24 participants had noticed the colour change in and so these participants' expectations of the effect of the conditions were not controlled. Experiment 4 aimed to obtain a larger sample of participants who did not notice the colour change to conduct a *t*-test of their performance in placebo-controlled conditions.

Factors affecting the likelihood of change detection are not well understood. Simon et al. (2000) reported different percentages of change detection among different gradually changing, change blindness designs, but did not identify factors that affected change detection. Therefore, in the absence of relevant theory, it was assumed that reducing the number and magnitude of the change signal would increase the chance of the change signal going undetected. The colour change happened only once and the magnitude of the colour change was reduced by applying a neighbouring, 'placebo' hue of filter to the control condition. The chromaticity of this placebo filter that was calculated according to Wilkins et al.'s (2002) methods. Here, a CIE LUV colour difference of 78 was applied to the placebo filter from the chosen filter; this difference was claimed to eliminate the benefit of the chosen filter whilst being unnoticeable when applied to lenses (Wilkins et al., 1994).

The task duration in each luminance condition was increased to 8.5 minutes. This was thought to increase the likelihood of finding an effect of visual stress by increasing its power and by, according to Tyrrell et al. (1995) and Wilkins et al. (1996), allowing sufficient time for visual stress symptoms to develop.

A filter choosing process was conducted at the beginning of the experiment to acquire an 'effective', chosen filter because the effective filter is claimed to sometimes vary over time (Irlen, 1991a; Wilkins, 2003). The filter choosing process replicated the colorimeter process (Wilkins et al., 1992; Wilkins, 1997; Wilkins 2001) rather than the Irlen method (Irlen, 1991a) because chromaticities of this system had been applied to Wilkins et al.'s (2002) method of acquiring the placebo filter, and due to the availability of chromaticities (Wilkins, 2015, personal communication) and methodology (Wilkins, 1997; Wilkins, 2001; Wilkins & Sihra, 2001) for replication. The chromaticities of the colour choosing process were matched to those used by the colorimeter (Appendix XII), which had a smaller variation of luminance and saturation, and more uniform distances between choices of hue than other filter methods (Wilkins et al., 1992).

It is important to note here that the chromaticities that were used for the filter choosing process and placebo filters did not have constant luminance and saturation (Appendix XII). Colour space are only approximately uniform and so Euclidean distances can only provide approximate colour and saturation differences. The CIE 2000 colour-difference formula (Luo, Cui and Rigg, 2001) alongside the CIE 1976 La'b' colour space may have provided better approximations as these were shown to out-perform the CIE 1976 Lu'v' and are currently recommended by the CIE (Witt, 2007). However, Experiment 4 adhered to the colorimeter method (Wilkins et al., 1992; Wilkins, 2001) because it was claimed to be sufficiently accurate and has been used to provide various results that appear to demonstrate an effect of chosen filters for reading (Evans & Allen, 2016).

In this experiment, tasks were the cloze task used in Experiment 3 and a visual search task, (Neisser, 1964). Neisser (1964) had reported that a 'Z' target in a visual search was found twice as quickly when distractor letters were round (e.g. 'C', 'O', 'R') than when they were angular ('M', 'N', 'X'), and suggested that more extensive operations were needed to distinguish 'Z' from the distractors in the latter. This task was adapted to the current experiment to ascertain whether a chosen filter affected overall reaction times or the percentage difference between reaction times of curved and angular letters.

4.1.1.2 Stroop designs

The second part of Experiment 4 used performance indicators of the Stroop interference effect to control for the placebo effect. The interference effect is the time delay of completing a task due to interference from other processes and is calculated as the reaction time difference between congruent and incongruent trials in a Stroop design (Stroop, 1935).

The interference effect has been related to reading ability by using Stroop tasks that use words but don't require reading; for example, in Stroop's (1935) original experiment, the task was to name, not read, the colour of a word. The automaticity hypothesis predicted that increased interference on tasks such as this were due to increased automatic reading processes (automaticy) of the irrelevant word (Logan, 1997; MacLeod & MacDonald, 2000). In contrast, the 'blocking hypothesis' predicted that this increased interference reflected slower word reading due to a reduced attentional mechanism used to block the reading of the word (Cohen, Dunbar & McClelland, 1990; Roelofs, 2003).

Some recent evidence has supported the blocking hypothesis: Stroop interference was negatively correlated with reading ability (Protopapas, Archonti & Skaloumbajas, 2007) and reduced with word reading practice (Protopapas, Vlahou, Moirou & Ziaka, 2014) in children who were 12-13 years old. In addition, dyslexic teenagers, 15 years of age, (Kapoula et al., 2010) and dyslexic children, 7-11 years of age, (Bub, Masson and Lalonde, 2006) showed significantly larger interference effects than control participants who had not been identified as dyslexic.

Since the relationship between the interference effect and reading ability is unclear, and the literature cited here was conducted on children, the effect of reducing visual stress on the interference effect is unknown. Nevertheless, if a chosen coloured filter were to benefit the reading performance of a single word, a change in the interference effect would be expected. An increased interference effect would support the automaticity hypothesis and a decreased interference effect would support the blocking hypothesis. Regardless of the direction, the change would be at a scale that could not be attributed to conscious processing and thus would be placebo controlled.

The original Stroop (1935) task was inappropriate to investigate the effect of coloured filters on reading because it involved the manipulation of colour. Instead, picture-word (Rosinski, Golinkoff & Kukish, 1975; Glaser and Düngelhoff, 1984; La Heij, 1988) and location-word (Palef and Olson, 1975; Virzi and Egeth, 1985; Lu & Proctor, 1995) Stroop designs were applied. A word reading task was added to the design to provide a comparison of single word reading performance in unmasked conditions. For these Stroop tasks, percentage interference effect (Stroop tasks) and reaction time (single word task) were measured, and filter conditions were filter and no filter.

The location Stroop task (Shor, Hatch, Hudson, Landrigan & Shaffer, 1972). presented a noun word in one of four congruent or incongruent locations. For example, if the word ‘left’ was presented at the left location, this would be a congruent trial, and if it was presented at any other location (bottom, right or top) it would be an incongruent trial (Figure 4.6). A trial required the word not to be read and the location (‘up’, ‘down’, ‘left’ or ‘right’) to be named. Similarly, the picture word Stroop task (Rosinski, Golinkoff & Kukish, 1975) presented a noun word in front of a congruent or incongruent picture. For example, if the word ‘cat’ was presented in front of a picture of a door, this would be an incongruent trial. A trial in the picture word Stroop task required the word not to be read and the picture to be named.

4.1.1.3 Measure of phonic ability

Dyslexia has been reported to predominantly involves impaired or delayed phonic acquisition (Ramus et al., 2003; Saksida et al., 2016) (Section 1.6), and so a separate measure of phonic ability was incorporated to more accurately identify people with dyslexia.

4.1.2 Method

4.1.2.1 Participants

Twenty nine filter users and 29 control participants (21 female), were recruited according to the criteria used in Experiment 3. Participants were between 18-35 years of age. Sixteen filter users used overlays and 13 used lenses. Sixteen filter users used Irlen filters; 10 filter users used Intuitive filters; three filter users used filters acquired from a dyslexia assessment and the method by which these were obtained was unknown.

Participants in each group were matched by sex and matched, as close as possible, by age. Table 4.1 displays their ages and the number of participants who self-identified as dyslexic.

Table 4.1.

Number of participants in each participant group, of 29 people, who were dyslexic and the means and standard deviations (S.D) of their ages (Experiment 4).

Participant group	Dyslexia	Age	
		Mean	S.D.
filter users	16	24.45	6.28
Control	1	19.8	1.21

4.1.2.2 Stimuli

Luminance conditions

There was no baseline condition since this experiment transitioned from a chosen to placebo filter. Chromaticities from the higher, not lower[11], saturation

[11] Chromaticities of the higher saturation setting were used because several chromaticities in the lower saturation setting had saturations that were too low to calculate a placebo filter of similar saturation and luminance with the required colour difference (communication, A. Wilkins, 02.2016), see Appendix XV.

setting of the colorimeter were applied to the background of text, a representation of the WRRT, to be viewed for comparison. The luminance of these chromaticities had a range in of 10.52–16.66 cd/m^2 and an average of 12.81 cd/m^2. Their saturations had a range of 0.048-0.114 S_{uv} , and an average of 0.081 S_{uv} (Appendix XV).

Filter choosing process

Figure 4.1 illustrates the filter choosing process that was used. Key presses were used to view, choose and fine-tune a filter which the participant thought most effectively reduced their visual stress symptoms according to Wilkins' (1994;2001) method. Although the colorimeter is designed to quantify the effective luminance and saturation of the filter in addition to the hue (Wilkins, 1994; 2001), this experiment only assessed the effect of a chosen hue and so did not vary the luminance and saturation from those originally set (Appendix XV).

Once a chosen filter had been acquired, a *placebo filter* was automatically generated by applying Wilkins et al.'s (2002) methods (communication, A. Wilkins, 02.2016): one of two chromaticities was randomly allocated; these chromaticities had been calculated to have similar saturation to the chosen filter by being equidistant to the white-point, whilst having a CIE 1976 LUV colour difference of 78 from the chosen filter (Wilkins et al., 2002) (Appendix XVI).

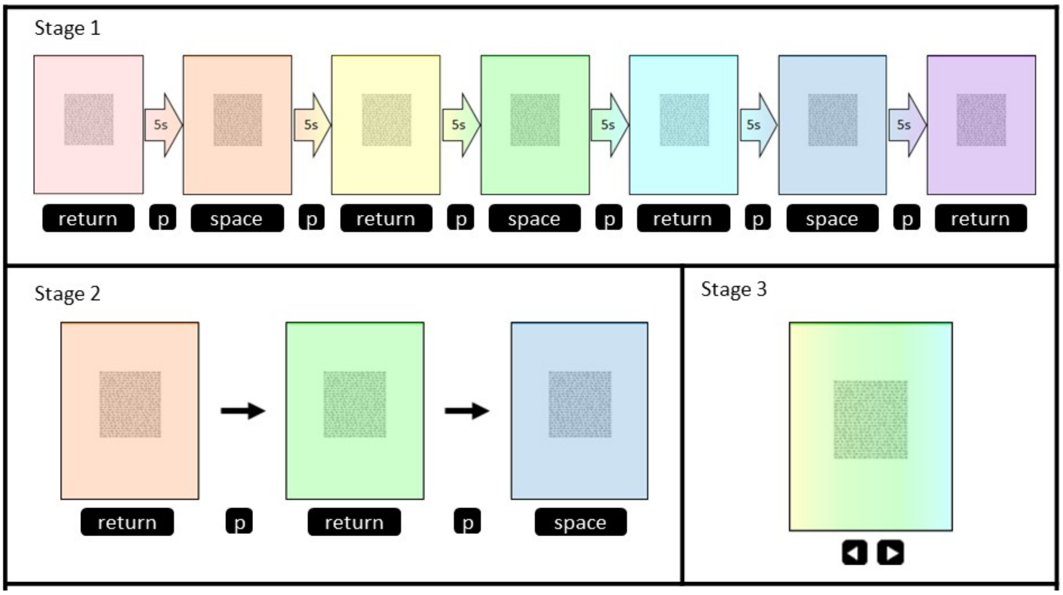


Figure 4.1. Procedure for acquiring chosen filter, including key presses. WRRT text was presented on twelve different chromaticities of filter (see Appendix XV for their chromaticities). A 'p' key-press changed the background filter. A 'return' keypress eliminated non preferred filters and a 'space' keypress saved preferred filters for later comparison. Stage 1 presented seven filters with five-second gradual changes between them. Stage 2 presented the saved filters from stage 1. Stage 3 presented the chosen filter, to be fine-tuned with right arrow and left arrow keypresses.

Change blindness tasks

To reduce the likelihood of noticing the colour change, the chosen filter was always applied at the beginning of the first task for continuity with the colour choosing process, and at the end of the second task. The task order was counterbalanced so that equal numbers of participants started each task with the chosen or placebo filter. This resulted in two orders of task and filter condition, as illustrated in Figure 4.2 and Table 4.2.

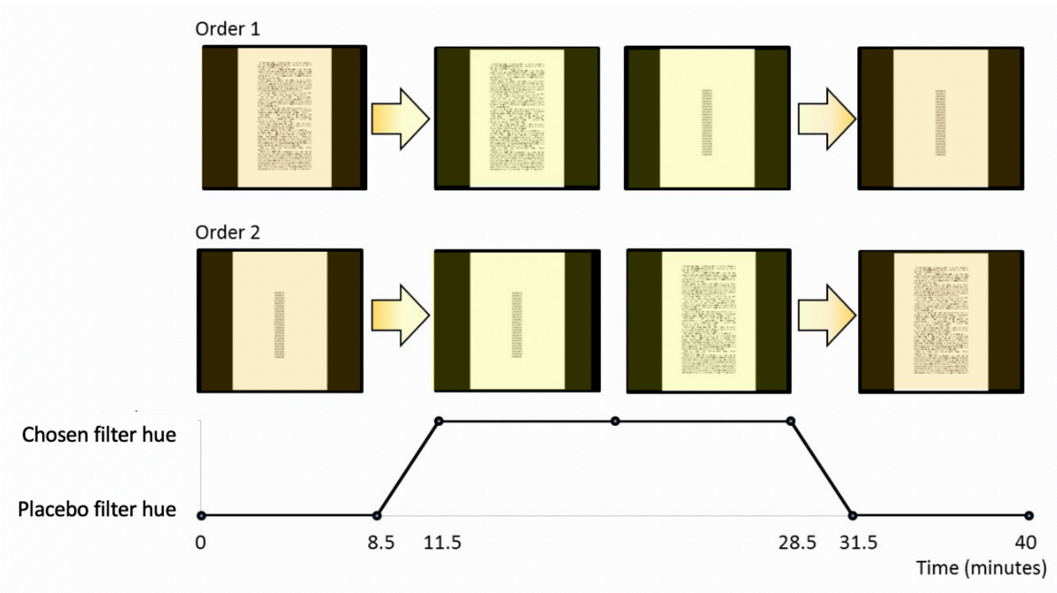


Figure 4.2. Order of tasks and conditions in the change blindness design (Experiment 4) Filter conditions always began and ended with the chosen filter. Order 1 began with the cloze task and order 2 began with the search task. The task duration of the cloze task was 8.5 minutes and the colour change was three minutes. The task duration of the search task depended upon how quickly the trials were completed; this was roughly the same.

Table 4.2

Orders of task and filter condition (Experiment 4)

	Task order		Filter condition order for cloze task		Filter condition order for search task	
Order	First	Second	First	Second	First	Second
1	Cloze	Search	Chosen	Placebo	Placebo	Chosen
2	Search	Cloze	Placebo	Chosen	Chosen	Placebo

Note: Tasks were the cloze and search task. Filters were chosen and placebo filters.

The cloze task (Figure 4.3) contained 11 randomly ordered articles to reduce the likelihood of an order effect. The articles were non-fiction, which meant that the correct choice words to complete the text were less ambiguous. Articles covered a range of topics to appeal to a wide range of interests. They were obtained from English as a Second Language assessments and so were assumed to be roughly the same level of reading difficulty.

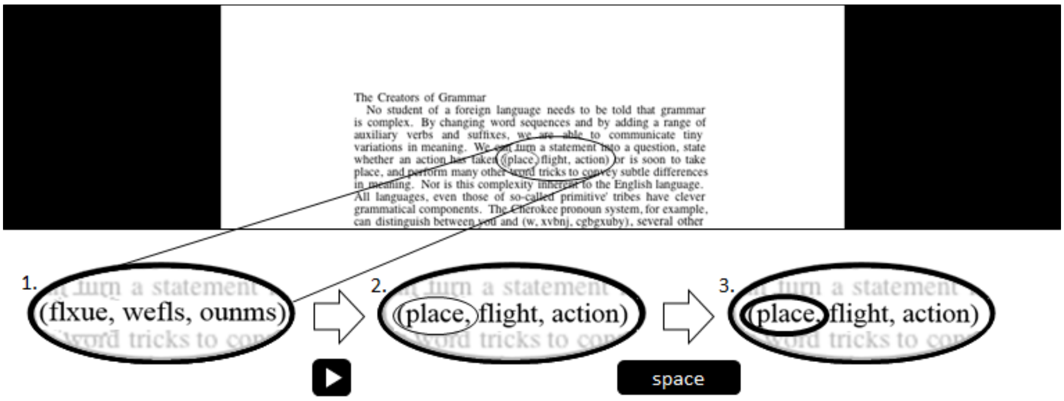


Figure 4.3. Stimulus and key presses used for one trial of cloze task in Experiment 4. This task was the identical to the cloze task used in Experiment 3 with the exception that non-fiction articles were used. Stimulus and key presses of one trial of the cloze task in Experiment 4.

The visual search task (Figure 4.4) had fifteen lines of seven capital letters, with one line containing the target letter ‘Z’. Distractor letters within one trial were either round or angular. The task was to locate the target by mouse-click. To prevent the target from immediately ‘popping out’, it was randomly located within the letter display and not on the first or last line, or on the left or right edge. The reaction time and whether the chosen location was correct was recorded. The search task was presented on the chosen and unchosen filter for 70 trials each, with 60 trials allocated to the gradually changing filter between the two conditions. The colour change procedure different slightly from the cloze task because it was designed to complete a set number of trials, rather than be time dependent. Therefore, the colour changed slightly after each mouse click.

Figure 4.4. Two trials of the visual search task. Left: Distractors were curly letters.

Right: Distractors were angular letters. The selection was made by mouse click, which circled the choice with a thick circle.

Voice activated reaction times of single word tasks

Voice activated reaction times were used to obtain reaction times of four different single word tasks: a location Stroop task; a picture Stroop task; a single word task and a phonic task. The use of voice activated reaction times was particularly important for recording spatial Stroop tasks as these have been reported to only produce interference effects when recorded verbally (Hilbert, Nakagawa, Bindl & Bühner, 2014; White, 1969).

The stimuli of these tasks are presented in Figure 4.6. The location Stroop task (Shor, 1970) presented a noun word in one of four congruent or incongruent locations. For example, if the word 'left' was presented at the left location, this would be a congruent trial, and if it was presented at any other location (bottom, right or top) it would be an incongruent trial (Figure 4.6). A trial required the word not to be read and the location ('up', 'down', 'left' or 'right') to be named. The picture Stroop task (Rosinski, Golinkoff & Kukish, 1975) presented a noun word over a congruent or incongruent picture, with the task requiring the picture to be named. Pictures were screened to be easily and instantly recognisable. The single word reading task required participants to read a noun that was displayed in front of a pattern with similar spatial properties to the Stroop stimuli. The phonic task was used to measure phonic ability (Snowling, 1996; Snowling, Stothard, & McLean, 1996). 12 non-words were individually presented and the task was to read these as quickly and accurately as possible.

Data storage was minimised by using the default quality of recording set by MATLAB; a sample rate of 8000 hertz, a depth of eight bits (eight bits per sample), and a single audio channel (“Mathworks Documentation”, 2018). Naming or word reading response latencies were recorded (in milliseconds) from the appearance of the stimulus to the onset of the voice response. Reaction times from sound waves were activated using a moving window that recorded the median amplitude every 100ms. Voice threshold was set to an amplitude range of 0.1 – 1. This range had been tested to exclude most background noise such as slamming of doors or interference from the microphone. Once the moving window threshold had been achieved, the time stamp of the onset of the word was acquired by moving the window backwards until the median amplitude was less than 0.01. The timing of the onset of the spoken words was optimized by screening the words that were to be spoken to only begin with hard consonants. Examples of time stamps reaction times are presented in Figure 4.5.

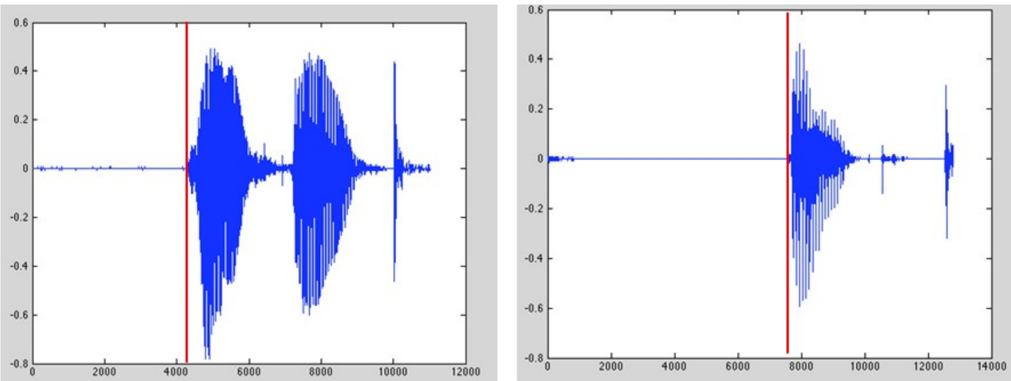


Figure 4.5. Recorded sound waves of spoken words in audio recorded task and the location of the voice activated reaction times (time stamp), (Experiment 4). Words were “Panther” (left) and “Belt” (right). Red vertical line indicates the location of the voice activated reaction time.

There were 100 trials in each task. Fifty trials were completed with and without the participants' filter. The order of filter conditions was counterbalanced so that, by the end of the experiment, an equal number of participants in each participant group had completed the first half of the tasks with and without their filter. The four tasks were completed in a counterbalanced order so that at the end of the experiment, an equal number of participants in each participant group had completed the experiment in every possible order. Each running of the experiment presented the same pictures, words and numbers of congruent or incongruent trials, but in a different order and no words were repeated within one set.

Questionnaire

Participants were given a questionnaire to record demographic information, including any diagnosis of dyslexia. For the latter, as with Experiment 3, participants were asked if they had had extra time in exams. If they answered yes, they were then asked if this was due to a diagnosis of dyslexia.

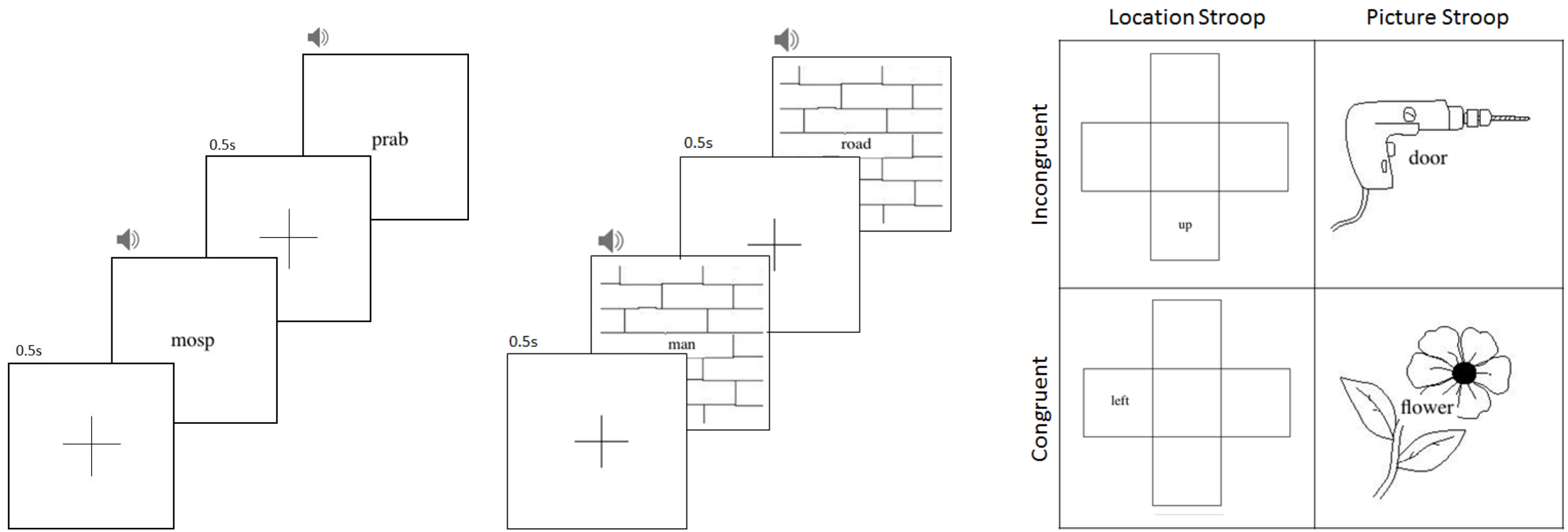


Figure 4.6. Stimuli of word tasks (Experiment 4): For each trial, a fixation point appeared on the screen for 500ms and, immediately at the offset of the fixation, the stimuli appeared in the centre of the screen and remained there until a spoken response was initiated. *Left.* Two trials of the phonic task. *Middle.* Two trials of the noun task. *Right.* The stimuli used for incongruent and congruent trials of the location and picture Stroop tasks. Words were presented in Times New Roman, point 18 and pictures were presented within a 5cm square window.

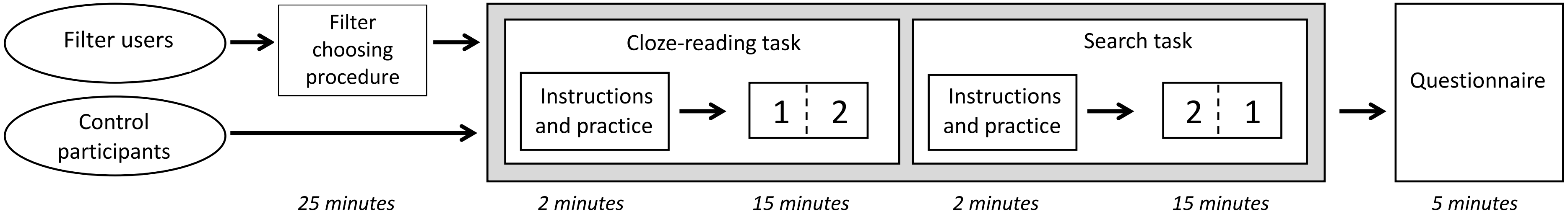
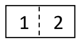
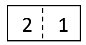


Figure 4.7. Sequence and timings of change blindness tasks (Experiment 4). Filter users obtained their chosen filter (1) by completing the filter choosing procedure prior to the tasks. Then a placebo filter (2) was allocated. Control participants were allocated the filters that their matched participant had received. The dotted line indicates a gradual transition between the filters. The grey border illustrates the counterbalancing of tasks so that the first task moved from the chosen to placebo filter , and the second task reversed this order  .Timings are in italic font.

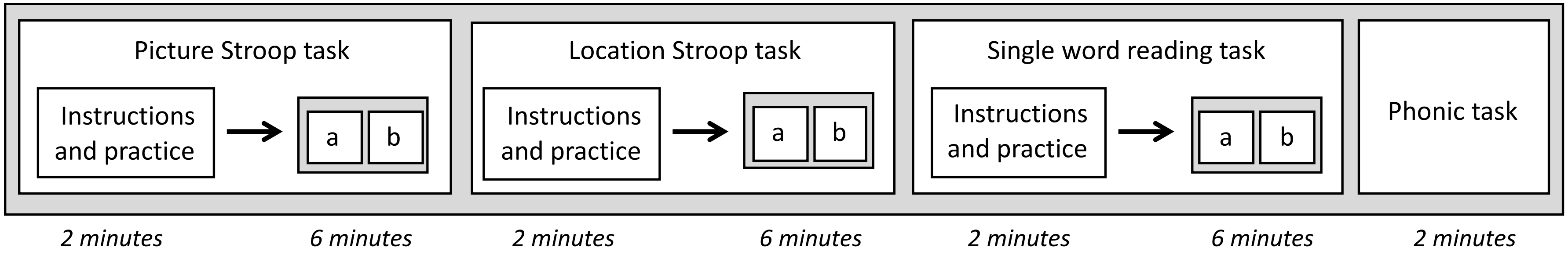


Figure 4.8. Sequence and timings of single word tasks (Experiment 4). [a] and [b] represent the filter and no filter conditions. Grey border shows counterbalancing of the tasks and filter conditions. Timings are in italic font.

4.1.2.3 Procedure

The change blindness tasks (

Figure 4.7) took place before the single word tasks (Figure 4.8). A filter was on the screen throughout the tasks to keep contrast and brightness constant and allow fair comparisons between them. filter users completed the filter choosing procedure to select their chosen filter and generate a placebo filter (Figure 4.1). The procedure took 25 minutes and was simplified from the Intuitive colorimeter procedure (Wilkins, 1997; 2001) to exclude manipulation of saturation and luminance. In stage 1, participants were instructed to inform the experimenter which filter, out of each successively presented pair, made the text appear most comfortable, clear and easiest to read; preferred filters were retained and non-preferred filters were removed. In stage 2, participants were presented with the preferred filters they had selected and asked to choose one preferred filter, if necessary, by forced choice. In stage 3, the experimenter, and then the participants fine-tuned the tint until the participant believed it most effectively reduced their visual stress symptoms.

The placebo filter condition was generated, but not displayed or referred to. The resulting filter conditions were used for each matched control participant. The chosen filters, unchosen filters and change blindness chromaticities are displayed in Appendix XVII.

The chosen filter covered the instructions of the first task. For the visual search task, participants were instructed to locate the ‘Z’ on each trial as quickly and accurately as possible by mouse-click, upon which the next trial would appear.

There were five practice trials. The procedure for the cloze reading task (Figure 4.3) followed that of Experiment 3.

Naming or word reading response latencies were recorded using a microphone headset. Instructions were to name the picture or word location, to read the single words and to read the non-words aloud as quickly and accurately as possible. Each audio task contained three practice trials in which the sound recording was checked and, if necessary, participants were asked to adjust the volume of their voice. Twelve filter users' data did not record sufficiently due to unexpected disrupted sound-recording conditions[12]. Seven of these participants completed their data sets by returning to the experiment and repeating the tasks.

To elicit whether participants noticed the colour change, they were asked if reading was easier or faster in some parts than others, and if they noticed a change in the screen at any time during the course of the experiment. Those who did not mention a colour change were recorded to have not noticed it. They then completed the questionnaire.

4.1.3 Analysis

4.1.3.1 Chosen Filters

Appendix XVIII displays the luminance and contrast measurements of chosen filters used in this experiment. Appendix XVII illustrates the distribution of chromaticities of the chosen and placebo filters in CIE (1976) and MacLeod-Boynton (1979) chromaticity diagrams. The chromaticities span the ordinate and

[12] Echoes from slamming of one of the corridor doors.

abscissa of both colour spaces. The chromaticities were well distributed in both diagrams with no clearly observable pattern or clustering. Saturations between the chosen filters of Experiment 4 were more uniform than those used in Experiment 3, as demonstrated by the absence of chromaticities around the whitepoint in the centre of the chromaticity diagrams (Appendix XVII). Appendix XXVII demonstrates that the saturation of filters in Experiment 4 spanned a shorter range than Experiment 3. Restrictions in saturation range was due to the use of the high saturation setting of the colorimeter.

The number of filter users who chose each colour, along with their average reading times, are displayed in Appendix XIX. The most popular chromaticities were mint-green and blue-purple, chosen by 37.5% of the sample. There was no observable pattern in baseline reading speeds or increased reading speeds across the chromaticities. The lack of a distinct pattern within each colour space could be due to the small sample size (n = 28). Chapter 7 continues the analysis by combining data from the three experiments involving chosen filters (Experiment 3, 4 and 6).

4.1.3.2 Change blindness tasks

Thirteen people in the experimental group and 16 people in the control group failed to notice the colour change (Table 4.3). This difference was not significant ($\chi^2(1) = .621$, $p = .431$, $\phi = 0.103$). In each participant group, since the numbers of participants who noticed and did not notice the colour change were similar, the analysis followed that of Experiment 3, and whether participants noticed the colour change (change detection) was introduced as another factor in the design. A 2x2 mixed ANOVA with the within-subjects factor of filter type (chosen filter and unchosen filter) and between-subjects factor of change detection (noticed/did not notice) was conducted for each participant group (experimental/ control) for each task.

Table 4.3.

Number of participants who noticed and did not notice the colour change in filter user and control participant groups (Experiment 4)

Participant group	Total	Noticed	Did not notice
Filter users	29	16	13
Control	29	13	16

The data collected during the transitioning colour phase approximated to the average of the chosen and unchosen filter conditions and were removed from the analysis, leaving four datasets according to filter condition (chosen and unchosen filter) and participant group (experimental and control) for each task.

4.1.3.3 Cloze task

Median reading speeds were obtained using the criteria outlined in Experiment 3 (section 3.2.3.3). There were fewer than 30% incorrect trials, where an incorrect word had been selected, per participant. These incorrect trials were equally distributed across the groups (Figure 4.9) and removed from the analysis.

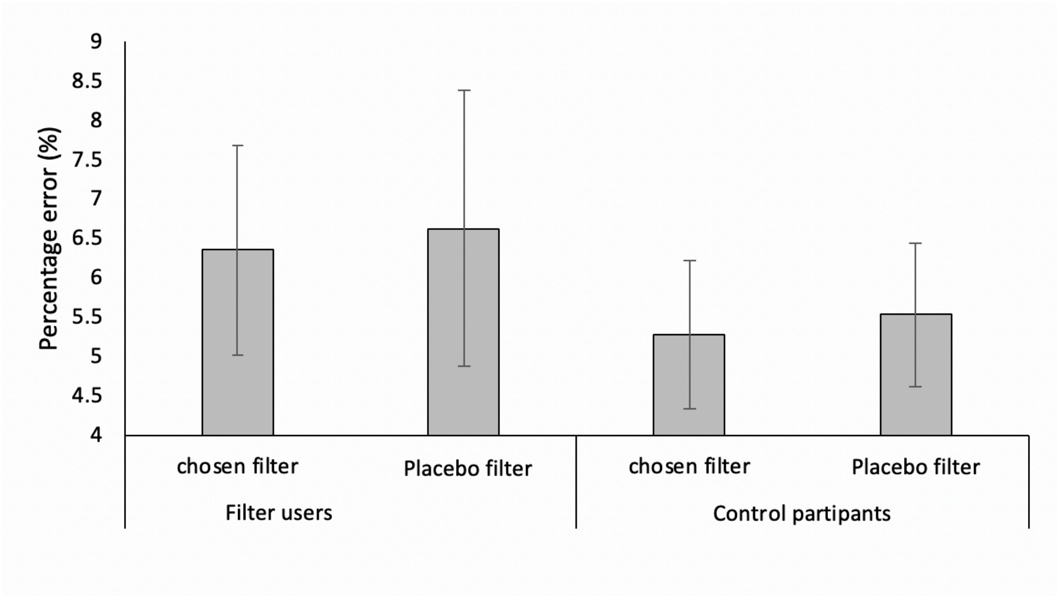


Figure 4.9. Percentage errors of cloze task across the filter conditions and participant groups (Experiment 4)

Participant 12 was classified as an outlying participants with a median reading speed of more than 2.5 standard deviations from the mean of all the participants within a luminance condition and within each participant group. This participant's data were removed. Table 4.4 displays the resulting numbers of participants in each group who noticed and did not notice the colour change.

Table 4.4.

Number of participants who noticed and did not notice the colour change in the experimental and control participant groups with (cloze) outlying participants removed (Experiment 4)

Participant group	Total	Noticed	Did not notice
Filter users	28	15	13
Control	28	12	16

The data were split into eight data sets, according to filter type, noticing the colour change and participant group (Figure 4.10). The filter using group displayed a significant interaction between the factors of change detection and filter use (F(1,26) = 5.270, p = 0.030, η_p^2= 0.162), but this was not found for the control participant group (F(1,26) = 0.134, p = 0.717, η_p^2= 0.05). To explore this interaction further, a test of simple effects was conducted. Those who noticed the colour change read significantly quicker with their chosen filter (M = 143.00, SD = 30.61) than their placebo filter (M = 127.48, SD = 32.94), F(1,26) = 7.006, p = .014, η_p^2= 0.212. yet no such effect was found for those who did not notice the colour change F(1,26) = 0.452, p = .507, η_p^2 = 0.017. Participants' average reading speeds illustrated that the control participant group (M = 153.78, S.D. = 28.84) read faster than the filter using group (M = 140.04, S.D = 28.53), but this did not reach significance t(54) = -1.792, p = .079, d = 0.479.

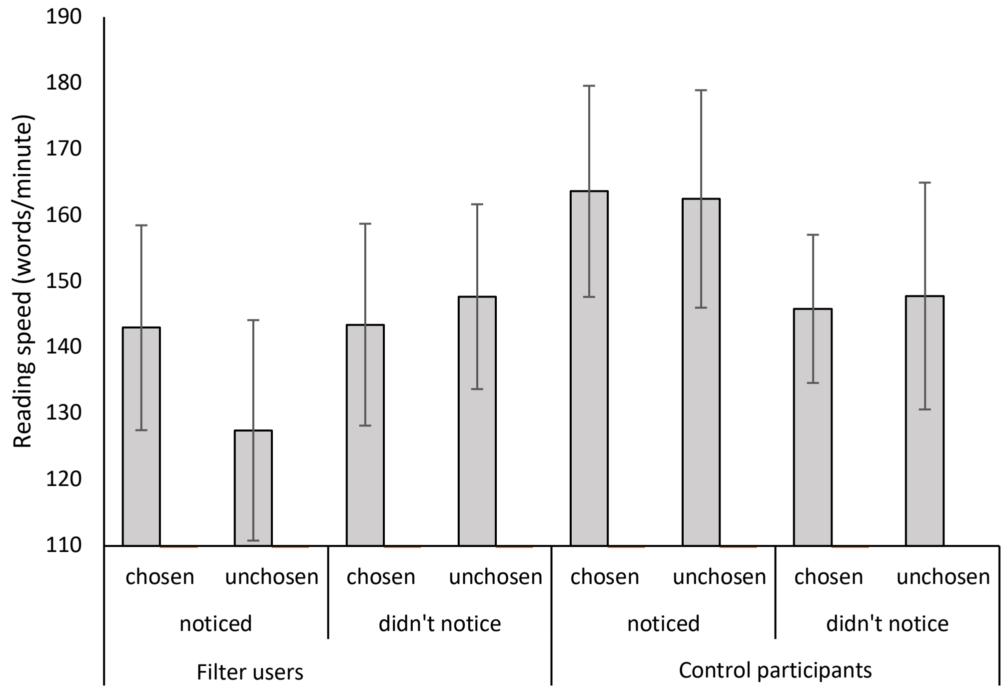


Figure 4.10. Reading speeds of cloze task categorised according to filter condition, participant group and whether participants noticed the colour change (Experiment 4). There was a significant interaction between filter type and change detection for filter users but not for normal participants.

4.1.3.4 Visual search task

Participant 9 did not complete the experiment and so the data of this participant, and the matched control participant, were removed from the analysis. All remaining participants' error rates were below 13%. Error trials were evenly distributed across the whole data set (Figure 4.11). These error trials were removed from the analysis so that analysed reaction times corresponded to trials when the 'Z' had been successfully located. The resulting numbers of participants in each group who noticed and did not notice the colour change for the visual search task analysis is presented in Table 4.5.

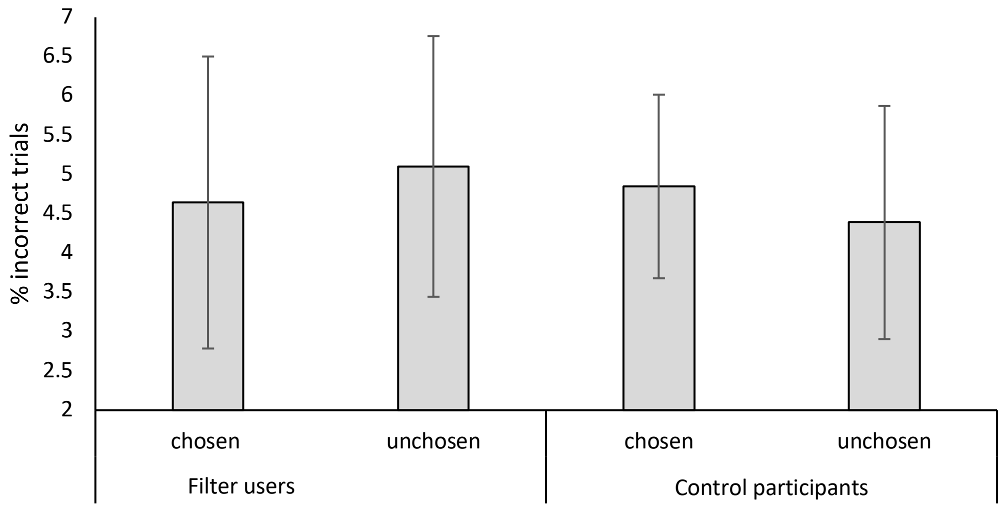


Figure 4.11. Percentage errors of search task across the filter conditions and participant groups (Experiment 4).

Table 4.5.

Number of participants who noticed and did not notice the colour change in the experimental and control participant groups with (visual search) outlying participants removed (Experiment 4)

Participant group	Total	Noticed	Did not notice
Filter users	28	15	13
Control	28	13	15

Three measures were assessed: i) the median reaction time, ii) the percentage difference between the reaction times obtained when the distractors of the search were curly and angular (percentage time difference), and iii) the number of times that the search reaction time was more than two standard deviations from the mean of each participant (number of spikes). It was assumed that the latter measure would indicate how many times the participant did not find the letter during their first search. No outlying participants were identified because all measures were less than 2.5 standard deviations from the mean of all the participants within a luminance condition and within each participant group.

For each measure, the data were split according to filter type, change detection and participant group. Each measure was assessed to contain one data set that was not normally distributed according to Shapiro-Wilk's test ($p > .05$), (Table 4.6). For all measures, there was no significant interaction found between noticing the colour change and filter-use and no main effect of filter-use in either participant group (Table 4.7; Figure 4.12.).

Table 4.6.

Visual search data sets that were assessed not to be normally distributed (Experiment 4)

Measure	Participant group	Change detection	Filter type	*p* (Shapiro-Wilk)
RT	Filter users	did not notice	Placebo filter	.005
Spikes	Filter users	did not notice	Placebo filter	.001
Difference	Control	noticed	Placebo filter	< .001

Table 4.7.

F statistics of mixed ANOVA with factors of filter use and change detection for all visual search measures and participant groups (Experiment 4)

Participant group	Measure	Factor (s)	*F*(1,26)	*p*	η_p^2
Filter users	1	Filter use*change detection	0.409	0.528	0.015
Filter users	1	Filter use	1.687	0.205	0.061
Filter users	2	Filter use*change detection	0.268	0.609	0.01
Filter users	2	Filter-use	0.366	0.55	0.014
Filter users	3	Filter use*change detection	0.014	0.906	0.001
Filter users	3	Filter use	3.971	0.057	0.132
Control	1	Filter use*change detection	0.13	0.722	0.006
Control	1	Filter-use	0.368	0.551	0.018
Control	2	Filter use*change detection	0.147	0.706	0.007
Control	2	Filter use	2.181	0.155	0.098
Control	3	Filter use*change detection	0.544	0.469	0.026
Control	3	Filter-use	0.27	0.609	0.013

Note. The measures were: 1. the median reaction time, 2. the percentage difference between the reaction times obtained when the distractors of the search were curly and angular and 3. The number of times that the search reaction time was more than two standard deviations from the mean of each participant.

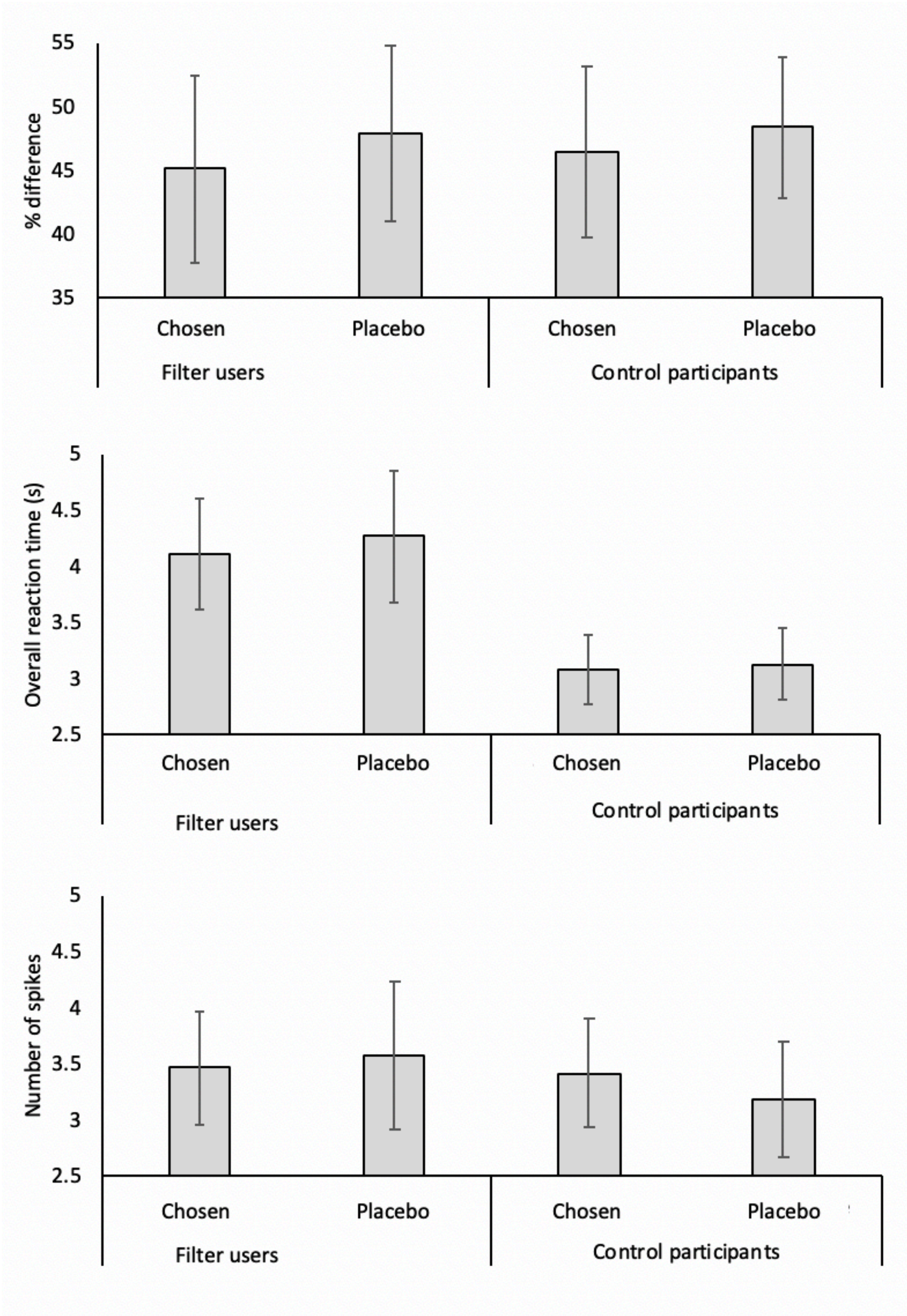


Figure 4.12. Visual search measures categorised according to filter condition (chosen and placebo) and participant group (Filter users and Control participants) (Experiment 4). Top: Reaction times. Middle: Percentage reaction time difference when distracters were curly and angular. Bottom: Number of reaction times that were above two standard deviations from the mean of each participant (number of spikes). `

Participants' average reaction times illustrated that the experimental group had significantly lower reaction times than the control group but did not differ significantly on the other measures (Table 4.8).

Table 4.8.

Independent t-test between the data of each participant group for the visual search measures (Experiment 4)

Measure	Participant group	Mean	S.D	*T* (48)	*p*(2-tailed)	*d*
1	filter users	4.190	1.377	3.413	.001	-1.011
	control participants	3.105	0.638			
2	filter users	46.430	16.267	0.217	.829	0.063
	control participants	47.366	13.502			
3	filter users	3.518	1.118	0.753	.455	-0.217
	control participants	3.295	0.921			

Note. Measures are described in Table 4.7.

4.1.3.5 Order effects

Analyses were conducted to determine if the order of tasks and filter condition affected 1. change detection 2. reading speeds or 3. search reaction times.

1. *Did order affect change detection*

Statistically more participants detected the colour change when they had completed the experiment in order 2 than order 1 ($\chi^2(1) = 4.419$, $p = .036$, ϕ = 0.069), (see Table 4.9 for participant numbers). The association between order and change detection was weak, ϕ = .276, $p = .036$, and the association was not found for either the filter users ($p = .198$) or control participants ($p = .089$) separately.

Table 4.9

Distribution of participants according to participant group, change detection (Noticed and did not notice) and task order (Experiment 4)

	Order 2(search)		Order 1(cloze)	
Participants	Noticed	Did not notice	Noticed	Did not notice
Control	8	5	6	10
Experimental	10	6	4	9
All	18	11	10	19

2. *Did order affect task measures?*

For each participant group, cloze reading speeds and search reaction times were reorganised according to the order of filter condition they had been administered. Paired *t*-tests were conducted to determine if the order of administration had affected these measures. All groups were normally distributed, as assessed by Shapiro Wilk ($p > .05$). Filter users performed better (read significantly quicker on the cloze task and had significantly faster reaction times on the search task) in the second test ($p < .05$), and this was not observed for the control participants (Table 4.10; Figure 4.13).

Table 4.10.

Order effects of cloze reading speeds and search reaction times for both participant groups (Experiment 4)

Measure	Participant group	$t(27)$	p(2-tailed)	d
Cloze reading speeds	Filter users	-2.953	0.006	0.401
	Control	-0.492	0.627	0.065
Search reaction times	Filter users	2.233	0.034	-0.300
	Control	0.785	0.441	-0.112

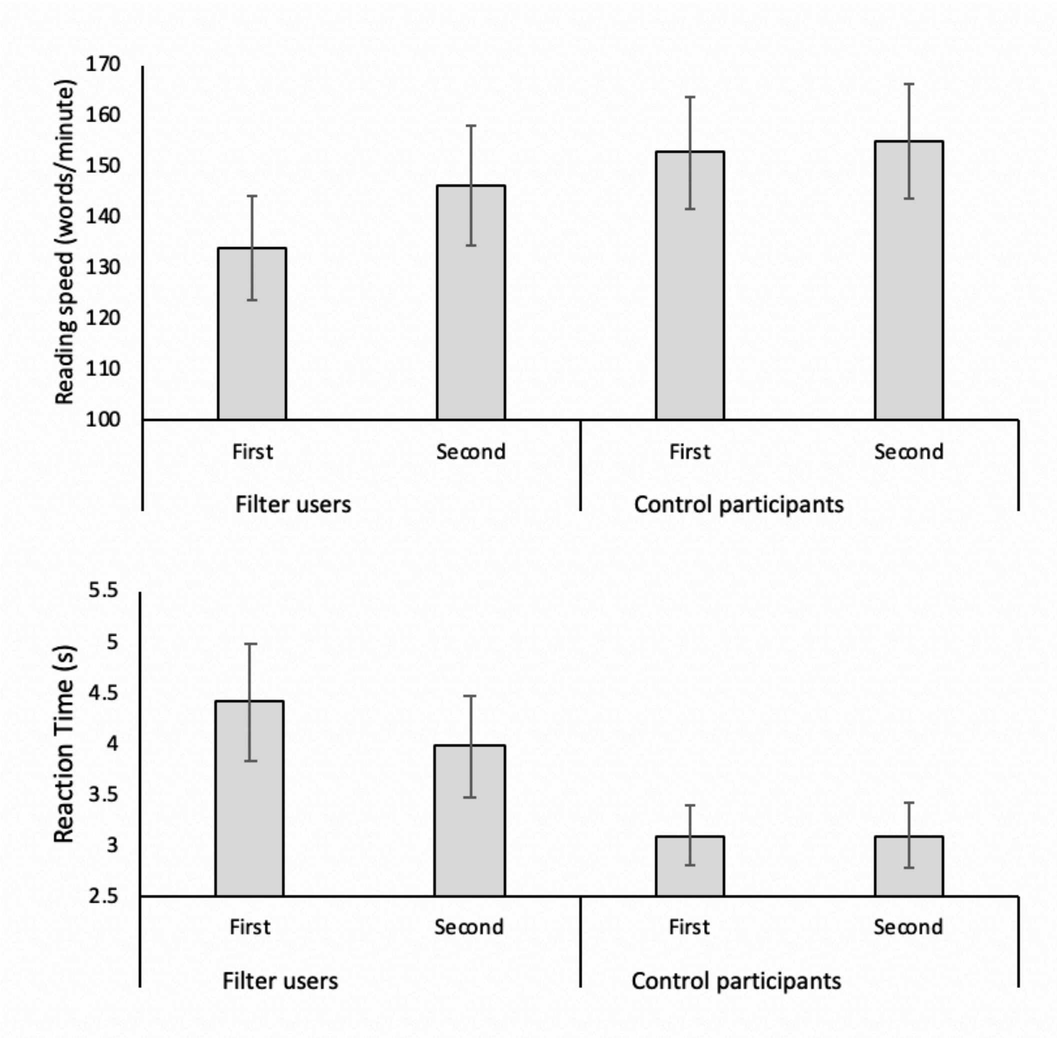


Figure 4.13. Cloze and search task performance when categorised according to participant group and the order in which the filter conditions had been completed (Experiment 4).

To investigate if this order effect of filter users affected the significant interaction that had been observed in their cloze reading speeds (section 4.1.3.3), the analysis was repeated as an ANCOVA with covariate of order. When order was included as a covariate, there was no significant interaction between filter type and change detection (p = .110, η_p^2 = .099).

4.1.3.6 Reading difficulties

Not all participants who had self-reported to be dyslexic demonstrated difficulties with phonics on the non-word task. Therefore, participants with reading difficulties were categorised into three groups: those who had difficulties with phonics, those who self-reported as dyslexic, and those who had difficulties with phonics and self-reported as dyslexic (Table 4.11). Participants who had difficulties with phonics were categorised using the non-word task data, with the threshold of a median response time of more than one second or more than one incorrectly read word.

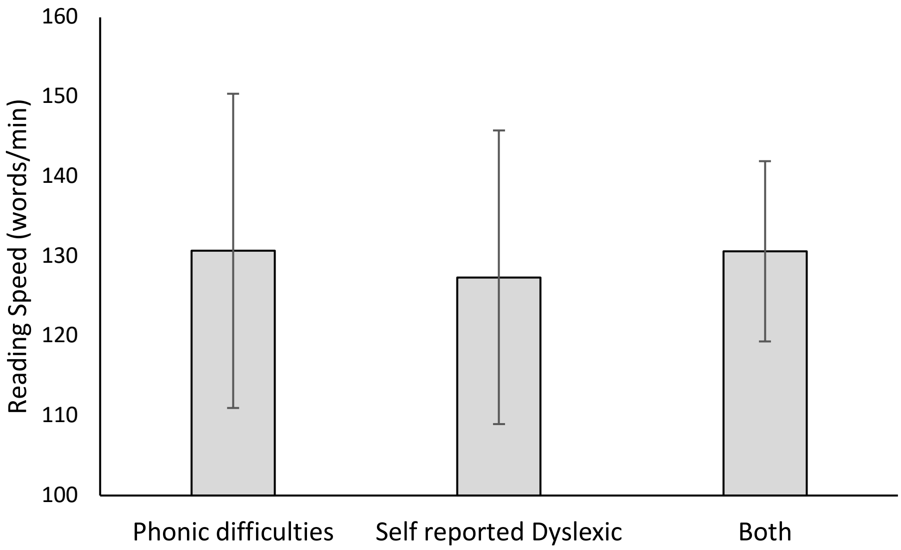
Table 4.11.

Number of participants who had reading difficulties (phonics difficulties, self-reported dyslexia and both of these) and did not have reading difficulties in each participant group (Experiment 4)

	Experimental	Control	All participants
Reading difficulties	23 (22)	4	2
• Phonic difficulties	7 (6)	3	10 (9)
• Self-reported dyslexia	7	1	8
• Both	9	0	9
No reading difficulties	6	25 (24)	31 (30)

Note. Numbers in brackets represent when outliers have been removed.

To determine if categorisation of reading difficulty had an effect on reading speeds on the cloze task, a one-way ANOVA was conducted between the participants in each of the three reading difficulty groups. There was no significant interaction between the different categories of reading difficulties, $F(2) = .047$, $p = .954$, $\eta_p^2 = .004$, (Figure 4.14). When the data of these three groups were combined to form a data set of participants with reading difficulties, these participants were found to read significantly slower (M = 129.66 , S.D. = 24.24) than participants without reading difficulties (M = 161.86 , S.D. = 24.88), $t(54)$ = 4.889, p = .727, d = 1.311 (Figure 4.14).



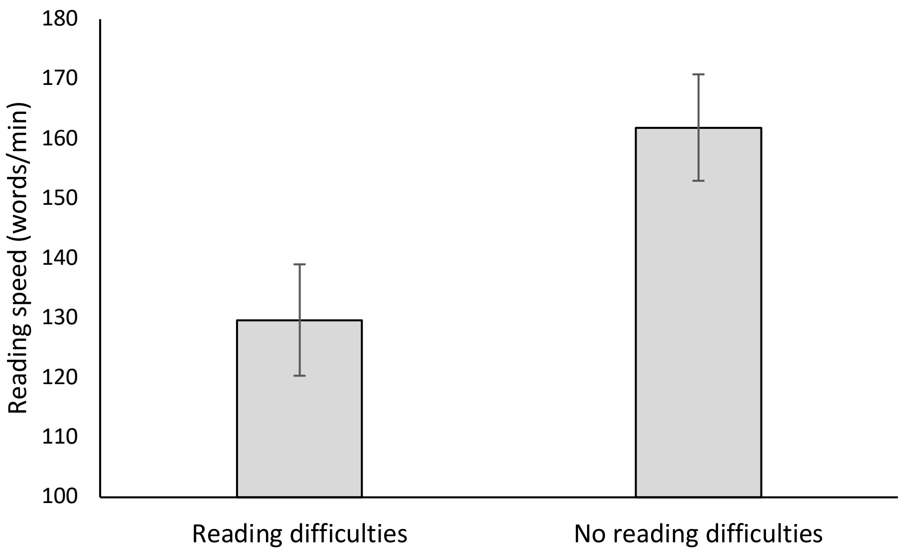


Figure 4.14. Cloze reading speeds when categorised according to type and presence of reading difficulty.

4.1.3.7 Single word tasks

There were complete data sets for 23 participants in the picture Stroop task, and 27 participants in the location Stroop and noun task. The audio files were listened to so that incorrect trials and reaction times could be identified. Incorrect trials, in which the correct word was not immediately spoken, were removed from the analysis. If possible, incorrect reaction times that had been triggered by background noise, were corrected, or otherwise removed. Removed reaction times did not exceed 10% of the trials for each participant and were evenly distributed across the conditions.

To assess if filter use had affected the picture and location Stroop interference effect differently in each participant group, median percentage interference of each luminance condition in each participant group was calculated, and a two-way mixed ANOVA was conducted for the data in each task. Percentage interference was calculated as follows:

$$\frac{I-C}{I}\ x\ 100 \qquad (7)$$

where I = Median incongruent reaction time, and C = Median congruent reaction time.

To retain statistical power and allow a comparison between two of the task, the participant outliers criteria was set to a z-score of 3. This criteria identified no outlying participants, but a criteria of 2.5 did not change the outcome of the analysis. Assumptions of a mixed ANOVA were violated (Table 4.12).

Table 4.12.

Violations of assumptions of mixed ANOVA for the three single word tasks (Experiment 4)

Task	Normality *p* (Shapiro-Wilk)	Homogeneity of variances *p* (Levene's)	Homogeneity of covariances *p* (Box's M)
Location Stroop	FU = .014		
Picture Stroop		NF = .026	< .001
Noun	C < .001	F = .012, N = .001	< .001

Note. FU = filter users, C = control participants, F = filter condition, NF = no filter condition. Unreported data represents *p* > .05.

Two data sets were not normally distributed (Table 4.12). Therefore, paired sample t-tests, which are robust against violations of normality, were conducted to assess whether there was an effect of filter use on the three measures. There was no significant difference between the filter and no filter condition of either of the Stroop interference tasks or for the noun reading task when completed by control participants (*p* > .05). However, filter users read the single words significantly more quickly than the filter users (*p* = 0.017), see Table 4.13. and Figure 4.15 for results summary.

Participants' average performance showed no significant difference between the location Stroop interference of the two participant groups. However, filter users completed the picture Stroop interference and single word reading task slower than the control participants (*p* = .006, *p* = .001), see Table 4.14 for results summary.

Table 4.13.

Paired t-tests comparing measures with and without filter for both participant groups (Experiment 4)

Measure	Participant group	Filter		No filter		Paired *t*-test			Cohen *d*
		M	SD	M	SD	df	*t*	*p*	*d*
Location Stroop interference (%)	Filter users	10.76	8.17	9.61	7.49	26	0.831	0.413	-0.147
	Control	8.58	5.79	9.14	6.43	26	0.353	0.727	0.091
Picture Stroop interference (%)	Filter users	6.87	8.26	3.94	14.58	22	0.677	0.505	-0.257
	Control	16.87	6.70	17.16	6.59	22	0.174	0.864	0.044
Single word reading reaction times (ms)	Filter users	652.13	131.55	743.26	240.64	26	2.530	0.017*	0.490
	Control	598.08	79.61	588.29	72.48	26	1.509	0.143	-0.129

Table 4.14.

Independent t-tests comparing measures of participant groups (Experiment 4)

Measure	filter users		Control participants		Independent *t*-test					Cohen *d*
	M	SD	M	SD	Levene's	Variances	df	*t*	*p*	*d*
Location Stroop interference (%)	10.19	6.97	8.86	4.53	0.125	assumed	52	0.832	0.409	-0.231
Picture Stroop interference (%)	697.69	169.84	593.19	74.24	0.001	not assumed	36	2.930	0.006*	-0.856
Word reading reaction times (ms)	5.40	5.68	17.01	5.31	0.985	assumed	44	7.165	< 0.001	2.114

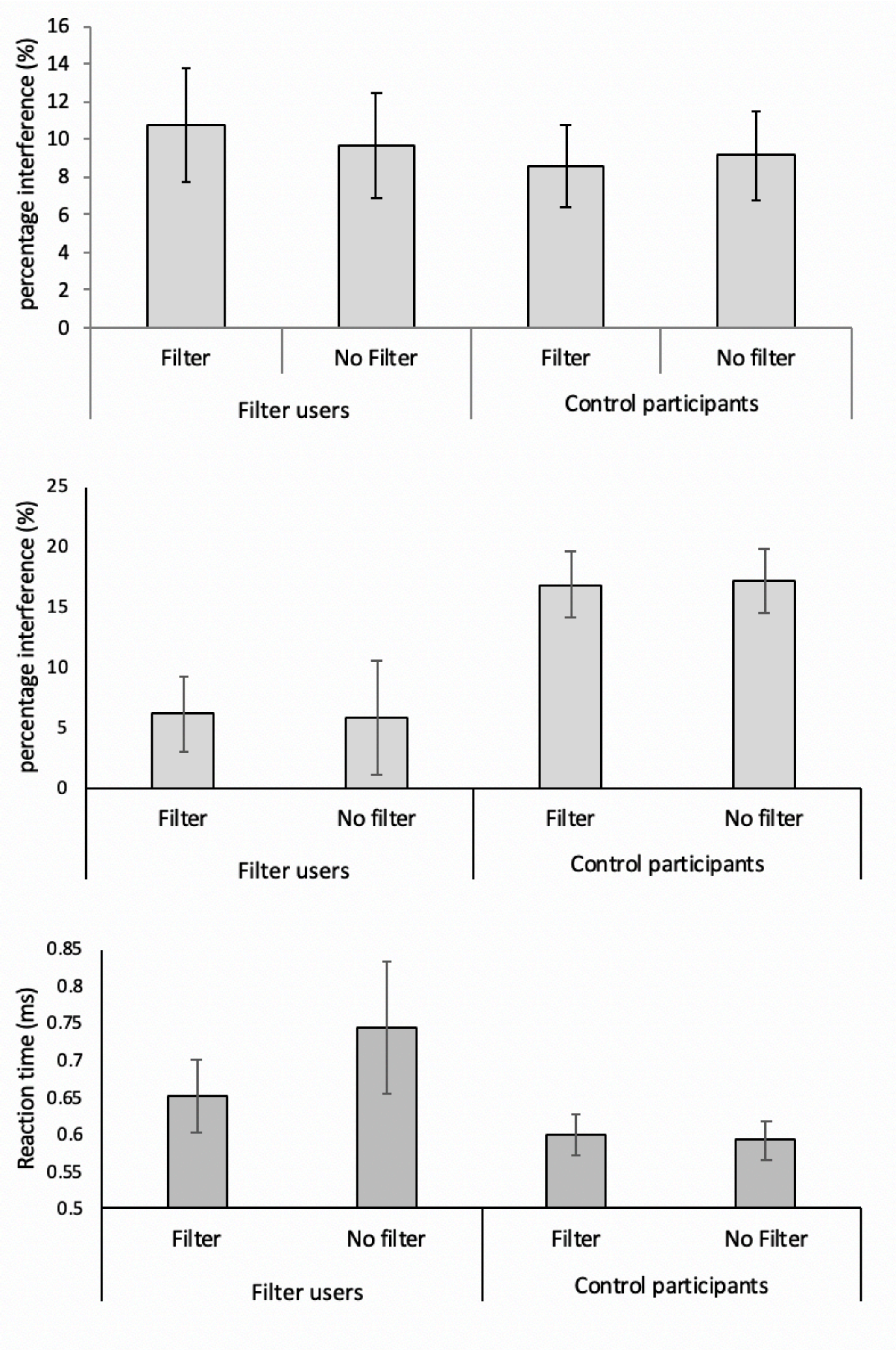


Figure 4.15..Measure of single word tasks categorised according to filter use and participant group. Top: Percentage interference of Location Stroop task. Middle: Picture Stroop task. Bottom: Reaction times of single word reading task

4.1.4 Discussion

Performance of the change blindness tasks was not affected by whether the filter was the chosen or placebo filter for either experimental or control participants. This was demonstrated by no effect of filter type on either participant group. The percentage of filter users who noticed the colour change was not statistically greater than control participants. This indicates that filter users we're not more reactive to the unchosen colour than control participants. Thirteen filter users did not notice the colour change, and so had completed the experiment in placebo controlled conditions. This group did not read significantly quicker with the chosen filter, indicating that a chosen filter did not benefit reading beyond placebo for these participants.

Seventy-nine percent of the filter users had reading difficulties (Table 4.11). Filter users are likely to have reading difficulties because filters are often discovered or administered in the quest to resolve their reading difficulties. There is a possibility that these reading difficulties prevented reading performance to be improved with the chosen filter of filter. This is an argument that has often been used by proponents of coloured filters (Irlen, 2011). However, the search task did not involve reading and demonstrated no effect of a chosen coloured filter on performance either. Fifteen filter users noticed the colour change. This participant group read significantly quicker on their chosen filter than the placebo filter, and this pattern was not observed for the matched controls. At first, this result appeared to demonstrate that noticing the colour change triggered a placebo response or that this was a group of participants with more severe visual stress symptoms. However, this was found not to be the case because when the analysis was repeated whilst controlling for order, no effect was found. It appears that this result is entirely due to an order effect experienced by the filter users that was not experienced by the control participants.

Filter users appear to have experienced a practice effect in both tasks and the control participants did not (Table 4.10). They also exhibited poorer performance on the cloze and search task. Poor performance and practice effects of the filter users is not unexpected and could be due to poorer word reading skills. This is supported by the filter users' slower reading times on the single word reading task (Section 4.1.3.7) and the higher numbers of participants with reading difficulties (Table 4.11).

The high proportion of people with reading difficulties in the experimental group (Table 4.11.) demonstrates that reading difficulties and the presence of dyslexia appears to be a confound. It is possible that, even if filter users' symptoms of visual stress had been alleviated by using their chosen filter, they did not demonstrate improved reading speeds due to their underlying reading difficulties. Since the comorbidity between visual stress and dyslexia in this sample is higher than previously reported (section 1.8.1) this may reflect that dyslexic individuals are more likely to have exposure to coloured filters during their diagnosis of dyslexia or due to their motivation to solve their reading problems.

The order in which the tasks and filter conditions had been administered appears to have affected whether participants noticed the colour change. Since the colour change was coded differently in the two tasks, a likely explanation could be that the colour change in one task was more noticeable than the other and took place in a time where participants were more likely to notice it. It could be assumed that most participants noticed the colour change when it first occurred and the order that was attributed to most change detection was when the search task had been administered before the cloze task. The colour change of the search task had been in a series of miniscule steps that were triggered by each mouse click. By comparison, the colour change of the cloze task had been time dependent and was unrelated to the participants' speed of completing the task. Therefore, perhaps more participants noticed the colour change when the search task was administered before the cloze task because the colour had depended upon their participation and was not as gradual as the cloze task.

Verbal picture and location Stroop tasks revealed no significantly increased percentage interference effect when comparing filter to no-filter condition in the either participant group. filter users demonstrated less picture word reading interference than control participants, and this appears to support the blocking hypothesis.

The single word reading times revealed significantly slower reading response times in the no filter than the filter condition for the experimental group, and this was not observed in the control group. Since no change in interference effect was observed in the Stroop tasks with their filter, this result is probably a placebo effect.

The change blindness tasks in this experiment demonstrated that a chosen coloured filter does not improve reading or search performance for people who use chosen coloured filters for reading. The single word tasks demonstrate that, although the filter users displayed poorer performance on some measures and demonstrated an improvement in single word reading in an uncontrolled task, there was no observed benefit of their filter when the task was controlled using Stroop performance measures. These results raise questions about the efficacy of filters and their recommendation as a remedy for visual stress.

4.2 Experiment 5

4.2.1 Introduction

Experiment 5 repeated the cloze task of Experiment 4 for filter users to clarify whether the interaction observed in the cloze task was due to a placebo effect or visual stress. It was designed and conducted prior to recognising that there was no significant interaction when the order effect was controlled (section 4.1.3.5), implications of which are incorporated in the discussion.

Filter users who had noticed (group 1) and had not noticed (group 2) the colour change in Experiment 4 would be made aware of the colour change prior to repeating the cloze task of Experiment 4. If group 1 were to read quicker on their chosen than placebo filter and this was not observed for group 2, then this suggests that the effect observed in Experiment 4 was not a placebo response to the awareness of the colour change but a stable effect of the chosen filter on these participants. This might be the result of group 1 having more symptoms of visual stress 13. However, if participants who did not notice the colour change in Experiment 4 changed their behaviour in Experiment 5 by reading quicker in their chosen filter condition, this could indicate that the results of Experiment 4 were

[13] It would be unlikely that such a result would be due to more effective hue choices of group 1 since the participants had undergone the same colour choosing procedure.

dependent upon the awareness of the colour change and therefore a placebo effect.

4.2.2 Method

4.2.2.1 Participants

20 participants (15 female) were recruited to complete the cloze experiment in exchange for payment. Twelve of these participants reported that they had dyslexia. Eleven participants used Irlen filters, eight participants used Intuitive filters and one participant used an unknown coloured filter obtained from a disability service.

11 participants had noticed the colour change in Experiment 4 (group 1) and nine participants had not noticed the colour change (group 2). The remaining eight filter users from Experiment 4 were unsuccessfully recruited; these participants completed the visual stress symptomatic questionnaire remotely.

4.2.2.2 Stimuli

Cloze task

The cloze task of Experiment 4 was used with the same order of filter conditions as in Experiment 4. Articles were new and presented in a random order for each participant.

Visual stress (VS) symptoms questionnaires

To determine if group 1 and group 2 differed in the number of visual stress symptoms they experienced, two visual stress questionnaires (Appendix IV) were administered. Existing questionnaires involved items relating to pattern-observation tasks (Conlon, Lovegrove, Chekaluk & Pattison, 1999), could be considered to be leading (Henderson et al., 2013) or too brief (Wilkins, 2003). Therefore, two questionnaires were created to elicit the number and type of visual stress symptoms experienced when reading. The first was designed to be quick to administer and was based upon Wilkins' (2003) measure. The second contained pseudorandomised positive and negative sentences in relation to visual stress symptoms (Henderson et al., 2013; Conlon, et al., 1999).

4.2.2.3 Procedure

Cloze task

Figure 4.16 displays the sequence and timings of Experiment 5. Participants were instructed to complete the same reading task that they had previously completed. They were reminded of the order in which the conditions would be administered and that the colour would gradually change after 8.5 minutes of reading. Participants completed four practice trials to remind them of the task. They then completed the experiment in silence.

The 20 participants who were able to return for the repeat cloze experiment completed the VS questionnaires in the lab using an online survey (Appendix IV). The remaining eight participants were sent a link to the online survey to complete remotely. Visual stress measure 1 and 2 were conducted in counterbalanced order so that an equal number of participants completed the measures in each order.

To gain some insight into the participants' perception of the effectiveness of their filters, participants who returned to complete the cloze task were asked what other colours improved their reading; see section 4.2.3.3 for their responses.

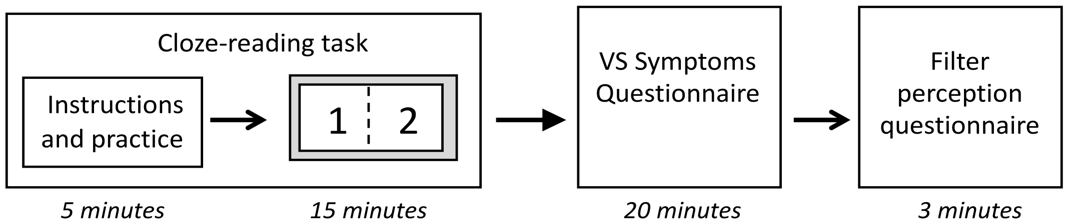


Figure 4.16. Sequence and timings of Experiment 5. Background filters 1 (chosen) and 2 (unchosen) were replicated from Experiment 4. The grey border represents the counterbalancing of the order of chosen and placebo filters. The dotted line indicates a gradual transition between the chromaticities of these filters. Timings are in italic font.

4.2.3 Analysis

4.2.3.1 Cloze task

The data collected during the transitioning colour phase were removed from the analysis. Median reading speeds were obtained using the criteria outlined in section 3.2.3.3. There were fewer than 20% incorrect trials per participant, and these were equally distributed across the groups and removed from the analysis.

A 2x2 mixed ANOVA with the within-subjects factor of filter type (chosen and placebo) and between-subjects factor of whether they had noticed the colour change in Experiment 4 (group 1 and 2) was conducted. There was no interaction between filter type and change detection, or effect of filter type on reading speed (p > .05), (Table 3.11). There was no effect of order of filter conditions, t(19) = -0.679, p = .505, d = 0.112.

The interaction of Experiment 4 was still present with this reduced number of participants. Group 1 had read significantly quicker with the chosen filter (M = 149.43, SD = 27.65) than the placebo filter (M = 130.08, SD = 32.83), (p = .025, 0.411), and Group 2 had not read significantly quicker with the chosen filter (M = 140.34, SD = 21.1) than when it had (M = 151.29, SD = 19.48), *(p* = .124, 0.27), (Table 4.15.; Figure 4.17). The order effect was also still present t(19) = 2.327, p = .031, d = .451.

Table 4.15.

F statistics of mixed ANOVA with factors of filter type and change detection when participants of Experiment 5 were tested in Experiment 4 and 5

Data	Factor (s)	F(1,18)	p	η_p^2
Experiment 4	Filter type*change detection	9.280	0.007	0.340
Experiment 4	Filter type	0.711	0.410	0.038
Experiment 5	Filter type*change detection	0.476	0.499	0.026
Experiment 5	Filter type	1.135	0.301	0.059

Note. Eleven people who did not notice the colour change and nine people who noticed the colour change returned for the experiment.

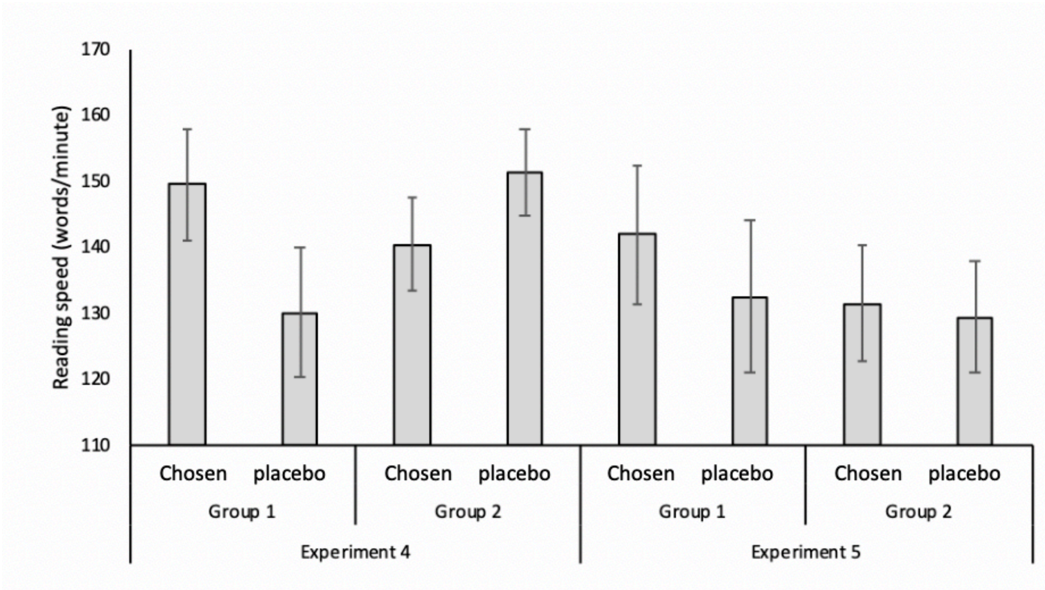


Figure 4.17. Reading speeds of Experiment 4 and 5 using only Experiment 5's participants' data and categorised according to whether a chosen or placebo filter was used and whether they had noticed the colour change in Experiment 4. Group 1 were participants who had noticed the colour change and group 2 were participants who had not noticed the colour change.

4.2.3.2 Visual stress questionnaire

Scores from the two questionnaires largely described different symptoms, with the exception of two items concerning difficulty of reading over time and symptoms of blur (Appendix IV). Symptoms were totalled, with one item removed when it had been repeatedly ticked across the two questionnaires. There were no circumstances in which participants had ticked one of these items without ticking the item that corresponded to it in the other questionnaire. All participants reported more than four symptoms of visual stress. There was no significant difference between the number of symptoms that Participants who noticed (M = 8, SD = 1.89) and did not notice the colour change (M = 7.92, SD = 2.019) experienced, $t(26) = .104$, $p = .918$, $d = 0.039$.

4.2.3.3 Views on colour

Twelve out of 20 participants said that their preferred colour was not the only colour that helped their reading speed. These participants were evenly distributed across the groups with seven participants in the group that noticed the colour change, and five in the group that did not notice the colour change. When asked what other colours improve their reading, participants responses were:

- ‘blue, purple, yellow glasses- make everything less intense. It may not be the colour.
- grey and yellow’
- ‘blue-grey, anything in the blue/green/grey spectrum’
- ‘There were other colours identified from the ophthalmologist at the time. Sometime just reducing the contrast so using grey helps when I do not have my lenses.’
- ‘light yellow’
- ‘pink/reds can help a little’
- ‘Yellow’
- ‘all colours seem to improve compared to white - I use Flux on my laptop’
- ‘most other colours, white background and black text is the worst. blue is good too’

- ‘Orange is what I usually use as it is easy to mind on computers’
- ‘Anything of a similar shade, particularly when coloured: yellow, pink, green, blue are all good, probably all fine.’
- ‘line spacing’.

Nine out of 20 of the participants had tried reading on grey or a reduced contrast setting rather than colour. One of these participants reported that this had been more beneficial than colour, three participants reported that it had been worse than colour, and the last five of participants reported that it had the same effect.

4.2.4 Discussion

Participants’ knowledge of which condition was their chosen filter did not affect their reading performance. This suggests that there was no placebo effect in relation to the expectancy of the chosen filter on reading performance. There was no order effect which implies that the order effect that was reported in Experiment 4 was a practice effect of the task. All participants reported four or more symptoms of visual stress, which suggests that people who use chosen coloured filters to read experience visual stress symptoms.

Surprisingly, 12 out of 20 of these participants reported that they did not believe that one colour was necessary to reduce their visual stress. Ten of these participants mentioned other colours, and two referred to factors involving reducing the contrast.

Chapter 5: The role of placebo in the diagnosis and treatment of visual stress with chosen coloured filters (Experiment 6)

5.1 Introduction

Collectively, the literature appears to suggest that WRRT reading speed increases were reported when chosen coloured filters were administered with leading instructions regarding their effectiveness (Table 5.1). With no blinding, this leaves the possibility that increased WRRT reading speeds in these studies were a placebo response to a chosen treatment. There is no convincing evidence to support the claim that WRRT reading speed increases directly correlate with measures of visual stress severity or reading benefit (section 1.7.3). Blinded studies indicate that a reduced text contrast, not chosen filter, may reduce visual stress and benefit reading (section 1.12).

To investigate if WRRT reading speed increases associated with chosen coloured filters are a placebo effect, either leading or non-leading instructions were given prior to completing the WRRT with and without a chosen or assigned colour (yellow), or grey filter (see Figure 5.1). A questionnaire measure of visual stress severity was also developed and incorporated to identify participants who had visual stress.

Table 5.1.

Reading speed increase according to filter-type and instruction-type in the literature

Filter type	Instruction type	Papers	Increased reading speeds
Assigned colour	Neutral	Jeanes et al,1997; Wilkins and Lewis, 1999; Wilkins et al, 2001; Wilkins et al, 2005	No
Assigned colour	Leading	Bouldoukian et al., 2002	No
Assigned grey	Neutral	Jeanes et al, 1997; Wilkins and Lewis, 1999	No
Assigned grey	Leading	Wilkins and Lewis, 1999	No
Chosen colour	Neutral	Ritchie et al., 2011; Ritchie et al., 2012	No
Chosen colour	Leading	Jeanes et al., 1997; Wilkins and Lewis, 1999; Wilkins et al., 2001; Bouldoukian et al., 2002, Wilkins et al, 2005	Yes

Note: These studies are described in section 1.9.1.3 and 1.9.2.1.

5.2 Method

5.2.1 Participants

120 native speaking undergraduates (18-31 years old, 96 female) with normal (79) or corrected to normal (41) vision were recruited in exchange for course credits. No participants identified themselves to be colour blind or had previous exposure to coloured overlays for reading. Six participants were self-diagnosed dyslexic.

5.2.2 Stimuli and Procedure

See section 2.4.2 for detail on how the baseline text and background luminance was measured. The effect of filters was mimicked by modifying the chromaticity of the text background to match the effect of Intuitive overlays (Wilkins, pers comm). These types of filter are designed to acquire the chosen colour in a systematic manner because they have approximately equal luminance and saturation, and approximately evenly distributed hues (Wilkins, 1994). The twelve chromaticities used for the chosen filter condition including the filter used for the yellow filter condition.

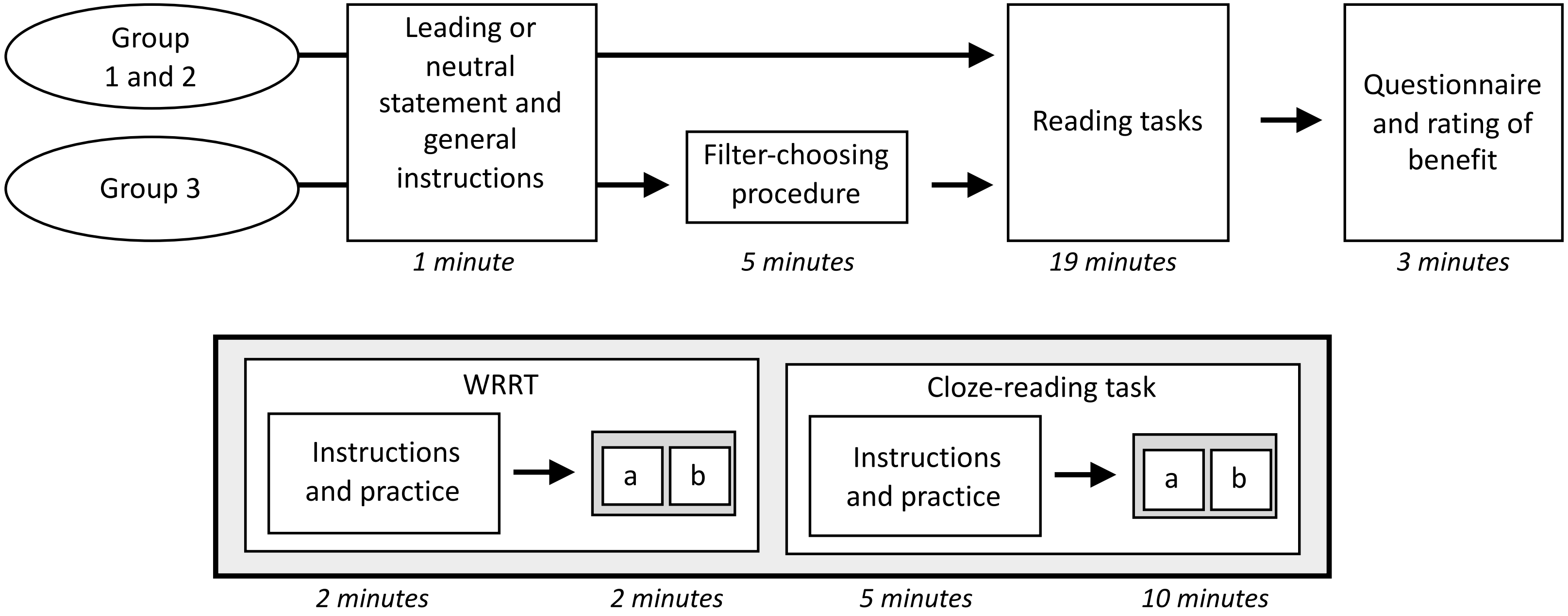


Figure 5.1. Sequence and timings of Experiment 6. Top: Group 1, 2 and 3 refers to participants who were allocated to receiving a yellow, grey or chosen filter, respectively. All groups were given leading or neutral sentences and general instructions regarding the nature of the tasks and order of completion. Group 3 completed the filter-choosing procedure. Bottom: Reading tasks were the Wilkins Rate of Reading test (WRRT) and cloze-reading task, conducted with (a) and without (b) the allocated filter. Grey border indicates counterbalancing of tasks and use of filter. Timings are in italic font.

5.2.2.1 Leading and non-leading instructions

Leading or non-leading instructions for each condition were in the form of a leading or neutral sentence (Table 5.2) presented in pt. size 16 font.

Table 5.2.

Leading or neutral sentences in yellow, grey or chosen filter conditions (Experiment 6)

Condition	Sentence
Leading/yellow	'A yellow background has been found to help reading'
Non-leading/yellow	'Sometimes text is displayed on a yellow background'
Leading/grey	'A grey background has been found to help reading'
Non-leading/grey	''Sometimes text is displayed on a grey background'
Leading/chosen	'A chosen coloured background has been found to help reading'
Non-Leading/chosen	'Sometimes text is displayed on a coloured background'

5.2.2.2 Filter choosing procedure

The filter choosing procedure was a simplified version of the Intuitive overlay procedure (Wilkins, 1994). Participants were asked to compare the appearance of text with two randomly chosen filter backgrounds (Figure 5.2). Preferred colour options were retained for later comparison and non-preferred colour options were eliminated. If there was no preference, both colours were retained. The colour choosing process was repeated for the retained colours and participants were ultimately told that they must choose only one preferred colour, which would be allocated as their chosen filter.

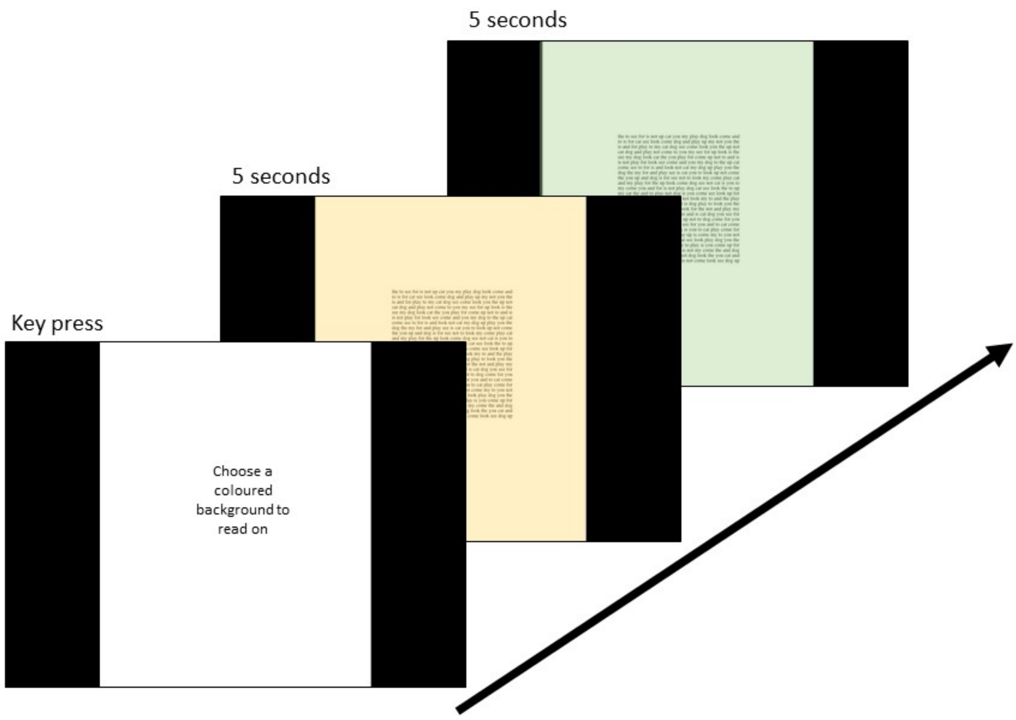


Figure 5.2. Filter choosing procedure for Experiment 6: Participants were asked to choose a coloured background to read on by repeatedly comparing the appearance of the WRRT on two coloured backgrounds for five seconds each.

5.2.2.3 The cloze task: a measure of reading benefit

The cloze task (Figure 4.3) was adapted to continue for five minutes before timing out.

5.2.2.4 The WRRT

The WRRT (Appendix VII) was presented with the spatial properties that were stipulated by Wilkins et al. (1996). The order of words in each line was pseudorandomised within each set of 20 participants, and within the testing of each participant.

5.2.2.5 Questionnaire

The questionnaire is described in section 4.2.2.2 and shown in Appendix IV.

5.2.2.6 Rating of benefit

The rating of benefit stimulus was a five-point Likert scale presented on a sheet of card (Figure 5.3).

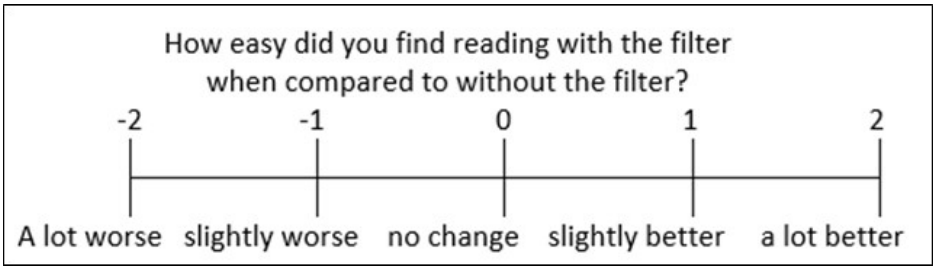


Figure 5.3. Stimuli used to rate the level of perceived benefit of filter (Experiment 6)

5.3 Procedure

It was verbally clarified that one section would involve audio recording for marking accuracy purposes, after which data would be deleted.

The sequence of the experiment is illustrated in Figure 5.1. Twenty participants were randomly assigned to one of the six conditions characterised by instruction type (leading or non-leading) and filter type (yellow, grey or chosen). Participants viewed the screen containing a leading or neutral statement regarding the effect of the allocated filter (Table 5.2) and instructions about the experiment (Appendix VI). The sentence was read aloud by the experimenter to ensure that it was understood.

Participants in the chosen filter condition completed the filter choosing procedure (section 5.2.2.2). All participants completed the cloze task (section 5.2.2.3) and WRRT (section 5.2.2.4), with and without their allocated filter. The order of tasks and use of filter was counterbalanced within each set.

The WRRT was conducted according to Wilkins' guidelines (Wilkins, 1997; 2001; 2002; 2003). Participants were instructed to read the text aloud as quickly as possible, whilst retaining accuracy, for one minute. One practice trial was followed by the two audio-taped trials, conducted with and without the allocated filter and lasting one minute before timing out.

In the cloze task, participants completed four practice trials, amounting to half a page of text. They then independently and silently completed the task with and without their allocated filter. The task was continued until it timed out after five minutes.

Participants completed the questionnaires in silence. To complete the rating of benefit measure (Figure 5.3), participants were asked to point at which number most accurately represented how easily they found reading with the filter when compared to without the filter.

5.4 Analysis

5.4.1 Chosen filters

Appendix XXI displays the luminance and contrast measurements of chosen filters used in this experiment; Appendix XX illustrates their chromaticities of the chosen and placebo filters in CIE (1976) and MacLeod-Boynton (1979) chromaticity diagrams. The chromaticities were well distributed across the 12 available filter choices in both diagrams. Appendix XXVII demonstrates that the saturation of filters used in Experiment 6 were lower than those used in Experiment 4 due to the use of the low, rather than high saturation setting of the colorimeter.

The number of participants who chose each colour, along with their average reading times, are displayed in Appendix XXII. The most popular chromaticities were aqua and blue, chosen by 37.5% of the sample. There was no observable pattern in baseline reading speeds or increased reading speeds across the chromaticities. The lack of a distinct pattern within each colour space could be due to the restricted number of colour choices. Chapter 7 continues the analysis by combining data from the three experiments involving chosen filters (Experiment 3, 4 and 6).

5.4.2 Calculation of reading speeds

WRRT reading speeds were calculated by totalling the number of accurately read words in one minute from the audio recordings. When a line was missed, one word was deducted from this total. Cloze reading speeds were calculated using times corresponding to correct word options. The time taken to select each word option was included because text may have been re-read or understood during this time. However, this inclusion in the median reading speed made no difference to the outcome of the analysis. All participants' accuracy was more than 70% in all conditions.

Each participants' median reading speed was calculated with and without the filter for all conditions. The data were split by task (cloze, WRRT), filter use (with filter, without filter), filter type (chosen, yellow, grey) and instruction type (leading, non-leading). Reading speeds more than 2.5 standard deviations from the mean of each of these data sets were labelled as outliers. Five participants were identified to have outlying reading speeds. Their data were evenly distributed across all conditions and so were removed from the analysis.

5.4.3 Order effects

To check for order effects, independent *t*-tests were used to compare reading speeds when each task was administered first or second, and paired sample *t*-tests to compare reading speeds obtained from the first and second administrations of each reading task. There was no effect of task order on the cloze task, or significant difference between first and second administrations within each task ($p > .05$), (Table 5.3) When the WRRT was conducted before the cloze task, WRRT reading speeds were significantly faster than when it was conducted after the cloze task ($p = .014$), (Table 5.4).

Table 5.3.

Paired samples t-test comparing the reading speeds obtained from the first and second administration within each task (Experiment 6)

Task	Administration	Mean	S.D.	*t*(114)	p(2-tailed)	*d*
WRRT	First	163.03	25.96	-1.82	.071	-0.065
Second		164.70	25.70			
Cloze	First	148.29	31.82	-0.367	.714	-0.032
	second	149.30	31.78			

Table 5.4.

Independent samples t-test comparing task reading speeds of both task orders (Experiment 6)

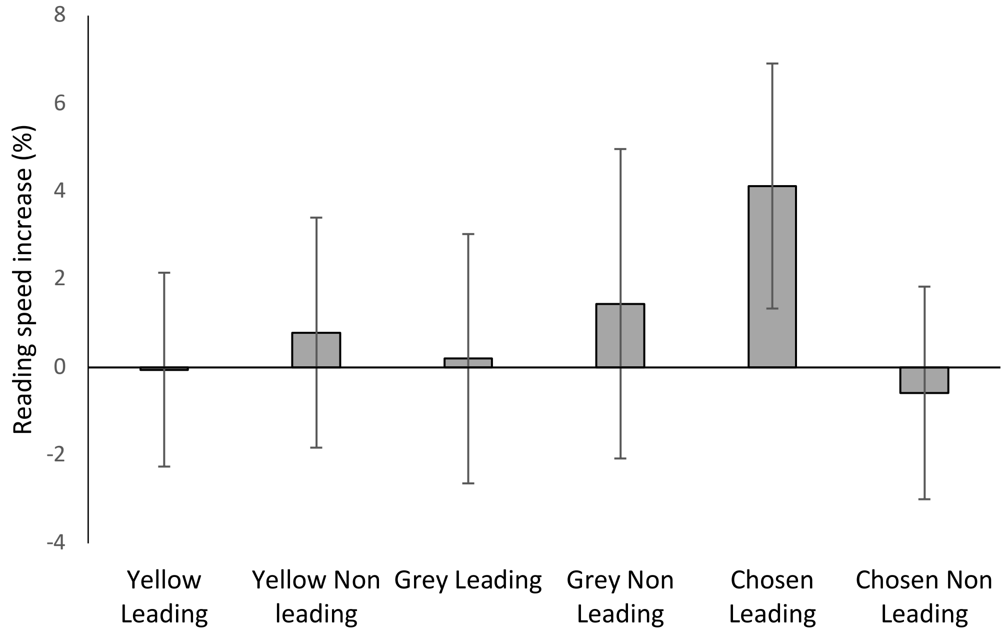
Task	Task order	Mean	S.D.	n	*t*(113)	p(2-tailed)	*d*
WRRT	1	169.81	28.37	56	2.506	.014	0.131
WRRT	2	158.22	20.85	59			
Cloze	2	148.14	29.29	56	-0.241	.810	0.034
Cloze	1	149.41	27.31	59			

Note. 1 = WRRT conducted first. 2. = cloze task conducted first.

5.4.4 Are WRRT reading speed increases with a chosen filter a placebo effect?

Percentage reading speed improvements made with a filter were calculated for each reading task according to the filter and instruction type, and one sample *t*-tests were used to determine whether any of the conditions in each reading task had a significant reading speed increase. Only the chosen filter/leading condition in the WRRT, not cloze task, had a reading speed increase significantly above zero ($p < .01$, 1-tailed). An independent *t*-test revealed that reading speeds were significantly increased when leading instructions were given (M = 4.13, S.D. = 6.36) compared to when neutral instructions were given (M = -0.58, S.D. = 5.52), $t(36) = 2.433$, $p = .01$, $d = .813$ (one-tailed, see Figure 5.4).

a.



b.

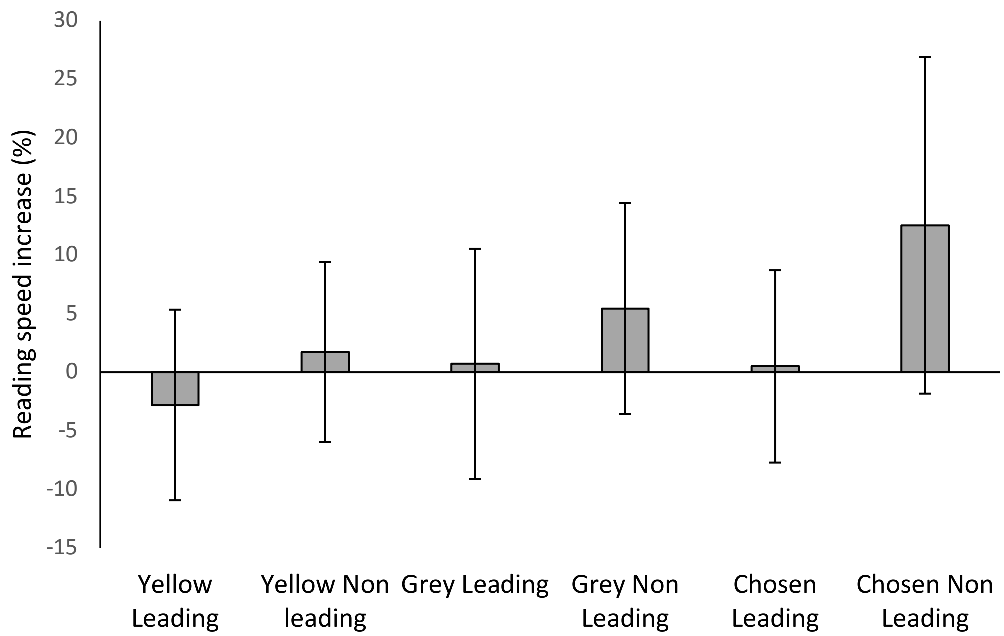


Figure 5.4. Reading speed increases when different filter types (yellow, grey or a chosen colour) and instruction (leading or non-leading) were used with the Wilkins Rate of Reading Test (Top) and the cloze task (bottom).

5.4.5 Measuring visual stress severity

To see if just one measure of visual stress could be used to determine the level of symptoms, the two sets of scores were correlated. As the two measures were not normally distributed (Shapiro-Wilk, $p < .001$), this calculation was based on Spearman's rho. The scores were highly correlated ($r = .627$, $n = 116$, $p < .001$), but this correlation only explains 39% of the variance. It therefore appeared that one measure could not be used to determine the level of symptoms. The two questionnaires largely described different symptoms, with the exception of two items (Appendix V). It was not appropriate to remove these items in one questionnaire since participants had sometimes responded differently on the two questionnaires. Therefore, an average of both questionnaires' normalised scores was calculated to represent a measure of visual stress (VS score) for each participant.

5.4.6 Correlations

Correlations are presented in Table 5.5. Increased WRRT reading speeds did not significantly correlate with the VS score. Although WRRT and cloze reading speeds did weakly correlate, reading speed increases on the two tasks were not found to correlate. Ratings of benefit were correlated with reading speed increases of both the cloze task and WRRT. Increased reading speeds on the cloze task with any filter type were found to be significantly correlated with VS scores, and also inversely correlated with cloze reading speed.

Table 5.5.

Bivariate correlations between perception scores, visual stress scores, reading improvement and reading speed on a white background, measured by the WRRT and cloze task. (Experiment 6)

	Cloze increase	WRRT increase	Cloze speed	WRRT speed	Rating of benefit score
VS score	0.272**	0.033	-0.143	-0.112	0.111
Cloze increase		0.104	-0.494***	-0.124	0.202*
WRRT increase			0.052	-0.188*	0.282**
Cloze speed				0.231*	- 0.038
WRRT Speed					-0.111

*p<.05. **p < .01. ***p < .001

Given that VS scores were correlated with increased reading speeds on the cloze task, participants categorised with visual stress were predicted to read quicker on the cloze task with any filter. A VS score of more than 0.5, equating to four total symptoms of visual stress across the two questionnaires, was used to categorise participants. Twenty-two participants were categorized with visual stress (visual stress group). Eight of these participants had been given leading instructions and 14 had been given non-leading instructions.

A mixed ANOVA was administered with the between-subjects factor of participant type (visual stress / non-visual stress group) and the within subjects factors of reading task (WRRT/cloze task) and filter application (filter/no filter). There was a main effect of measure, with significantly quicker reading speeds for the WRRT (M = 163.87 , S.D. = 25.73) than for the cloze reading task (M = 148.79, S.D = 31.66), $F(1,114) = 7.10$, $p = .009$.

A three way interaction was found $F(1,113) = 14.42$, $p < .001$. Simple effects analysis revealed a significant interaction between the filter and participant group for the cloze task, $F(1,113) = 14.13$, $p < .001$, that was not present for the WRRT, $F(1,113) = 0.045$, $p = .832$. Post-hoc *t*-tests revealed that the VS group read significantly slower without a filter (M = 132.36, S.D = 30.52) than with a filter (M = 153.49 , S.D = 27.34), $t(21) = 4.283$, $p <.001$, $d = .730$, and significantly slower than the non-VS group without a filter (M = 152.05 , S.D. = 30.76), $t(113) = -2.681$, $p = .025$, $g = .641$, See Figure 5.5.

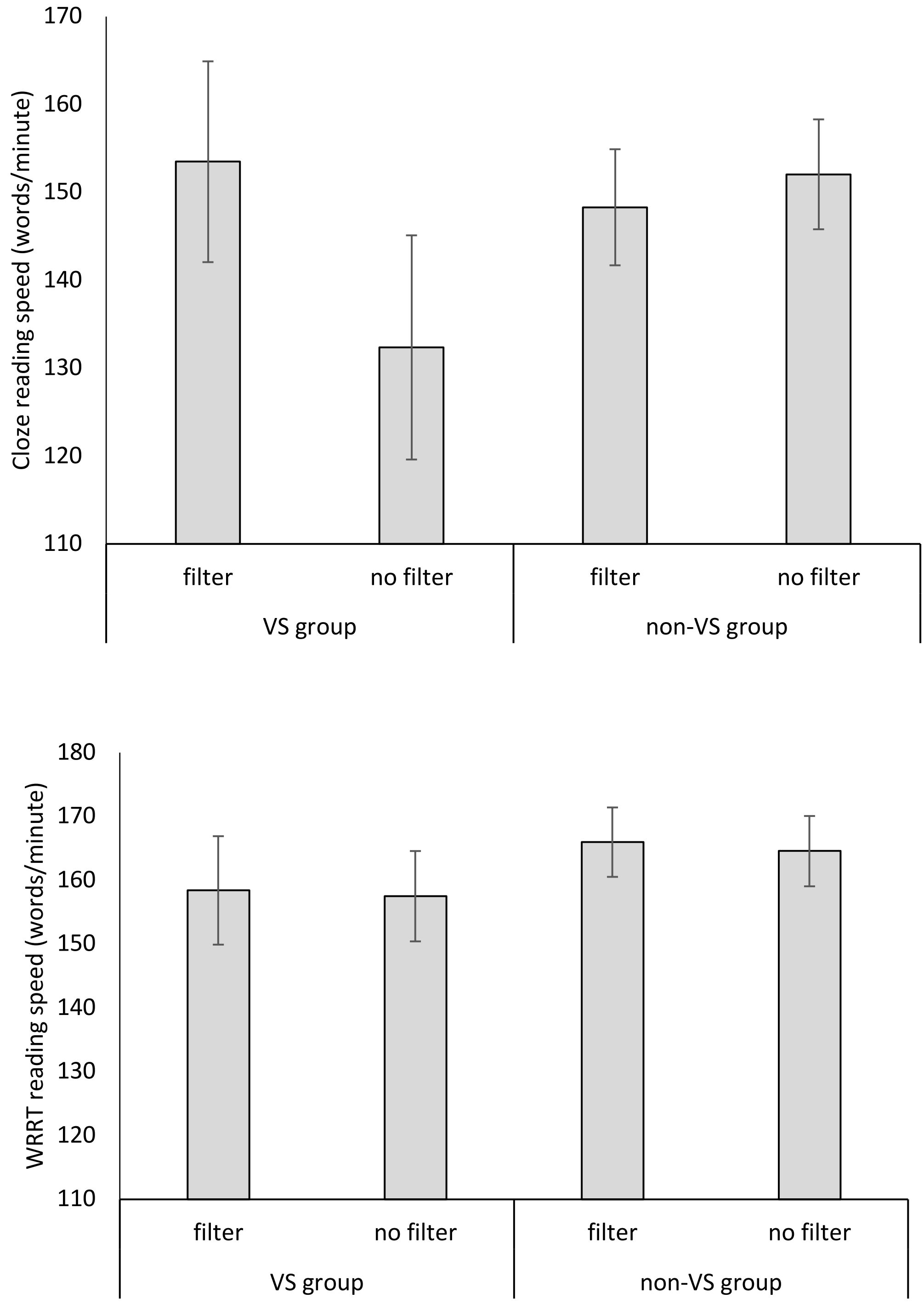


Figure 5.5. Reading speeds with and without a filter in participants with (VS group) and without (non-VS group) visual stress symptoms when using the cloze task (top) and Wilkins Rate of Reading Test (bottom).

It was possible to isolate the effect of a chosen coloured filter on the reading speeds of visual stress participants for both reading tasks because 12 participants with visual stress happened to be in the chosen coloured filter condition; five participants had been given leading instructions and seven had been given non-leading instructions, and the order of filter/no filter condition was evenly split. Paired sample *t*-tests were conducted to determine if reading speeds were faster in the filter than no-filter condition for both reading tasks for VS participants who used chosen coloured filters for reading. These participants read significantly quicker with their chosen coloured filter on the cloze task but not on the WRRT (Table 5.6).

Table 5.6.

Paired samples t-tests comparing reading speeds with and without the filter for the group of 12 people with visual stress who had been allocated a chosen filter for each task (Experiment 6)

Task		M	S.D	*t*(11)	*p*(2-tailed)	*d*
Cloze	filter	156.7	21.77	2.73	.02	0.702
	no filter	136.76	35.04			
WRRT	filter	160.92	17.36	0.999	.34	0.199
	no filter	158.08	11.12			

5.5 Discussion

5.5.1 Task Order effect of the WRRT

Whilst there was no order effect within each task, the WRRT retrieved significantly faster reading speeds when it was conducted after the cloze task. Cloze reading times were significantly slower than WRRT times, which indicates that the task was more arduous. Therefore, this task order effect may have been the result of tiring from having read for five minutes.

5.5.2 Increased reading speed on the WRRT with a chosen coloured filter reflects placebo

With a self-chosen coloured filter, reading speed improvements were boosted when participants were led to believe that their chosen colour would be beneficial (Figure 5.4). This cannot be interpreted as an effect of visual stress since there were an equal number of participants with visual stress in the groups with leading and non-leading instructions. This apparent placebo effect is specific to the WRRT (there was no similar effect of instructions on the cloze task), and specific to chosen coloured filters (there was no effect with other filter types). It thus appears that leading instructions induce a placebo response on the WRRT when a filter is chosen, but not assigned.

This is of interest because leading instructions are used in the filter choosing procedure of visual stress diagnosis: participants are told to find the colour which benefits them, thereby inadvertently communicating that one colour is beneficial. The present findings also indicate that designs that have found increased WRRT reading speeds with chosen coloured filters but not with assigned 'control' filters with leading instructions (Bouldoukian et al., 2002; Wilkins & Lewis, 1999) did not discount, and were likely to be the result of a placebo effect. Studies and diagnosis methods that conducted symptomatic questionnaires prior to the filter choosing procedure, or used participants with reading problems who were motivated to improve their reading, may also have enhanced this placebo effect.

5.5.3 WRRT reading speed increases do not diagnose visual stress or predict a reading benefit

The number of reported visual stress reading symptoms was not found to correlate with increased reading speeds on the WRRT (Table 5.5). Furthermore, participants who had visual stress, identified on the basis of their responses to the symptomatic questionnaire, did not read significantly quicker on the WRRT (Figure 5.5). This was also the case when only the participants who had been allocated a chosen filter were analysed (Table 5.6). It appears that WRRT reading speed increases do not identify people who experience symptoms of visual stress when reading.

Reading speed increases of both tasks did not correlate. This indicates that increased WRRT reading speeds do not reflect a benefit to reading and so, according to the simple model of reading, do not measure a benefit to decoding text by reduced visual stress.

The significantly faster reading speeds retrieved on the WRRT than the cloze task may demonstrate the comparative lack of attentional resources needed for a reading task when comprehension is not measured. The reading speeds of both reading tasks correlated, which is unsurprising since they both required word decoding processes. However, the correlation was weak, which suggests that the tasks are largely measuring different things. Word frequency, length and predictability have all been suggested to affect the speed of processing words (Rayner, Reichle, Stroud, Williams & Pollatsek, 2006), and these are all represented differently in the two reading measures. On the WRRT, words are separately identified and read aloud over one minute, rather than read silently, in context, over a prolonged period of time. Eye-movements would need to focus on each word, rather than focus on a few points in each sentence (Clifton et al, 2015).

5.5.4 Any type of filter appears to reduce visual stress

Regardless of filter-type, the number of reported visual stress reading symptoms was found to correlate with increased reading speeds on the cloze task (Table 5.5), participants with visual stress read significantly slower without than with their filter (Figure 5.5), and this pattern extended to participants who had been allocated a chosen filter (Table 5.6). The strong correlation between reading speed and reading speed improvement on the cloze task for all filter types (Table 5.5) indicates that slower readers are more likely to have visual stress and benefit from a filter. When participants with visual stress read with a filter, they read at similar speeds to participants without visual stress (Figure 5.5).

In this thesis, text contrast was determined by substituting the minimum luminance of the Michelson contrast formula (2) with the average text pixel luminance (section 2.4.6). If minimum pixel luminance had been used, then these findings might indicate that reduced text contrast reduces visual stress. However, the use of the average text pixel luminance to represent the minimum luminance resulted in increased text contrast with filter use (Appendix XI). Since luminance was not controlled in this experiment, perhaps the performance of these participants was improved due to reduction of contrast between the screen and the dark surround?

5.5.5 Ratings of benefit

There was a weak correlation between the rating of benefit of the filter and WRRT reading speed increases (r = 0.282) and cloze reading speed increases (r = 0.202) (Table 5.5). Since the WRRT appears to measure a placebo response and the cloze task appears to measure reduced visual stress symptoms, this may suggest that people's perception of reading benefit does not clearly distinguish between a placebo effect and reduced visual stress.

Chapter 6: Contrast studies

Some research appears to support that a reduction in text contrast may reduce visual stress (section 1.12). Coloured overlays were reported to reduce the contrast of the underlying text by 2-5% (Wilkins et al., 2001) but no WRRT reading speed benefit of reduced contrast on visual stress was found using assigned and chosen grey overlays (Jeanes et al., 1997; Wilkins et al., 2001). This finding overlooked the uncertainty over whether the WRRT reading speed increases were a placebo effect, and Chapter 5 indicated that the increased reading speeds associated with the WRRT are indeed a placebo effect. Therefore, researching the effect of contrast on visual stress is a necessary step towards understanding the nature of the apparent benefit of coloured filters to reading.

Typical readers appear to prefer, and read better in, high text contrast conditions. High contrast colour combinations, including black on white text, yield faster reading rates (Tinker and Paterson, 1931), and higher legibility scores (Poffenberger, 1925; Preston, Schwankl & Tinker, 1932) than lower contrast combinations. However, the effect of text contrast on reading performance for those with visual stress is under researched.

Visual stress has been subjectively reported to be triggered by high contrast stimuli in the peripheral field, for example by headlights of oncoming traffic when driving or glare and high contrast environmental patterns in the classroom (Wilkins, 2003). Irlen testing specifies that the full page should be covered by the overlay and no area of white should show (Irlen, 2010), again indicating that contrast in the periphery field may contribute towards the development of visual stress.

This chapter tests the contrast hypothesis by measuring how text contrast affects preference rating (Experiment 7) and reading performance (Experiment 8).

6.1 Experiment 7: paper preferences for reading

Experiment 7 used a public event to obtain data. The aim was to investigate which paper types were preferred for reading for people with and without visual stress. Paper types were chromatic or achromatic and had one of three levels of luminance: light, intermediate (equivalent to that achieved by applying a coloured filter), and dark.

6.1.1 Method

6.1.1.1 Participants

80 participants (23 female) were recruited at a public event (@Bristol, 02.2016). Participants were between the ages of 18 and 35. Four participants self-identified to be dyslexic. Four participants used coloured filters for reading, one with and three without dyslexia. All participants had had an eye test within the last two years. Forty participants had normal visual acuity and 38 had corrected to normal vision. Two participants self-identified themselves as colour blind.

6.1.1.2 Stimuli

Display

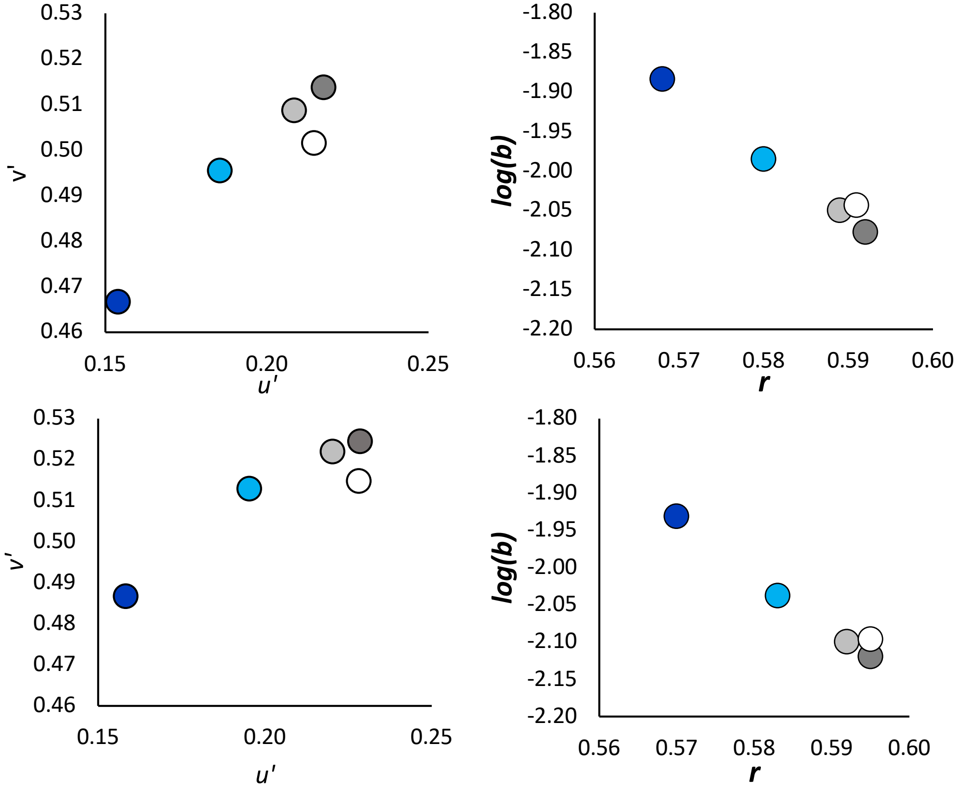
Five copies of the same text (Appendix VIII) were printed onto five different pages of A4 paper (paper conditions). Texts were backed on black card and displayed on a A0 black foam board with Velcro in an arrangement that did not draw attention to any one page (Figure 6.1). The arrangement of paper conditions was pseudorandomised so that they were presented in each of the five possible locations an equal number of times.

Figure 6.1. Arrangement of paper conditions for Experiment 7. For each participant, the height of the display was adjusted so that their eye level was at the centre of the display. The presentation of paper conditions were offset so that no condition was directly in line with eye level vertically or horizontally. Although this example has 'light blue' at the top of the display, the position of the paper conditions were rotated throughout the experiment so that their order was counterbalanced across participants.

The paper conditions were white, light blue, light grey, dark blue and dark grey. The event was held indoors with a mixture of natural and LED lighting. Under these lighting conditions, the light blue and light grey, and the dark blue and dark grey, paper conditions were approximately the same luminance, and the two shades of blue and grey were approximately the same hue (Figure 6.2).

Handout

There were five handouts (Appendix IX) which contained a matching, pictorial representation of the corresponding display, and a measure of visual stress. The criteria for categorising participants with visual stress was simplified from previous experiments. Visual stress measure 1 (Appendix IV) was used because it was quick to administer, did not contain questions relating to brightness of the page that may have influenced preference decisions and, although it was not as thorough, it had been highly correlated with visual stress measure 2 (section 5.4.5).



	CIE (1976)				MacLeod-Boynton (1979)		
Colour	L	u	v	S_{uv}	r	B	Log b
DG	29.8	0.218	0.514	0.019	0.592	0.008	-2.077
DB	34.1	0.154	0.467	0.095	0.568	0.013	-1.884
LG	55.9	0.209	0.509	0.029	0.589	0.009	-2.05
LB	56.4	0.186	0.495	0.054	0.58	0.01	-1.985
W	81.1	0.215	0.502	0.025	0.591	0.009	-2.043

	CIE (1976)				MacLeod-Boynton (1979)		
Colour	L	u	v	S_{uv}	r	b	log b
DG	17.2	0.229	0.524	0.014	0.595	0.008	-2.12
DB	14	0.158	0.487	0.083	0.57	0.012	-1.931
LG	25.3	0.220	0.522	0.019	0.592	0.008	-2.1
LB	28.4	0.195	0.513	0.041	0.583	0.009	-2.038
W	39.5	0.228	0.515	0.009	0.595	0.008	-2.096

Figure 6.2. Chromaticities and luminance properties of paper conditions in Experiment 7. 'DG', 'DB', 'LG', 'LB' and 'W' represent the paper conditions of 'dark grey', 'dark blue', 'light grey', 'light blue' and 'white'. CIE (1976) and MacLeod-Boynton (1979) diagrams displaying the chromaticities of papers that were used in Experiment 7. Measurements were taken at two different times for comparison due to the variable natural lighting. Corresponding values of chromaticities, luminance and saturations are displayed in the adjacent tables

6.1.1.3 Procedure

The experiment took place over two hours of an afternoon at a public event. Participants were asked to stand on a marked spot in front of the display. The height of the display was adjusted so that the middle of the display was at eye-level. Once positioned, they were asked to look at each page of text, decide the order in which the displays were easiest to read and indicate their order of preference by writing a number, from 1-5, on a diagrammatical presentation of the display. One would indicate their highest preference and five would indicate their lowest preference. They then filled out the short questionnaire. The order that the pages were displayed, and corresponding handouts, were changed after testing every five participants so that at the end of the experiment, an equal number of participants had completed the task with each display.

6.1.2 Analysis

One participant wrote the same number on each presentation of text. These data were removed from the analysis.

A threshold of two or more symptoms of visual stress was applied to categorise participants with (VS group) and without visual stress (no VS group). Fifty-nine participants reported 0-1 symptoms of visual stress, and 20 participants reported two or more symptoms of visual stress[14]. Only two participants in the VS group preferred a white background.

[14] A chi-squared test of association determined that this percentage was not significantly different to the percentage of participants with visual stress that had been recorded in Experiment 6.

Appendix XXIII displays the preference ratings of the colour conditions according to the number of visual stress symptoms for participants who were categorised in the VS group and the no VS group. A Mantel-Haenszel test of trend was run to determine whether a linear association existed between the number of symptoms and preference for white paper. The Mantel-Haenszel test of trend showed a statistically significant linear association between number of symptoms and preference for the white page, $\chi^2(1) = 27.718$, $p < .001$, $r = .596$. Higher preferences for white paper were associated with a lower number of symptoms and vice-versa.

To assess if the there was a difference in preference between the papers of the same luminance, a Wilcoxon signed-rank test was conducted between the preferences of the papers with the same luminance. Both participant groups demonstrated no significant difference in preference when the paper luminance was the same ($p < .05$), (Table 6.1).

The preferences of the papers with the same luminance were averaged to result in three preference groups corresponding to the three paper luminance conditions; Lum1 (white), Lum2 (light grey and light blue) and Lum3 (dark grey and dark blue). A Friedman test was run to determine which paper luminance was most preferred for the VS and no VS group. Preferences were statistically different in the VS group, $\chi^2(2) = 12.371$, $p = .002$, and no VS group, $\chi^2(2) = 47.459$, $p < .001$. Table 6.2 and Table 6.3 show the paired comparisons of each participant group. The VS participants preferred the light grey/blue paper, and there was no significant difference between their preferences of the white and dark grey/blue paper. In contrast, the no VS group preferred the white paper to the light grey/blue paper, and the light grey/blue paper to the dark grey/blue paper.

Table 6.1.

Wilcoxon signed-rank tests between the preference of paper types with the same luminance in each participant group (Experiment 7)

Participant group	Paper types (medians)		z	Adjusted target alpha	*p*
VS group	LB (2)	LG (1)	-0.953	0.05	.340
VS group	DB (4.5)	DG (3)	-1.684	0.025	.092
No VS group	LB (3)	LG (3)	-0.424	0.0167	.671
No VS group	DB (4)	DG (4)	-1.932	0.0125	.053

Note. LB = light blue, LG = light grey, DB = dark blue, DG = dark grey. Numbers in brackets are medians.

Table 6.2.

Paired comparisons of Freidman test between the preference of paper types of VS group (Experiment 7)

Comparisons		z	Adjusted target alpha	*p*
Lum1(4)	Lum2(1.75)	-2.925	.05	.003*
Lum2(1.75)	Lum3(4)	2.767	.025	.006*
Lum1(4)	Lum3(4)	0.158	.0167	.874

Note. Lum1 = White, Lum2 = luminance of light grey or light blue paper, Lum3 = luminance of dark blue or dark grey paper. Numbers in brackets are medians.

Table 6.3.

Paired comparisons of Freidman test between the preference of paper types of no VS group (Experiment 7)

Comparisons		z	Adjusted target alpha	p
Lum1(2)	Lum2(2.5)	-6.444	.05	< .001*
Lum2(2.5)	Lum3(4)	-3.498	.025	.003*
Lum1(2)	Lum3(4)	-6.444	.016667	< .001*

Note. Numbers in brackets are medians.

6.1.3 Discussion

The participants' preference rating was not affected by whether the paper was coloured and appeared to be affected by its luminance and contrast. Participants categorised with no visual stress preferred text condition 1 (white paper) and demonstrated decreasing preference as the paper became darker. This result is supported by previous research showing that high contrast text is preferable for normal readers (Poffenberger, 1925; Tinker and Paterson, 193; Preston, Schwankl & Tinker, 1932).

Participants categorised with visual stress rated text condition 2 the highest and appeared to dislike the darkest and lightest paper equally. This supports the contrast hypothesis by indicating that a reduced contrast text, with luminance and text contrast of paper with a filter, is preferred by visual stress sufferers.

The effect is unlikely to be a placebo effect because the participant groups were from the same population. Most of these participants were unlikely to be invested in the idea that one text condition might benefit their reading because they did not use coloured filters to read or have dyslexia and so were unlikely to have been exposed to the notion of coloured filters for reading. The four dyslexic participants were evenly split across the participant groups. The four filter users were in the visual stress participant group, but their removal did not make a difference to the analysis. Experiments that manipulate contrast directly are needed to verify if the contrast hypothesis is correct.

6.2 Experiment 8: How does reducing text contrast affect reading performance of filter users?

Experiment 8 tested the contrast hypothesis by investigating if reducing the text contrast improved cloze reading performance of filter users. It was assumed that filter users had visual stress because all filter using participants in Experiment 4 had reported four or more symptoms of visual stress. The text contrast was manipulated by adjusting the luminance and by increasing the spacing between the lines.

6.2.1 Method

6.2.1.1 Participants

Twenty-four filter users (20 female), aged 17-36, who used coloured filters for reading were recruited by online and paper advert, and website in exchange for payment. The recruitment of filter using participants was exceedingly slow and so the criteria was relaxed to include people who used coloured filters on computers or tablet devices for reading (screen filters) or had found filters to be beneficial at some point in their lives. Eleven participants used lenses, 16 used overlays and 11 used screen filters, with two of these participants using only screen filters and not lenses or overlays. One participant used an orange (blue light blocking) filter on their phone and said that it helped their reading. Five participants did not use filters anymore but had reported them to be beneficial in school for at least two years. All other participants were currently using their filters and had been using them from between one month to 14 years.

Twenty-four control participants (20 female), aged 18-25, were recruited from the psychology undergraduates. All control participants reported three or fewer symptoms of visual stress.

One filter user, and no control participants self-identified as colour blind. All participants had received an eye test within the last two years. Participants with reading difficulties were identified as they had been in previous experiments (Table 6.4).

Table 6.4.

Number of participants who had reading difficulties (phonics difficulties, self-reported dyslexia and both of these) and did not have reading difficulties in each participant group (Experiment 8)

	Filter users	Control	All participants
Reading difficulties	20	3	23
• Phonic difficulties	7	2	9
• Self-reported dyslexia	9	0	9
• Both	4	1	5
No reading difficulties	4 (3)	21 (20)	25 (23)

Note. Numbers in brackets represent number of participants with outliers removed

6.2.1.2 Stimuli

Reading performance was measured with the cloze task (Figure 4.3). Visual stress symptoms and phonic ability were measured with the questionnaire (Appendix IV) and phonic task (Figure 4.6) according to the previous specifications.

The stimuli in this experiment were displayed on an uncorrected screen to achieve a text luminance of 96.552 cd/m^2; this acquired a normal text condition with a closer approximation to the measured text luminance (97.9cd/m^2), and therefore the text contrast, of printed normal text (section 2.4.2) than previous experiments. There were three conditions:

1. *Normal (baseline) condition:* luminance properties to match normal text without an overlay
2. *Reduced contrast condition*: luminance to match normal condition and RMS contrast to match an overlay condition
3. Increased linewidth condition: RMS contrast to match the normal condition, overall luminance to match the reduced contrast condition and increased line spacing

Unintended luminance properties of Condition 2

As in Experiment 1 and Experiment 2, Condition 2 had matched RMS contrast to an overlay condition but Michelson text contrast, determined to be the most accurate way to define screen contrast post testing (section 2.4.6.2), calculated the text contrast of Condition 2 to be 71% lower than intended. luminance and contrast measurements of the three conditions are displayed in Appendix XXIV.

6.2.1.3 Procedure

Figure 6.3 displays the sequence and timings of Experiment 8. The cloze task was completed for eight minutes in each condition. The order of conditions was counterbalanced so that each order was completed the same number of times. Participants completed the questionnaire (Appendix IV).

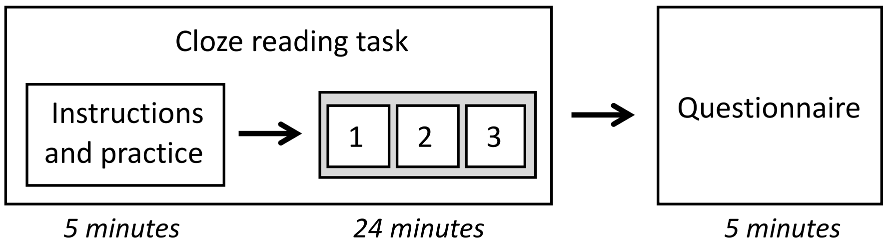


Figure 6.3. Sequence and timings of Experiment 8. 1, 2 and 3 are normal, reduced contrast and increased linewidth text conditions respectively (see section 6.2.1.2). Grey border represents counterbalancing of conditions across participants. Timings are in italic font.

6.2.2 Results

Comprehension answers were over 70% correct, and comprehension levels were evenly distributed across the data. Reading speeds were calculated as in previous experiments (section 3.2.3.3) and categorised according to text condition and participant group. One participant in each group was found to be more than 2.5 standard deviations from the mean of a group. The data of this participants were removed from the analysis, but the removal did not change the outcome of the analysis.

6.2.2.1 Test for order effects

The reading speed data were organised into the order of conditions that had been administered for each participant, and a one-way within groups ANOVA was conducted for the data of each participant group (Figure 6.4). Reading speeds in the three order groups were statistically different in both experimental, $F(2,44) = 13.122$, $p < .001$, $\eta_p^2 = .374$, and control groups $F(2,44) = 8.627$, $p = .001$. $\eta_p^2 = .282$. For both participant groups, pairwise comparisons revealed that the condition that had been completed first elicited significantly quicker reading speeds than the second and third condition ($p < .05$), and there was no significant difference between the reading speeds of the second and third condition ($p > .05$), see Table 6.5 and

Table 6.6.

Table 6.5.

Pairwise comparisons for order of conditions for filter users (Experiment 8)

Text condition comparisons		*t*(22)	Adjusted target alpha	*P* (2-tailed)	*d*
1	2	4.744	.05	< .001*	.450
1	3	4.387	.025	< .001*	.431
2	3	-0.271	.0167	0.789	.023

Table 6.6.

Pairwise comparisons for order of conditions for control group (Experiment 8)

Text condition comparisons		*t*(22)	Adjusted target alpha	*P* (2-tailed)	*d*
1	2	3.687	.05	.001	.476
1	3	3.444	.025	.002	.640
2	3	1.062	.0167	.300	.195

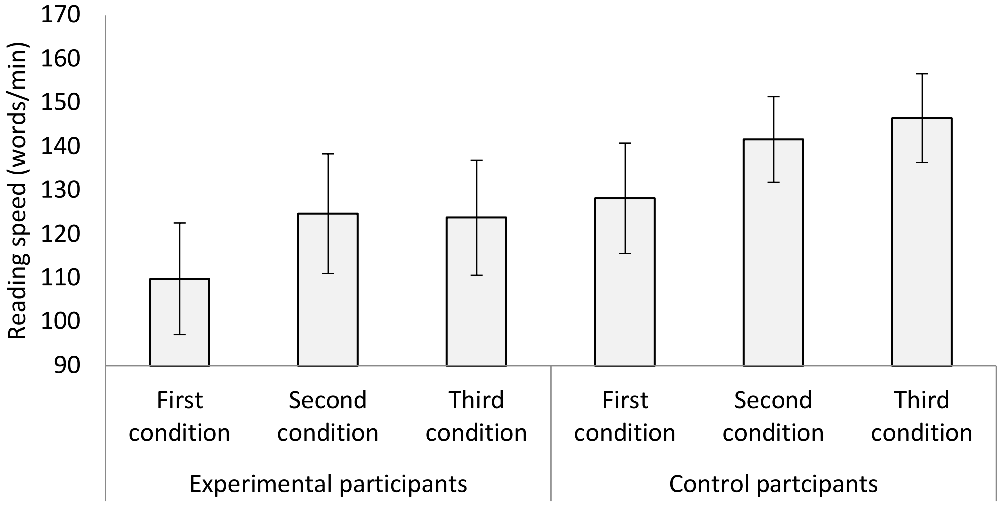


Figure 6.4. Reading speeds when categorised according to the order that the text conditions had been completed (Experiment 8).

6.2.2.2 Test for effects of text condition

A one-way within groups ANOVA was conducted on the reading speeds of the experimental and control participant groups (Figure 6.5). Reading speeds in each text condition were not found to be statistically different in the experimental group, $F(2,44) = 0.67$, $p = .937$, $\eta_p^2 = .003$. However, reading speeds were found to be statistically different in the control group, $F(2,44) = 4.566$, $p = .016$, $\eta_p^2 = .172$. Pairwise comparisons revealed that Text condition 3 elicited significantly quicker reading speeds than Text condition 1 ($p = .014$[15]) , and there were no significant differences between other text conditions (Table 6.7).

Table 6.7.

[15] .017 when the outliers was not excluded

Pairwise comparisons of one-way ANOVA (Experiment 8)

Text condition comparisons		*t*(22)	Adjusted target alpha	*P* (2-tailed)	*d*
3	1	-.598	.05	.014*	.472
3	2	-2.674	.025	.032	.431
1	2	-2.287	.016667	.556	.099

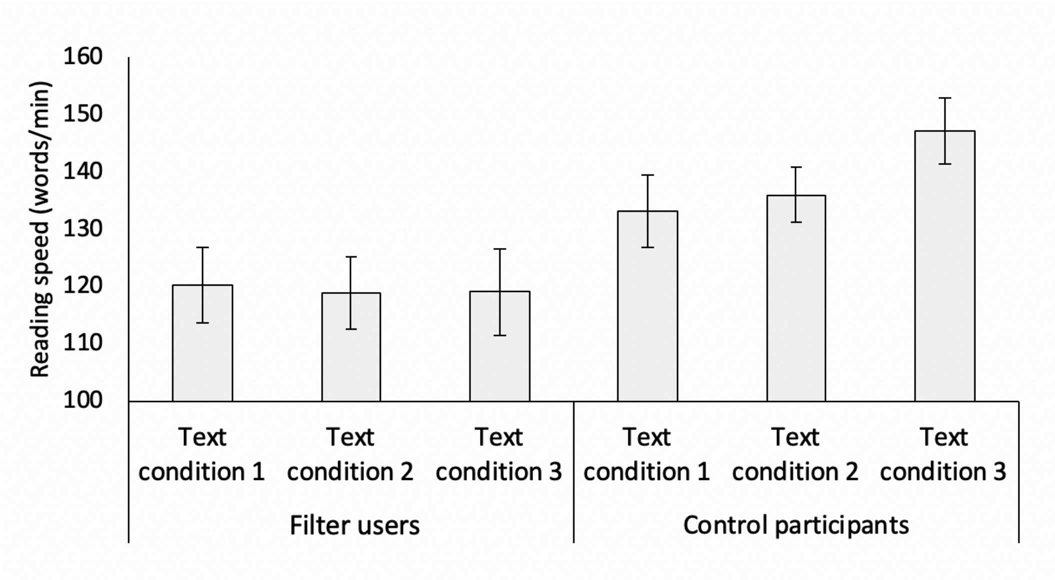


Figure 6.5. Reading speeds of cloze task categorised according to normal, reduced contrast and increased linewidth text conditions for experimental and control participants.

6.2.2.3 Test for effects of participant group

Filter users read significantly slower (M= 122.28, S.D. = 33.22) than control participants (M= 141.85, S.D.= 27.36), $t(44) = 2.327$, $p = .025$, $d = .643$.

6.2.2.4 Questionnaire

With the exception of two participants, all filter users reported four or more symptoms of visual stress. The two participants in question reported two and three symptoms of visual stress respectively. The first was using a screen filter on their phone, and the second was using a pair of blue glasses that had been administered by the Intuitive method in an opticians. Both these participants read quicker in the normal text condition, but their removal did not change the outcome of the analysis.

6.2.3 Discussion

The results of Experiment 8 did not support the contrast hypothesis because reading performance of filter users was not improved when text contrast was reduced in text condition 2. However, this experiment suffered from the same methodological flaw that resulted in unintended luminance properties of conditions in Experiment 1 (section 3.1.2.2) and 2 (section 3.1.5.2). The RMS contrast formula (3) rather than the Michelson text contrast formula (4)[16] had been used to quantify the contrast value of condition 2, resulting in the text contrast of condition 2 being 71% lower than intended (or if calculated by Michelson text contrast). The result of this miscalculation may be that the contrast that was applied to condition 2 may be lower than the range of text contrast that is beneficial to people with visual stress. It should be noted, however, that RMS contrast was an appropriate measure for calculating an equivalent contrast to overlay text of the increased linewidth condition (condition 3), which had the

[16] Section 2.4.6 describes the rational behind using Michelson text contrast rather than RMS contrast to determine the contrast of text that is displayed on screen.

same luminance values as normal text (condition 1) but a higher percentage of pixels allocated to the (white) text background.

It could also be argued that the reduced clarity (blurring) of condition 2, due to its lower range of greyscales, reduced its legibility and impeded any improvement in reading speed that may have occurred from reduced visual stress. The reduced clarity of condition 2 did not affect reading performance of the control participants (Figure 6.5). However, filter users may have been more affected by the clarity of the stimuli that the control participants because they were less fluent readers than control participants, reading about 20 words/minute slower (section 6.2.2.3) and having more reading problems (Table 6.4.). Since non-fluent readers spend more time fixating on words than fluent readers (Rayner, 1986), it would follow that filter users in this experiment spent more time focussing on the words to process them. Therefore, the clarity of the text in condition 2 may have been more relevant to the reading performance of filter users than control participants.

Condition 2 reduced text contrast whilst maintaining equal page luminance to normal text (condition 1). This resulted in a high contrast edge between the screen and surround that may have triggered visual stress and prevented the reduced text contrast from yielding an effect. Although macular (central) vision is highest resolution, contrast is also detected in the peripheral fields (Virsu & Rovamo, 1979). It seems appropriate to consider the effect of contrast in the peripheral field since the stimulus used to trigger visual stress is a page of text, rather than a single word.

The absence of findings in this experiment may also be due to the high proportion of people with reading difficulties (87% with the outlying participant removed). Even if reduced text contrast had reduced their symptoms of visual stress, these participants' reading difficulties might have prevented an increase in their reading speed.

It is also plausible that the participant criteria used in this experiment did not successfully recruit people with visual stress. The 'filter users' criteria was relaxed due to difficulties with recruitment: unlike Experiment 3 and 4, participants who had used filters any time and participants who changed the colour of their computer screens (screen filters) were included. Two participants reported below the threshold that was being used to categorise visual stress (three or less symptoms of visual stress), one of whom had recently been diagnosed with the Intuitive method. This highlights the question over the validity of the diagnosis methods of visual stress. It is possible that more filter users did not have visual stress and reported inflated numbers of symptoms due to their awareness and previous exposure to the filters.

An order effect was observed in the cloze task across for reading speeds of both filter users and control participants between the first and second condition. This underlines that a practice effect appears to exist for the cloze task. However, there was no order effect between the second and third condition for either participant group, which suggests that once the task is learnt, there are no practice effects of the cloze task.

Control participants read quicker when spacing between lines was increased and this effect was not observed for filter users. This suggests that increasing line spacing benefits reading performance of typical readers, possibly due to reduced crowding: an unexpected finding. Pelli, Tillman, Freeman, Berger and Majaj's, 2007 demonstrated that Rapid serial visual presentation (RSVP) reading rate is determined by the length of visual span (the number of letters that can be recognised in one fixation), and visual span is determined by the number of characters that are not crowded in the text. Increased letter spacing was found to increase visual span (Yu, Cheung, Legge & Chung, 2007) and quicken oral RSVP (Chung, 2002; Yu, Cheung, Legge & Chung, 2007) and sentence reading speed (Yu, Cheung, Legge & Chung, 2007) in typical readers (aged 19-34) and sentence reading speed in dyslexic readers (aged 8-14) (Zorzi et al., 2012). The stimuli used in these experiments also increased the line spacing. Zorzi et al. (2012) did not find an effect in control subjects and suggested that dyslexic readers were abnormally affected by crowding. The results of Experiment 8 suggest that the reading performance of typical readers are also affected by reducing crowding due to increased line spacing when the task is to read text silently for understanding.

Chapter 7: Assessing the chosen filters

7.1 Chosen filters of Experiment 3, 4 and 6

Throughout the experiments, the chosen chromaticities of filter span the ordinate and abscissa of the MacLeod-Boynton (1979) cone-excitation diagram (Appendix XXV). As such, it was not possible to examine the effect of chromaticities that lay on lines of constant S or L-M cone. However, this random distribution indicates that colour preferences are not related to a change in sensitivity of S or L-M cone excitation.

Figure 7.1 displays the number of participants who chose each chromaticity of filters in Experiment 3, 4 and 6. Overall, aqua and orange were the most popular colours, accounting for 33% of the filter choices; yellow-orange, yellow, lime-green and purple were the least popular colours, accounting for 6.5%.

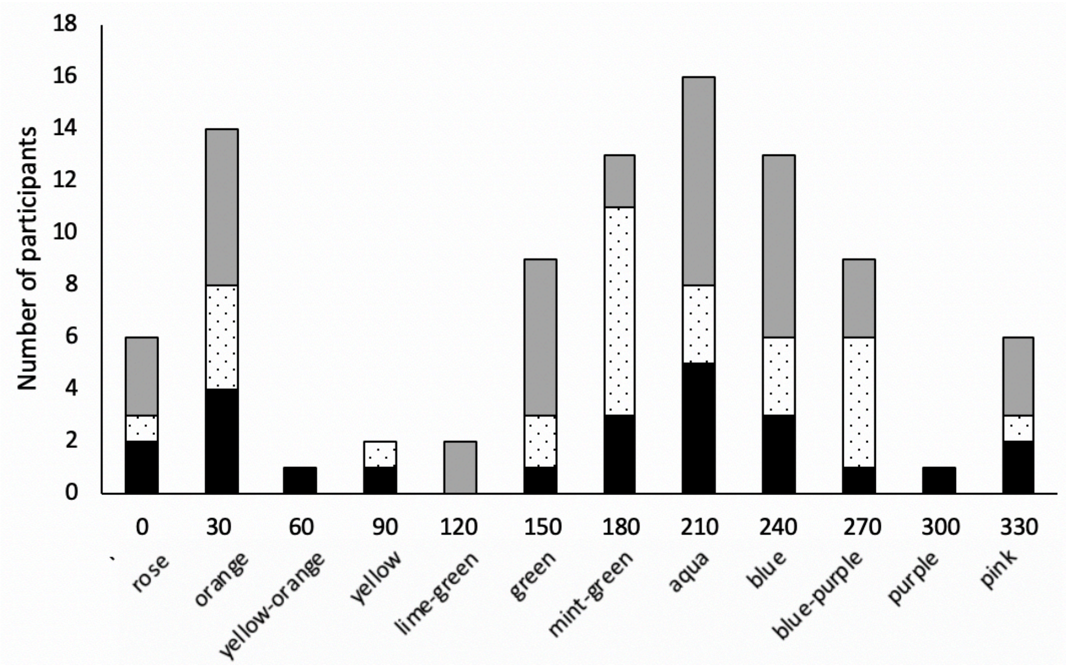


Figure 7.1. Stacked bar chart categorising chosen filters of participants in Experiment 3 , Experiment 4 and Experiment 6 . Chosen filters (Appendix XXV) were categorised according to their hue-angle proximity to 12 hue angles obtained from the colorimeter (Appendix XV). Colour descriptors are for guidance only.

7.2 An effect of gender?

Conway et al. (2016) proposed that colour choices of filter were influenced by gender (see section 1.11.3). By referring to previous research on stereotypical gender-based colour preferences in western populations (LoBue & Deloache, 2011; Al-Rasheed, 2015), Conway et al. (2016) categorised the Intuitive overlay colours of ‘rose’, ‘pink’ and ‘purple’ as ‘female colours’, ‘blue’, ‘green’ and ‘blue-purple’ as ‘male colours’, and the remaining colours as ‘neutral colours’.

Conway et al.'s (2016) study appears to have oversimplified the cited literature, overlooking the impact of age and culture on gender based colour preferences; the effect and interaction of gender, age and culture on colour preferences are poorly understood and difficult to disentangle (Hurlbert and Ling, 2017). For example, the preference of pink amongst females appears to be both culture and age dependant. Culturally, Al-Rasheed's (2015) demonstrated in undergraduates that, although Arabic females were more likely to prefer red-pink colours from a pair of randomly presented colour options, English females were more likely to prefer purple or blue-green colours. With regard to the effect of age, LoBue and Deloache's (2011) categorisation of pink as a female colour was based on the choices of four year olds; Sorokowski, Sorokowska and Witzel (2014) demonstrated colour preferences for Swiss females depended upon their age: pink (or purple) was more likely to be preferred by girls than boys (aged 10-14 years), but red (not pink) was more likely to be preferred by woman (aged 17-48 years) than men.

Conway et al.'s (2016) categorisation of gender-based colour preferences relied on research that used subjective names to categorise the colours of overlays, preventing the accurate comparison of colour choices between studies. Colours with the same colour name can vary dramatically (especially with regard to colours of the same hue but different saturation; such as red and pink). A universal colour preference model (Ling and Hurlbert, 2007) has allows more accurate comparison between studies. Using this model, adults (aged 20-26) from the UK demonstrated a preference for the blue end of the blue-yellow axis, with females preferring colours on the red end of the red- green axis more than men (Hurlbert and Ling 2007; Ling, Hurlbert & Hurlbert, 2011).

7.2.1 Method

Two analyses were adopted to investigate the effect of gender on colour preferences of chosen filters in this thesis.

The first analysis conducted a chi-squared analysis to determine if there was a statistically significant association between gender and colour preference according to Conway et al.'s (2016) gender stereotypical categorisation of filters. Since Conway et al.'s (2016) effect was only found over a large sample size, data from Experiment 3, 4 and 6 were combined and analysed collectively. To allow colours of filters in Experiment 3 and 4 to be categorised, Conway et al.'s (2016) categorisations were used to inform three boundaries in the CIE (1976) chromaticity diagram forming three gender stereotypical zones of 'male', 'female' and 'neutral' colours. Chosen filters of Experiment 3, 4 and 6 were then categorised in these gender stereotypical zones (Appendix XXVI). Data corresponding to neutral colours was removed for the purpose of the analysis.

In response to Hurlbert and Ling's (2007) finding that females preferring colours on the red end of the red-green axis more than men, the second analysis assessed whether the distributions of cone excitations on the red (L) -green (M) axis of the abscissa of the MacLeod Boynton (1979) were significantly different for filters chosen by males and females.

7.2.2 Analysis

When chosen filters were categorised according to Conway et al.'s (2016) gender categorisation of colour, a chi-squared analysis determined that there was no statistically significant association between gender and colour preference, $\chi^2(1) = 0.011$, p = .917. Both males (72%) and females (71%) preferred male, rather than female colours. Interestingly, all but one participants who chose colours that had been categorised to be neutral were female.

The distribution of chromaticities chosen by males and females on the abscissa (*r*) of the MacLeod Boynton (1979) diagram were not normally distributed (p > .05). Since there was only a small sample (26) of male participants, to avoid the inflation of a type 1 error rate, a Mann-Whitney U test, was conducted. Distributions of the *r* values for males and females were not similar, as assessed by visual inspection. Values of *r* for males (mean rank = 38.28) and females (mean rank = 48.28) were not statistically significantly different, U = 993, *p* = .102.

7.2.3 Discussion

This thesis does not support that colour choice of filters are affected by gender. No effect of gender was observed when chosen filters were categorised according to Conway et al.'s (2016) gender categorisation of colour or when r values of chromaticities chosen by males and females were analysed on the MacLeod Boynton (1979) diagram.

Conway et al.'s (2016) effect was only observed for chosen lenses, not overlays; the authors suggested that this was due to the influence of how the lenses would appear when worn. If this hypothesis was true, then filter choices in this thesis would not be expected to be impacted by gender because they were presented over the text on the computer screen and not worn on the face.

Conway et al.'s (2016) effect was small (*p* = .04) and only found for a considerably larger sample size (n = 238) than the one used in the present thesis (n = 92). The populations in this thesis were predominantly female; 71% (Experiment 3), 72% (Experiment 4) and 70% (Experiment 6). This high proportion of females is unsurprising in Experiment 6, who were psychology undergraduates. However, the reason for the similar proportion of females amongst filter users (Experiment 3 and 4) remains unclear. Regardless, further investigations into the role of gender on colour choice should aim to recruit a larger sample size with equal numbers of males and females.

Chapter 8: General discussion

This thesis has provided a reading performance measure that can be used to assess the benefit of filters, found no benefit of a chosen filter, shown that the test used to diagnose visual stress is susceptible to placebo effects and found effects of contrast. An alternative hypothesis, that the benefit of coloured filters may be in the reduction of contrast, has been suggested. The findings in this thesis relate to the benefit of filters on the reading performance of adults, not children. Further research is needed to ascertain whether these findings generalise to studies involving children.

8.1 No observed effect of a chosen hue

Irlen (1991) and Wilkins' (2003) claimed that a specific colour of filter is needed to benefit reading for people who experience visual stress. However, there is little evidence in the literature (section 1.11.1), or this thesis, to support this assertion. Experiments 4-5 demonstrated no observable benefit to reading of a chosen hue for filter users. It could be argued that an improvement in reading performance was not observed in these experiments because there was a high percentage (79%) of participants with reading problems. However, there was also no observable benefit of chosen hue on Neisser's (1964) visual search task or the placebo controlled (Stroop) single word reading task (Experiment 4). Additionally, 12 out of 20 filter users who returned to complete Experiment 5 reported that they did not believe their chosen hue of filter was more beneficial to their reading than another. Lastly, Experiment 7 demonstrated that coloured paper was not preferred more than grey paper for people who reported visual stress.

Despite an absence of findings to support an effect of chosen hue on visual stress in this thesis, more experiments are required to ascertain if a particular hue can be of benefit. The methods used to acquire the chosen filters in the experiments of this thesis impeded controlled experiments of the effect of hue on visual stress. Filters were acquired by replication of filter users' existing chosen filters (Experiment 3) or replication of the Intuitive filter choosing method (Experiment 4 and 6). Therefore, their luminance, saturations and resulting text contrasts were not held constant. The variation of saturations across these experiments is demonstrated by box plots in Appendix XXVII: filters in Experiment 4 and 6 spanned a shorter range of saturation than Experiment 3 as they used the restricted higher and lower saturation setting of the colorimeter. To provide controlled experiments of the effect of hue on visual stress, performance measures obtained when using chosen coloured filters could be compared with the those obtained when using achromatic filters of equal lightness and contrast reduction (Veszeli and Shepherd, 2019) instead of comparing to baseline; as was the case in Experiment 3 and 6.

Veszeli and Shepherd (2019) reported that the colours of overlay that children chose to be most comfortable were not the same as the colours that benefited reading speed on the WRRT or the colours that they considered to be their favourite. This may indicate that the existing filter choosing methods do not acquire colours that would be of benefit to reading. Veszeli and Shepherd (2019) also suggested the manufacturing of aids that vary stimulation in the early cone-opponent pathways. Future experiments could apply this technique to the method of choosing a filters on screen; this would allow a more thorough investigation into individual cone excitation.

8.2 Placebo effects

Chosen coloured filters appear to result in placebo effects of reading speed on some reading tasks (WRRT and single word) and not others (cloze). When typical readers were instructed that their chosen coloured filter would benefit their reading, they demonstrated a placebo effect by reading quicker with their chosen filter on the WRRT, but not on the cloze task (Experiment 6). Placebo effects were not observed on the cloze task (Experiments 4 and 5). However, a placebo effect appeared to have been observed on the single word reading task because there was no control condition and these participants did not demonstrate any effect in the controlled (Stroop) task (Experiment 4). The presence of a placebo effect on the WRRT, but not on the cloze task, suggests that placebo effects offer no immediate benefit to natural reading in adult readers. Placebo effects may have occurred on the single word reading task and WRRT and not the cloze task because they were simpler tasks. They contained less information to process, using only one word rather than text, and required less working memory by identifying and not comprehending text. Therefore, the low cognitive load necessary to complete these task may have left them with more cognitive resources allowing them to improve by way of a placebo effect.

The placebo effect of the WRRT means that opticians could be falsely diagnosing patients, and those who have reading problems, and are invested in the idea that a chosen colour may help them, may be more susceptible to placebo effects during diagnosis. The Irlen method uses a subjective response to a chosen filter and so cannot be seen to be any more robust. This calls into question whether all participants (filter users) in Experiments 3, 4, 5 and 8 actually had visual stress. Potentially, their reported high symptoms of visual stress may have been primed by their diagnosis. A validated method of diagnosis remains to be developed.

8.3 Filter users and reading difficulties

Reducing visual stress by using chosen coloured filters is claimed not to lead to improvements in reading when other reading difficulties have not been addressed (Blaskey et al., 1990; Irlen, 1991a; Wilkins, 2003; Wilkins et al., 2016). Therefore, the high percentage (79 – 83%) of participants with reading difficulties in Experiments 4, and 8 may have confounded the outcome of these results and leaves a degree of uncertainty over the absence of effects found. An observable benefit to reading performance in these participants might be limited by the presence of reading difficulties (phonic or comprehension) that reduce reading performance by processes unrelated to visual stress.

Although it is plausible that some reading difficulties may stem from visual stress, the high proportion of people with reading difficulties amongst filter users may reflect that people with reading difficulties are more exposed to filters than people without reading difficulties. Filters are not widely available, yet sometimes distributed as part of a dyslexia assessment. Additionally, people with reading difficulties are more likely to discover filters as they search to find answers to their reading difficulties.

Due to the high proportion of filter users with dyslexia, future research would benefit from using cloze reading performance of general population samples rather than filter users. Eighteen and 25% of participants in Experiment 6 and 7 were found to experience visual stress in populations of 120 and 80 participants. This gives some indication about the numbers required to gain a sample with sufficient power.

Self-reported diagnosis of dyslexia (Experiment 3) or measures of phonic difficulties (Experiment 4, 5 and 8) were used to categorise dyslexia and reading difficulties in the experiments of this thesis. However, in light of the high proportion of participants with reading difficulties in this thesis, it would appear that these methods were too brief. A more robust dyslexia assessment (e.g. Saksida et al. ,2016) or adult dyslexia questionnaire (e.g. Vinegard, 1994) may categorise dyslexia more thoroughly and inform the test that is used to measure an effect of filters.

Since participants with dyslexia or other reading difficulties are unlikely to immediately improve their reading when visual stress has been reduced, it is crucial that the tests used to measure an effect of filters do not require reading processes that are affected by their dyslexia. The use of tests that contain high contrast patterned stimuli may be more appropriate than text stimuli that involves reading continuous text. For example. Singleton and Henderson's (2007) computerised task may provide a high contrast visual search task that is suitable: the task required the location of a target three letter words within a square matrix of distractor three letter words; the stimuli was high contrast because there were no spaces between the words.

8.4 Contrast hypothesis

The results of Experiment 7 support the contrast hypothesis: that a reduction of text contrast reduces visual stress and benefits reading. Experiment 7 demonstrated that participants categorised with visual stress preferred reading from paper which mimicked the effect of filters by reducing the text contrast. In contrast, participants without visual stress preferred reading from white paper. This indicates that people with visual stress prefer reading when text contrast is slightly reduced and people without visual stress prefer reading when text contrast is at normal levels. The absence of an effect of filter use on Stroop interference could be because the RMS contrast of the stimuli were too low to trigger visual stress due to the abundance of white space surrounding the image.

Although there was no finding to support the contrast hypothesis in Experiment 8, it is possible that the contrast of the low contrast condition (condition 2) was too low to benefit the reading performance of people with visual stress. The text contrast that was applied to this condition was incorrectly calculated using RMS contrast rather than Michelson text contrast, which has been justified to better present how humans perceive screen text (section 2.4.6). Clearly, more experiments are required to ascertain if optimum text contrast levels differ for people with and without visual stress.

8.4.1 High contrast edge

Experiment 6 demonstrated that any filter benefitted reading performance for participants categorised with visual stress. The number of VS symptoms correlated with reading speed improvements on the cloze task when any filter was used, and participants categorised with visual stress read quicker on the cloze reading task with any filter than with no filter. Analysis of the contrast values of filters using Michelson text contrast revealed that filters had increased the text contrast of these filters (Appendix XXI). It would be unlikely that increased text contrast improves visual stress because filters reduce contrast of printed text. Therefore, one explanation for this improved reading performance with filters, that increased text contrast, may have been due to the reduced contrast edge where the computer screen met the surround.

A high contrast edge was also present in the stimuli of Experiment 8; this feature was not present where effects of contrast have been found, both in this thesis and the literature. Future experiments should control for this by eliminating the high contrast edge of the stimulus; this could be achieved by applying achromatic filters to control conditions (Veszeli and Shepherd, 2019). The high contrast edge could also be eliminated by presenting white (normal contrast) or light grey (reduced contrast) text stimuli on a black or dark grey text background with text luminance controlled using Michelson text contrast (4) (not RMS, see section 2.4.6.3).

8.4.2 Measuring contrast

With the resources available, the actual perceived contrast of text stimuli in these experiments, and the contrast reduction afforded by filters, may be impossible to measure: a human participant reading black text on a white background will employ various mechanisms to process light entering the eye to a percept, for example, centre-surround inhibition will accentuate the contrast of black text on a white page. Therefore, it was important that the screen stimuli used in the experiments of this thesis replicated printed text as much as possible.

The screen text stimuli that was used in the experiments of this thesis were replicated from normal text in office environments (Section 2.4.2), but this replication was hindered by the grey-scales in text on a CRT (section 2.4.5). A crucial improvement to the text stimuli of future experiments would be to use a more modern screen with rendered text that appears as it would when printed on paper. Levien (2003) concluded that 170 dpi displays can sufficiently display very high contrast text under normal viewing conditions. Such a resolution would make grey scales in text imperceptible when viewing, not just reading, the text. This would allow more control over the manipulation and measuring of text contrast.

8.5 Measures

8.5.1 A measure of reading benefit

Previous research has lacked a test that assesses whether filters benefit natural reading. The cloze task (Experiment 3) measures reading speed of silent and sustained reading, whilst monitoring comprehension with word selection tests. This task was concluded to represent a naturalistic reading task for adults when a high correlation was found between median reading speeds calculated with and without the time taken to complete word selection tests (Experiment 3). Future studies may incorporate age appropriate reading materials into this test to to assess the reading benefit of filters to children.

The cloze task had a practice effect that disappeared after 20 minutes of task practice in a previous session (Experiment 5) or after 8.5 minutes in the same session (Experiment 8). Future experiments might address this practice effect by practicing the task for about ten minutes before completing the experiment.

8.5.2 Symptomatic questionnaires

The symptomatic questionnaires (Appendix IV) appear to have been successful in categorising participants with visual stress. However, due to their brevity, there may be symptoms which were missed from the measure that, if included, could have more accurately categorised participants with visual stress or a level of visual stress severity. Additionally, the questions of the first visual stress questionnaire could be considered to be leading. Future research should systematically develop a factor analysed questionnaire that categorises participants with visual stress accurately. Alternatively, Conlon et al.'s (1999) questionnaire could be used without the questions relating to pattern-observation tasks, which may be irrelevant to visual stress in reading.

8.6 In practice

Practitioners, and individuals with visual stress, can be encouraged to use cheap and practical solutions to visual stress that reduce text contrast, such as pastel or natural coloured paper, quick print settings, neutral density filters and anti-glare filters. Coloured filters could be used to reduce the contrast of text However, it is important to refute Irlen's (1991) claims that people with visual stress have a problem filtering the white spectrum of light and that glasses should be worn all the time. Typical readers' reading performance does not appear to be inhibited by a Michelson text contrast reduction of 71% (see section 3 and Figure 6.5). Since visual stress appears to affect a large proportion (18-25%) of people, all text could be printed on to paper that is not bright white to reduce its text contrast, such as recycled paper. General changes in the appearance of text like this may benefit the reading of individuals who, without knowing that they have visual stress, might not otherwise comfortably read.

Chapter 9: Conclusions

This thesis estimates that 18-25% of the general adult population may suffer from visual stress. Methods have been developed that can be used to investigate the effects of chosen coloured filters on reading beyond placebo. In particular, the cloze task provides a sensitive measure of reading performance that monitors reading speed and comprehension over a prolonged period of time.

There is some evidence in this thesis to suggest that the effective component of chosen coloured filters may be text contrast, not colour. This would explain the apparent effectiveness of coloured filters for reading when there is an absence of evidence to support their efficacy.

References

Albon, E., Adi, Y., Hyde, C., (2008). The effectiveness and cost-effectiveness of coloured filters for reading disability: A systematic review. West Midlands Health Technology Assessment Collaboration, Department of Public Health and Epidemiology, University of Birmingham.

Allen, P. M., Evans, B. J. W., & Wilkins, A. J. (2012). The uses of colour in optometric practice to ameliorate symptoms of visual stress. *Optometry in Practice*, *13*(1), 1-8.

Allen, P. M., Gilchrist, J. M., & Hollis, J. (2008). Use of visual search in the assessment of pattern-related visual stress (PRVS) and its alleviation by colored filters. *Investigative ophthalmology & visual science, 49*(9), 4210-4218.

Al-Rasheed, A. S. (2015). An experimental study of gender and cultural differences in hue preference. *Frontiers in psychology, 6*(30).

American Academy of Pediatrics. (2009). *Learning disabilities, dyslexia, and vision. Pediatrics, 124*(2), 837-844.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

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.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General, 139*(2), 299.

Atwood, M. (2014) Stone Mattress: Nine Tales. Bloomsbury Publishing

Baayen, R. H., Piepenbrock R., and van Rijn, H. (1993) The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania.

Badcock, D., & Lovegrove, W. (1981). The effects of contrast, stimulus duration, and spatial frequency on visible persistence in normal and specifically disabled readers. Journal of Experimental Psychology: *Human Perception and Performance, 7*(3), 495.

Beecher, H. K. (1955). The powerful placebo. *Journal of the American Medical Association, 159*(17), 1602-1606.

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. JOSA A, 19(6), 1096-1106.

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.

Brainard, D. H., (1997). The psychophysics toolbox. *Spatial vision, 10*, 433-436.

Brainard, D. H., Pelli, D. G., & Robson, T. (2002). Display characterization. *Signal Process, 80*, 2-67.

Brown, W. E., Eliez, S., Menon, V., Rumsey, J. M., White, C. D., & Reiss, A. L. (2001). Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology, 56*(6), 781-783.

Bub, D. N., Masson, M. E., & Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychological Science, 17*(4), 351-357.

Buchsbaum, G., & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. Proceedings of the Royal society of London. Series B. *Biological sciences, 220*(1218), 89-113.

Cardona, G., Borràs, R., Peris, E., & Castañé, M. (2010). A placebo-controlled trial of tinted lenses in adolescents with good and poor academic performance: reading accuracy and speed. *Journal of optometry, 3*(2), 94-101.

Carver, R. P. (1990). Reading rate: A review of research and theory. New York: Academic Press.

Chouinard, B. D., Zhou, C. I., Hrybouski, S., Kim, E. S., & Cummine, J. (2012). A functional neuroimaging case study of Meares–Irlen syndrome/visual stress (MISViS). *Brain topography, 25*(3), 293-307.

Chung, S. T. (2002). The effect of letter spacing on reading speed in central and peripheral vision. *Investigative Ophthalmology & Visual Science*, *43*(4), 1270-1276.

Clifton Jr, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *Journal of Memory and Language, 86*, 1-19.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological review, 97*(3), 332.

Conlon, E. G., Lovegrove, W. J., Chekaluk, E., & Pattison, P. E. (1999). Measuring visual discomfort. *Visual Cognition, 6*(6), 637-663.

Conway, M. L., Evans, B. J., Evans, J. C., & Suttle, C. M. (2016). Does gender influence colour choice in the treatment of visual stress? *PloS one, 11*(9) e0163326.

Cornelissen, P. L., Hansen, P. C., Gilchrist, I. D., Cormack, F., Essex, J., & Frankish, C. (1998a). Coherent motion detection and letter position encoding. *Vision Research, 38*(14), 2181-2191.

Cornelissen, P. L., Hansen, P. C., Hutton, J. L., Evangelinou, V., & Stein, J. F. (1998b). Magnocellular visual function and children's single word reading. *Vision research, 38(*3), 471-482.

Cornelissen, P. L, Richardson, A. R., Mason, A., Fowler, M. S., & Stein, J. F. (1994) Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics & controls. *Vision Research, 315*(10), 1483–1494.

Cotton, M. M., & Evans, K. M. (1990). A review of the use of Irlen (tinted) lenses. Clinical & Experimental *Ophthalmology, 18*(3), 307-312.

Danelli, L., Berlingeri, M., Bottini, G., Borghese, N. A., Lucchese, M., Sberna, M., ... & Paulesu, E. (2017). How many deficits in the same dyslexic brains? A behavioural and fMRI assessment of comorbidity in adult dyslexics. *Cortex, 97*, 125-142.

Dillon, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*(10), 1297-1326.

Dodd, S., Dean, O. M., Vian, J. and Berk, M. (2017) A Review of the Theoretical and Biological Understanding of the Nocebo and Placebo Phenomena. *Clinical Therapeutics, 39*(3), 469-476.

Elliott, J. G., & Gibbs, S. (2008). Does dyslexia exist? *Journal of Philosophy of Education, 42*(3-4), 475-491.

Elliot, D. B., & Wood, J. M. (2017). Coloured filters show gender differences and poor repeatability. *Ophthalmic and Physiological Optics, 37*(6), 635-639.

Evans, B. J. W. (2005). The need for optometric investigation in suspected Meares–Irlen syndrome or visual stress. *Ophthalmic and Physiological Optics, 25*(4), 363-370.

Evans, B. J. W., & Allen, P. M. (2016). A systematic review of controlled trials on visual stress using Intuitive Overlays or the Intuitive Colorimeter. *Journal of optometry, 9*(4), 205-218.

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.

Fawcett, A. J., & Nicolson, R. I. (1999). Performance of dyslexic children on cerebellar and cognitive tests. *Journal of motor behavior, 31*(1), 68-78.

Fawcett, A. J., Nicolson, R. I., & Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia, 46(*1), 259-283.

Fernandez, D., & Wilkins, A. J. (2008). Uncomfortable images in art and nature. Perception, 37(7), 1098-1113.

Fisher, C., Chekaluk, E., & Irwin, J. (2015). Impaired driving performance as evidence of a magnocellular deficit in dyslexia and visual stress. *Dyslexia, 21*(4), 350-360.

Fitzmaurice, C. (2016, September 30). Review slated for bias. Retrieved from https://www.opticianonline.net/opinion/letter-coloured-overlays-review-slated-for-bias

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: absence of inhibitory neighborhood effects in lexical decision and semantic categorization. Journal of Experimental Psychology: Learning, *Memory, and Cognition, 22*(3), 696.

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.

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLOS one, 9*(8), 1-12.

Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., & Helfer, S. G. (2005). Goal activation, expectations, and the placebo effect. *Journal of personality and social psychology, 89*(2), 143.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience, 4(*7), 563.

Gibbs, R., Davies, G., & Chou, S. (2016). A systematic review on factors affecting the likelihood of change blindness. *Crime Psychology Review, 2(*1), 1-21.

Giddings, E. H., & Carmean, S. L. (1989). Reduced brightness contrast as a reading aid. *Perceptual and motor skills, 69*(2), 383-386.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. Journal of Experimental Psychology: *Human Perception and Performance, 10(*5), 640.

Goodbourn, P. T., Bosten, J. M., Hogg, R. E., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2012). Do different 'magnocellular tasks' probe the same neural substrate?. Proceedings of the Royal Society of London B: Biological Sciences.

Gotzsche, P. C. (1994). Is there logic in the placebo?. The Lancet, 344(8927), 925-926.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education, 7(*1), 6-10.

Gould, J. D., & Grischkowsky, N. (1986). Does visual angle of a line of characters affect reading speed?. *Human Factors, 28*(2), 165-173.

Griffiths, P. G., Taylor, R. H., x , L. M., & Barrett, B. T. (2016). The effect of coloured overlays and lenses on reading: a systematic review of the literature. *Ophthalmic and Physiological Optics, 36*(5), 519-544.

Hall, R., Ray, N., Harries, P., & Stein, J. F. (2013). A comparison of two-coloured filter systems for treating visual reading difficulties. *Disability and rehabilitation, 35*(26), 2221-2226.

Haour, F. (2005). Mechanisms of the placebo effect and of conditioning. *Neuroimmunomodulation, 12*(4), 195-200.

Harries, P., Hall, R., Ray, N., & Stein, J. F. (2015). Using coloured filters to reduce the symptoms of visual stress in children with reading delay. *Scandinavian journal of occupational therapy, 22*(2), 153-160.

Harris, D., & MacRow-Hill, S. J. (1999). Application of ChromaGen haploscopic lenses to patients with dyslexia: a double-masked, placebo-controlled trial. *Journal of the American Optometric Association, 70*(10), 629-640.

He, Y., & Legge, G. E. (2017). Linking crowding, visual span, and reading. *Journal of vision, 17*(11), 11-11.

Henderson, L. M., Taylor, R. H., Barrett, B., & Griffiths, P. G. (2014). Treating reading difficulties with colour. *BMJ (British Medical Journal),* 349

Henderson, L. M., Tsogka, N., & Snowling, M. J. (2013). Questioning the benefits that coloured overlays can have for reading in students with and without dyslexia. *Journal of Research in Special Educational Needs, 13*(1), 57-65.

Hilbert, S., Nakagawa, T. T., Bindl, M., & Bühner, M. (2014). The spatial Stroop effect: A comparison of color-word and position-word interference. Psychonomic bulletin & review, 21(6), 1509-1515.

Hogben, J. H., Pratt, C., Dedman, K., & Clark, C. D. (1996). Blurring the image does not help disabled readers. *Vision Research,* 36(10), 1503-1507.

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.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics, 6*(2) 65-70.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing, 2*(2), 127-160.

Howell, E., & Stanley, G. (1988). Colour and learning disability. Clinical and *Experimental Optometry, 71*(2), 66-71.

Huang, J., Zong, X., Wilkins, A. J., Jenkins, B., Bozoki, A., & Cao, Y. (2011). fMRI evidence that precision ophthalmic tints reduce cortical hyperactivation in migraine. *Cephalalgia: an international journal of headache, 31*(8), 925–36.

Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current biology, 17*(16), 623-625.

Hurlbert, A. C., & Ling, Y. (2017). Understanding colour perception and preference. In Colour Design (pp. 169-192). Woodhead Publishing.

Iles, J., Walsh, V., & Richardson, A. R. (2000). Visual search performance in dyslexia. *Dyslexia, 6*(3), 163-177.

Irlen, H. (1983). Successful treatment of learning disabilities. Paper presented at the 91st annual Convention of the American Psychological Association, Anaheim, CA, USA.

Irlen, H. (1987). Irlen Differential Perceptual Schedule. Long Beach, California: Perceptual Development Corporation.

Irlen, H. (1991a). Reading by the colors. Avery Publishing Group.

Irlen, H. (1991b). Scotopic Sensitivity Syndrome: screening manual. Long Beach, CA: Perceptual Development Corporation.

Irlen, H. (1994). Scotopic sensitivity/Irlen syndrome: Hypothesis and explanation of the syndrome. Journal of Behavioral Optometry, 5(62), 65-66.

Irlen, H. (2010). The Irlen revolution: A guide to changing your perception and your life. New York: SquareOne Publishers.

Irlen, H. (2011). Factors that would account for lack of immediate change. Retrieved from http://pediatrics.aappublications.org/content/128/4/e932.comments#rebuttal-factors-that-would-account-for-lack-of-immediate-changes

Irlen, H. (2015). Sports, concussions and getting back in the game of life. Square One Publishers.

Irlen Institute (2012). "Irlen: Who do we help". Retrieved from http://www.irlenuk.com/who-do-we-help-overview.htm

Irlen Institute (2017a) "Where the science of colour changes lives". Retrieved from http://www.irlen.com

Irlen Institute (2017b) "About the Irlen Institute". Retrieved from http://irlen.com/about-irlen-institute/

Irlen Institute (2017c). "Irlen and the Brain". Retrieved from http://www.irlen.com/brain-research/

Jeanes, R. J., Busby, A. E., Martin, J., Lewis, E., Stevenson, N., Pointon, D., & Wilkins, A. J. (1997). Prolonged use of coloured overlays for classroom reading. *British Journal of Psychology, 88*(4), 541-548.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., & Passingham, R. E. (1994). Motor sequence learning: a study with positron emission tomography. *Journal of Neuroscience, 14*(6), 3775-3790.

Kapoula, Z., Lê, T. T., Bonnet, A., Bourtoire, P., Demule, E., Fauvel, C., ... & Yang, Q. (2010). Poor Stroop performances in 15-year-old dyslexic teenagers. *Experimental brain research, 203*(2), 419-425.

Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., … & Lembo A. J. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ, 336*(7651), 999-1003.

Kim, J. H., Seo, H. J., Ha, S. G., & Kim, S. H. (2015). Functional magnetic resonance imaging findings in Meares-Irlen syndrome: A pilot study. *Korean Journal of Ophthalmology, 29*(2), 121-125.

Klein, S. A., Hu, Q. J., & Carney, T. (1996). The adjacent pixel nonlinearity: problems and solutions. *Vision research, 36*(19), 3167-3181.

Knoblauch, K., Arditi, A., & Szlyk, J. (1991). Effects of chromatic and luminance contrast on reading. *JOSA A, 8*(2), 428-439.

Kriss, I., & Evans, B. J. (2005). The relationship between dyslexia and Meares-Irlen Syndrome. *Journal of Research in Reading, 28*(3), 350-364.

Kruk, R., Sumbler, K., & Willows, D. (2008). Visual processing characteristics of children with Meares–Irlen syndrome. *Ophthalmic and Physiological Optics, 28*(1), 35-46.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition, 16*(5), 400-410.

Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: New insights from an ideal-observer model of reading. *Vision Research, 42*(18), 2219-2234.

Legge, G. E., Parish, D. H., Luebker, A., & Wurm, L. H. (1990). Psychophysics of reading. XI. Comparing color contrast and luminance contrast. *JOSA A, 7*(10), 2002-2010.

Legge, G. E., Rubin, G. S., & Luebker, A. (1987). Psychophysics of reading—V. The role of contrast in normal vision. *Vision research, 27*(7), 1165-1177.

Leonard, C. M., Eckert, M. A., Lombardino, L. J., Oakland, T., Kranzler, J., Mohr, C. M., ... & Freeman, A. (2001). Anatomical risk factors for phonological dyslexia. *Cerebral Cortex, 11*(2), 148-157.

Le Jan, G., Le Bouquin-Jeannès, R., Costet, N., Trolès, N., Scalart, P., Pichancourt, D., ... & Gombert, J. E. (2011). Multivariate predictive model for dyslexia diagnosis. *Annals of dyslexia, 61(*1), 1-20.

Levien, R. L. (2003, December). Resolution-dependence of perceived contrast of textures. In Color Imaging IX: Processing, Hardcopy, and Applications (Vol. 5293, pp. 405-412). International Society for Optics and Photonics.

Lightstone, A., Lightstone, T., & Wilkins, A. (1999). Both coloured overlays and coloured lenses can improve reading fluency, but their optimal chromaticities differ. Ophthalmic and *Physiological Optics, 19*(4), 279-285.

Ling, Y., & Hurlbert, A. C. (2007, January). A new model for color preference: Universality and individuality. In Color and Imaging Conference (Vol. 2007, No. 1, pp. 8-11). Society for Imaging Science and Technology.

Ling, Y., Hurlbert, A., & Hurlbert, A. C. (2011). Age-dependence of colour preference in the UK population. *New directions in colour studies, 347.*

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*(18), 7943-7947.

LoBue, V., & DeLoache, J. S. (2011). Pretty in pink: The early development of gender-stereotyped colour preferences. *British Journal of Developmental Psychology, 29*(3), 656-667.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 13*(2), 123-146.

Lovegrove, W.J., Martin, F., Blackwood, M. & Badcock, D. (1980) Specific reading difficulty: Differences in contrast sensitivity as a function of spatial frequency. *Science, 210*, 439–440.

Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic bulletin & review, 2*(2), 174-207

Luo, M. R., Cui, G., & Rigg, B. (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Centre Français de la Couleur, 26*(5), 340-350.

MacLeod, D. I., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *JOSA, 69*(8), 1183-1186.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in cognitive sciences, 4*(10), 383-391.

Macpherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research, 84*(1), 24-62.

Mathsworks (2018, August 23). Record and Play Audio. Retrieved from https://www.mathworks.com/help/matlab/import_export/record-and-play-audio.html

McIntosh, R. D., & Ritchie, S. J. (2012). Rose-tinted? The use of coloured filters to tread reading difficulties. In Sergio Della Sala & Mike Anderson (Ed.), Neuroscience in Education: The good, the bad, and the ugly (pp. 230–243). Oxford University Press.

Meares, O. (1980). Figure/ground, brightness contrast, and reading disabilities. *Visible language, 14*(1), 13-29.

Metha, A. B., Vingrys, A. J., & Badcock, D. R. (1993). Calibration of a color monitor for visual psychophysics. *Behavior Research Methods, Instruments, & Computers, 25*(3), 371-383.

Mitchell, C., Mansfield, D., & Rautenbach, S. (2008). Coloured filters and reading accuracy, comprehension and rate: a placebo-controlled study. *Perceptual and motor skills, 106*(2), 517-532.

Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. *Annals of Internal medicine, 136*(6), 471-476.

Monger, L., Wilkins, A., & Allen, P. (2015). Identifying visual stress during a routine eye examination. *Journal of optometry, 8*(2), 140-145.

Neisser, U. (1964). Visual search. *Scientific American, 210*(6), 94-103.

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), 1729–35.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research?. *Cognition, 35*(2), 159-182.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (1995). Time estimation deficits in developmental dyslexia: evidence of cerebellar involvement. *Proceedings of the Royal Society of London Series B: Biological Sciences, 259*(1354), 43-47.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in neurosciences, 24*(9), 508-511.

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 Difficulties, 9*(2), 14-22.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 87.

Nosofsky, R. M., & Palmeri, T. J. (1997). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception & Psychophysics, 59*(7), 1027-1048.

O'Brien, B. A., Mansfield, J. S., & Legge, G. E. (2000). The effect of contrast on reading speed in dyslexia. *Vision research, 40*(14), 1921-1935.

Palef, S. R., & Olson, D. R. (1975). Spatial and verbal rivalry in a Stroop-like task. Canadian *Journal of Psychology 29(*3), 201.

Parra, J., Lopes da Silva, F. H., Stroink, H., & Kalitzin, S. (2007). Is colour modulation an independent factor in human visual photosensitivity? *Brain, 130*(6), 1679-1689.

Paradis, E., & Sutkin, G. (2017). Beyond a good story: from Hawthorne Effect to reactivity in health professions education research. *Medical education, 51*(1), 31-39.

Peli, E. (1990). Contrast in complex images. *JOSA A, 7*(10), 2032-2040.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of vision, 7*(2), 20-20.

Podell, S. M. (1990). An overview of the Irlen method. *Journal of Behavioral Optometry, 1*(7), 171-173.

Poffenberger, A. T. (1925). Psychology in advertising. New York, NY, US: McGraw-Hill Book Company.

Preston, K., Schwankl, H. P., & Tinker, M. A. (1932). The effect of variations in color of print and background on legibility. *The Journal of General Psychology, 6*(2), 459-461.

Protopapas, A., Vlahou, E. L., Moirou, D., & Ziaka, L. (2014). Word reading practice reduces Stroop interference in children. *Acta psychologica, 148*, 204-208.

Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology, 54*(3), 251-282.

Rae, C., Karmiloff-Smith, A., Lee, M. A., Dixon, R. M., Grant, J., Blamire, A. M., ... & Radda, G. K. (1998). Brain biochemistry in Williams syndrome Evidence for a role of the cerebellum in cognition? *Neurology, 51*(1), 33-40.

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current opinion in neurobiology, 13*(2), 212-218.

Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive neuropsychology, 29*(1-2), 104-122.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain, 126*(4), 841-865

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*(1), 283-293.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of experimental child psychology*, *41*(2), 211-236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological bulletin, 124(3), 372.

Rayner, K., & Juhasz, B. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology, 16*(1-2), 340-352.

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. Wiley Interdisciplinary Reviews: *Cognitive Science, 1*(6), 787-799.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and aging, 21*(3), 448.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological science, 8*(5), 368-373.

Ritchie, S. J., Della Sala, S., & McIntosh, R. D. (2011). Irlen colored overlays do not alleviate reading difficulties. *Pediatrics, 128*(4), 932–8.

Ritchie, S. J., Della Sala, S., & McIntosh, R. D. (2012). Irlen Colored Filters in the Classroom: A 1-Year Follow-Up. Mind, *Brain, and Education, 6*(2), 74-80.

Robinson, G. L., & Conway, R. N. (1994). Irlen filters and reading strategies: effect of coloured filters on reading achievement, specific reading strategies, and perception of ability. *Perceptual and motor skills, 79*(1), 467-483.

Robinson, G. L., & Conway, R. N. (2000). Irlen lenses & adults a small scale study of reading speed, accuracy, comprehension & self-image. *Australian Journal of Learning Difficulties, 5*(1), 4-13.

Robinson, G. L., & Foreman, P. J. (1999a). Scotopic sensitivity/Irlen syndrome and the use of coloured filters: a long-term placebo-controlled study of reading strategies using analysis of miscue. *Perceptual and motor skills, 88*(1), 35-52.

Robinson, G. L., & Foreman, P. J. (1999b). Scotopic sensitivity/Irlen syndrome and the use of coloured filters: a long-term placebo controlled and masked study of reading achievement and perception of ability. *Perceptual and motor skills, 89*(1), 83-113.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychological review, 110*(1), 88.

Rose, J. P., Geers, A. L., Rasinski, H. M., & Fowler, S. L. (2012). Choice and placebo expectation effects in the context of pain analgesia. *Journal of behavioral medicine, 35*(4), 462-470.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development,* 247-253.

Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J. F., Bricout, L., ... & George, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental psychology, 52*(10), 1503.

Scheiman, M., Blaskey, P., Ciner, E. B., Gallaway, M., Parisi, M., Pollack, K., & Selznick, R. (1990). Vision characteristics of individuals identified as Irlen Filter candidates. *Journal of the American Optometric Association, 61(*8), 600-605.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*(6), 1270-1281.

Schminder, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution. *European Research Journal of Methods for the Behavioral and Social Sciences, 6*, 147-151.

Shepherd, A.J. (1997). A vector model of colour contrast in a cone excitation colour space. *Perception, 26,* 455-470.

Shepherd, A.J. (1999). Remodelling colour contrast: implications for visual processing and colour representation. *Vision Research, 39*, 1329-1345.

Shor, R. E. (1970). The processing of conceptual information on spatial directions from pictorial and linguistic symbols. Acta Psychologica, 32, 346-365.

Shor, R. E., Hatch, R. P., Hudson, L. J., Landrigan, D. T., & Shaffer, H. J. (1972). Effect of practice on a Stroop-like spatial directions task. *Journal of Experimental Psychology, 94(*2), 168.

Simmers, A. J., Bex, P. J., Smith, F. K., & Wilkins, A. J. (2001). Spatiotemporal visual function in tinted lens wearers. *Investigative ophthalmology & visual science, 42*(3), 879-884.

Simmers, A. J., Gray, L. S., & Wilkins, A. J. (2001). The influence of tinted lenses upon ocular accommodation. *Vision Research, 41*(9), 1229-1238.

Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception, 29(*10), 1143-1154.

Singleton, C., & Henderson, L. (2007). Computerised screening for visual stress in reading. *Journal of Research in Reading, 30*(3), 316–331.

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). The magnocellular deficit theory of dyslexia: the evidence from contrast sensitivity. *Vision research, 40*(1), 111-127.

Skottun, B. C., & Skoyles, J. (2007). Yellow filters, magnocellular responses, and reading. *International Journal of Neuroscience, 117(*2), 287-293.

Smith, L., & Wilkins, A. (2007). How many colours are necessary to increase the reading speed of children with visual stress? A comparison of two systems. *Journal of Research in Reading, 30*(3), 332–343.

Snowling, M. J. (1996). The Graded Nonword Reading Test. Bury St Edmunds. Suffolk: Thames Valley Test Company.

Snowling, M. J. (2001). From language to reading and dyslexia 1. *Dyslexia, 7*(1), 37-46.

Snowling, M., Stothard, J., & McLean, J. F. (1996). *Graded non-word reading test*. London: Thames Valley Test.

Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Reviews Neuroscience, 8*(4), 276.

Sorokowski, P., Sorokowska, A., & Witzel, C. (2014). Sex differences in color preferences transcend extreme differences in culture and ecology. *Psychonomic bulletin & review, 21*(5), 1195-1201.

Stein, J. F. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia, 7*(1), 12-36.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18(*6), 643.

Suttle, C. M., Barbur, J., & Conway, M. L. (2017). Coloured overlays and precision-tinted lenses: poor repeatability in a sample of adults diagnosed with visual stress. *Ophthalmic and Physiological Optics, 37*(4), 542-548.

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.

Thompson, J. J., Ritenbaugh, C., & Nichter, M. (2009). Reconsidering the placebo response from a broad anthropological perspective. *Culture, Medicine, and Psychiatry, 33*(1), 112-152.

Tinker, M. A., & Paterson, D. G. (1931). Studies of typographical factors influencing speed of reading. VII. Variations in color of print and background. *Journal of Applied Psychology, 15*(5), 471.

Tyrrell, R., Holland, K., Dennis, D., & Wilkins, A. J. (1995). Coloured overlays, visual discomfort, visual search and classroom reading. *Journal of Research in Reading, 18*(1), 10–23.

Vellutino, F. R. (1979). Dyslexia: Theory and research. MIT Press.

Veszeli, J., & Shepherd, A. J. (2019). A comparison of the effects of the colour and size of coloured overlays on young children's reading. *Vision research, 156,* 73-83.

Vinegard, M. (1994). A revised adult dyslexia check list. *Educare. 48*, 21-23.

Virzi, R. A., & Egeth, H. E. (1985). Toward a translational model of Stroop interference. *Memory & cognition, 13*(4), 304-319.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research, 37*(3), 475-494.

Whiteley, H., & Smith, C. (2001). The use of tinted lenses to alleviate reading difficulties. *Journal of Research in Reading, 24*(1), 30-40.

Williams, M. C., & LeCluyse, K. (1990). Perceptual consequences of a temporal processing deficit in reading disabled children. *Journal of the American Optometric Association, 61*(2), 111-121.

Wilkins, A. J. (1994). Overlays for classroom and optometric use. *Ophthalmic and Physiological Optics, 14*(1), 97–99.

Wilkins, A. J. (1995). *Visual Stress.* Oxford University Press.

Wilkins, A. J. (1997). A system of precision ophthalmic tinting and its role in the treatment of visual stress. *John Dalton's colour vision legacy*, 695-708.

Wilkins, A. J. (2001). A System for Precision Ophthalmic Tinting Manual for the Intuitive Colorimeter Mk.2 and Precision Tints. 4th Edition. Retrieved from: https://www1.essex.ac.uk/psychology/overlays/mk2ukman.htm

Wilkins, A. J. (2002). Coloured overlays and their effects on reading speed: a review. *Ophthalmic and Physiological Optics, 22*(5), 448-454.

Wilkins, A. J. (2003). *Reading through colour*. Chichester: John Wiley & Sons.

Wilkins, A. J., Allen, P. M., Monger, L. J., & Gilchrist, J. M. (2016). Visual stress and dyslexia for the practicing optometrist. *Optometry in Practice, 17(*2).

Wilkins, A. J., Baker, A., Amin, D., Smith, S., Bradford, J., Zaiwalla, Z., ... & Fish, D. (1999). Treatment of photosensitive epilepsy using coloured glasses. *Seizure-European Journal of Epilepsy, 8(*8), 444-449.

Wilkins, A. J., Evans, B. J. W., Brown, J. A., Busby, A. E., Wingfield, A. E., Jeanes, R. J., & Bald, J. (1994). Double-masked placebo-controlled trial of precision spectral filters in children who use coloured overlays. *Ophthalmic and Physiological Optics, 14*(4), 365-370.

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*(6), 491–497.

Wilkins, A. J., & Lewis, E. (1999). Coloured overlays, text, and texture. *Perception, 28*(5), 641–50.

Wilkins, A. J., Lewis, E., Smith, F., Rowland, E., & Tweedie, W. (2001). Coloured overlays and their benefit for reading. *Journal of Research in Reading, 24*(1), 41–64.

Wilkins, A. J., & Neary, C. (1991). Some visual, optometric and perceptual effects of coloured glasses. *Ophthalmic and Physiological Optics, 11*(2), 163-171.

Wilkins, A. J., & Nimmo-Smith, I. (1984). On the reduction of eye-strain when reading. *Ophthalmic and Physiological Optics, 4*(1), 53-59.

Wilkins, A. J., & Nimmo-Smith, M. I. (1987). The clarity and comfort of printed text. *Ergonomics, 30*(12), 1705-1720.

Wilkins, A. J., Nimmo-Smith, I., & Jansons, J. E. (1992). Colorimeter for the intuitive manipulation of hue and saturation and its role in the study of perceptual distortion. *Ophthalmic and Physiological Optics, 12*(3), 381-385.

Wilkins, A. J., Nimmo-Smith, I., Tait, A., McManus, C., Sala, S. D., Tilley, A., … & 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. (2001). A colorizer for use in determining an optimal ophthalmic tint. *Centre Français de la Couleur, 26*(3), 246-253.

Wilkins, A. J., Sihra, N., & Myers, A. (2005). Increasing reading speed by using colours: Issues concerning reliability and specificity, and their theoretical and practical implications. *Perception, 34*(1), 109–120.

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.

Witt, K. (2007). *CIE color difference metrics. Colorimetry: understanding the CIE system*, 79-100. John Wiley & Sons.

Wolfe, J. M. (2010). Visual search. *Current biology, 20*(8), 346-349.

Wright, H. R., & Lack, L. C. (2001). Effect of light wavelength on suppression and phase delay of the melatonin rhythm. *Chronobiology international, 18(*5), 801-808.

Wu, J. H., & Yuan, Y. (2003). Improving searching and reading performance: the effect of highlighting and text color coding. *Information & Management, 40*(7), 617-637.

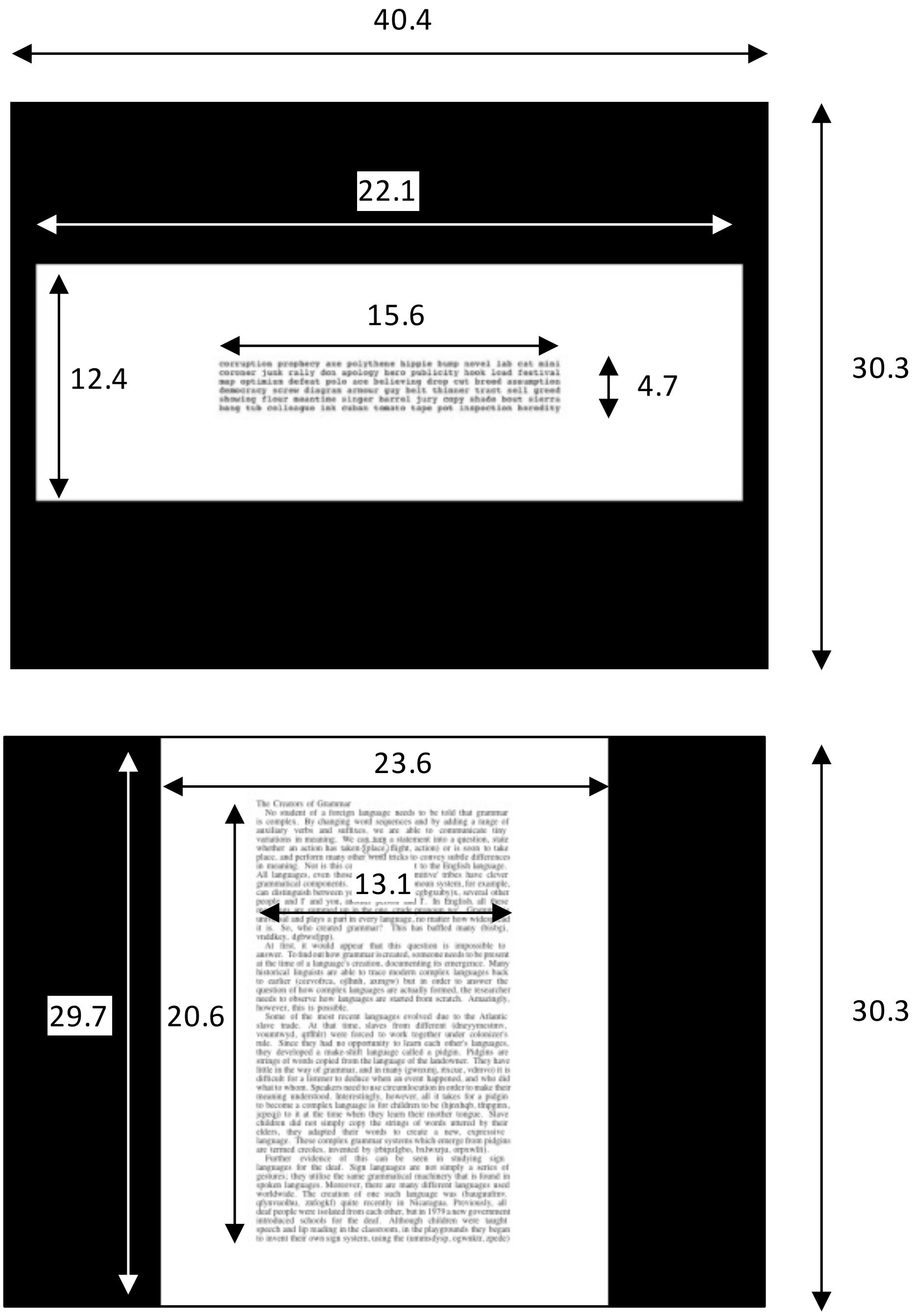
Xiao, Y., Wang, Y. I., & Felleman, D. J. (2003). A spatially organized representation of colour in macaque cortical area V2. *Nature, 421*(6922), 535.

Yu, D., Cheung, S. H., Legge, G. E., & Chung, S. T. (2007). Effect of letter spacing on visual span and reading speed. Journal of vision, 7(2), 2-2.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., ... & Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109*(28), 11455-11459.

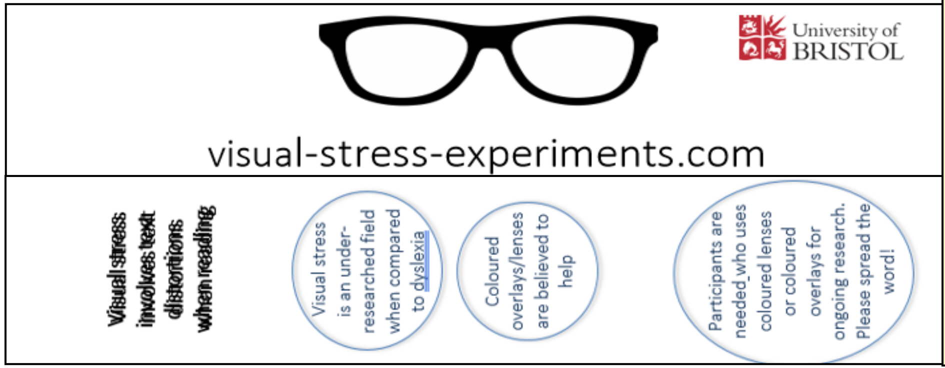
Appendix I. Dimensions of text stimuli

Experiment 1 (top) and other experiments (bottom). Scale is 1:25.



Appendix II. Recruitment methods

Top left: animated screen advert. Bottom left: bookmark to be used as fliers. Right: Website.



Appendix III. Formation of noun and animal word bank (Experiment 1 and 2)

Word were retrieved from lexical-databases of 3-11 letter words. The noun word bank contained medium to high frequency words (100-1000 lemma frequency) and the animal word bank included a wider frequency band of 22 -1000 lemma frequency so that longer animal words could be used. Words that were not easily identified as animals or non-animals were removed, for example foreign, repeated and offensive words, proper nouns, birds, fish and homonyms, and words with some associated with animals (e.g. beast, brute). A familiarity screening was undertaken to ensure that the animal words were immediately recognisable as animals; 34 students were asked 'which words people, in general, would not immediately recognise as animals?'. Nine or more participants selected 18 animal words, which left 63 remaining animal words to be used across both tasks

ape	frog	hyena	donkey	octopus
bat	goat	mouse	ferret	panther
cat	hare	otter	jaguar	antelope
cow	lamb	panda	kitten	elephant
dog	lion	sheep	monkey	hedgehog
fox	mole	snail	rabbit	kangaroo
pig	pony	snake	weasel	reindeer
rat	toad	tiger	buffalo	squirrel
bear	wolf	whale	gazelle	alligator
boar	worm	zebra	giraffe	crocodile
bull	camel	baboon	gorilla	chimpanzee
deer	horse	badger	hamster	rhinoceros
		beaver	leopard	hippopotamus

Appendix IV. Questionnaires (Experiment 5 -8)

Novel measures of visual stress were designed to be quick to administer and to elicit the number and type of visual stress symptoms when reading. The first was based upon Wilkins' (2003) measure. The second contained pseudorandomised positive and negative sentences in relation to visual stress symptoms (Henderson et al., 2013; Conlon, Lovegrove, Chekaluk & Pattison, 1999).

Visual stress measure 1

1. Does reading become harder the longer you read? Yes
   No
2. Do words in a book ever... (tick those which apply)
   - go blurred
   - go smaller/bigger
   - jump around
   - fade or disappear
   - none of the above

Visual stress measure 2

Tick the sentence which applies to you

1. the difficulty of reading stays the same as I read
   reading becomes more difficult the longer I read
2. I often lose my place in the text
   I don't often lose my place in the text
3. I usually understand what I read when I read it first time
   I often have to reread the same text to understand what I have read
4. letters or spaces begin to move
   letters and spaces remain still
5. letters remain clear
   letters become blurred or fuzzy
6. words are too close together
   words are far enough apart
7. the page is a comfortable brightness
   the page feels too bright or not bright enough

Questions about the perceived effectiveness of the chosen hue

1a. Do you believe that one specific colour is necessary to increase you reading speed?
1b. If not, what other colours improve your reading speed?
2a. Have you tried reading on grey, or reducing contrast, instead of a colour?
2b. If yes, how does this compare with reading on a colour?

Appendix V. Elicitation of whether participants noticed the colour change

Did (0) /Did not notice(1)	**Did you find any parts easier or faster to read than others? Did you notice any difference in the screen?**
0	No. I checked when there were two possible answers. Didn't notice.
1	When the words are at the end, it changed colour. It was easier to read when it was white.
0	Didn't notice, didn't notice at all.
0	It was easier to pick a word when you had all the info prior. Didn't notice any change in the appearance.
0	Didn't notice change in reading speed. Noticed a change in the screen but couldn't specify what it was.
0	no. Didn't notice anything
0	not on the reading task, no difference in reading speed, when it was more yellow I thought it was easier. I didn't notice a change in the screen when I was reading but did no the t and l task
1	When it was whiter, it was easier to see the Ts. Didn't notice as much with reading
1	Can't remember any bits in specific, but remember reading a couple of bits twice. Was the colour background was changing?.
0	Did think she was dyslexic as a child. Possible reading problems. Not noticably more difficult to read in part than others. Didn't notice anything about the appearance of the text when reading.
0	Didn't notice.
1	Found it easier when searching for Ls when the white wasn't so harsh.
1	I found it easier when it wasn't too dark, and when it wasn't too bright.
1	No. I noticed that I was reading quicker but couldn't tell you why. The screen seemed to go more peachy and back again.
0	Didn't find reading quicker or slower. Didn't notice any change in the apearance of the screen. Completely surprised.
0	no. I thought it got duller/ more saturated in the ts and ls. (when pushed) I noticed that it got bright at some point.For the reading, I found it harder to read, but didn't know why.
1	When it was changing it was a bit weird,
1	Yes. I found it easier to read when it was greenish whitish
1	I noticed the colour change throughout

1	Noticed colour change, particularly in the second part.
1	I found some parts easier and faster to read, and a change of brighness in the screen.
0	Didn't find any parts easier to read than others. I wasn't reading faster in some areas than others. Didn't notice any change in the colour of the screen when reading.
1	When it was dark or very bright, I found it more difficult to read. I could read easier towards the whiter end of the spectrum.
1	No. No. I think it was changing colour on both tasks.
0	no
1	Noticed colour change
0	didn't notice reading change but did notice word search change - startling bright and difficult to find target
0	can't remember when and where, there were bits which were quite smooth, and other bits where the screen was shaking around
0	When it wasn't in the middle of the paragraph, In the middle I was speeding up, at the end I was slowing down, as the words got more difficult to pick.
1	Found reading easier on the colour'
1	felt it was harder, and noticed the colour change - not sure if it effected speed of reading.
0	found border distracting
1	Noticed colour change- when it was reaching white..- more profound on second one
0	Towards the end it became really blurred, page 3 found difficult
1	When the vocabulary got a bit simpler. It seemed to get brighter at points, I lent in- I was trying to concentrate more.. Ts- dazed when it was white, and more comfortable when searching on a blue background. Glow on the outside of the Ts when it was white.
1	Noticed colour change
0	no
1	The colour changed, I found it too bright and that made me slow down. After a while, I got used to it.
1	When the screen was brighter and whiter I found the text more difficult to read- I had to squint to focus and slow down. I seem to miss things, I miss words.
1	When the screen went to orange, I was conscious of the colour change, it jarred me and I felt unsettled. It took me a while to refocus, as it was pale orange. It was difficult to read as it was becoming more orange, but once it was more orange, I found it easier. I was more interested by the time it went white. I noticed the colour change last time too.

1	I found it easier and read quicker in the middle. Noticed the colour change.
1	When the brightness was gone, I found it easier to read. I pulled away from the screen when it was bright.
1	It was painful when it was bright white, and that meant that I found it difficult to focus on the meaning of the sentences. When it was changing colour, it messed with my reading. I felt it went a bluey grey at the end
1	It was certainly easier to read when it was blue, when it changed it became harder. I think I was reading faster in the blue
1	I definitely found it easier to read when it was purple. I was covering parts of the text up when it was white but I could look at the whole thing when it was purple. It was much easier to understand - the text was just going into my brain better.
0	I picked up the flow and became more accustomed to the writing style. It got easier and more fluent as I got into it. I didn't notice any change in the screen when I was reading or on the visual search.
1	I found it easier to read when it was darker. I think I read at the beginning faster. I think it was darker. I had to squint when it was white because the words wobble.
1	I read at quite a constant rate. The colour of the screen was changing on both tasks.

Appendix VI. Non leading instructions applied to yellow filter condition (Chapter 5).

(All screen shots have been reduced to a ratio of 1:3.)

Text is sometimes presented on a yellow background.

You will complete two tasks on a both a yellow and white background.
Task 1 (5 mins): Read silently some text and choose words to complete the text.
Task 2 (1 min): Read aloud a passage of randomly ordered words as rapidly as possible.
In each task, part 1 will be conducted on a white background and part 2 on a yellow background.

Testing time is no longer than 20 minutes.
Press the right arrow key to begin practicing the first task.

Appendix VII. WRRT

the to see for is not up cat you my play dog look come and
to is for cat see look come dog and play up my not you the
is and for play to my cat dog see come look you the up not
cat dog and play not come to you my see for up look is the
see my dog look cat the you play for come up not to and is
is not play for look see come and you my dog to the up cat
come see to for is and look not cat my dog up play you the
dog the my for and play see is cat you to look up not come
the you up and dog is for see not to look my come play cat
and my play for the up look come dog see not cat is you to
my come you and for is not play dog cat see look the to up
my cat the and to play not dog is you come see look up for
you cat up see is dog for come not look my to and the play
not see and for up my come cat is dog play to look you the
dog is come to you cat see up look for the not and play my
come play the look up not my to and is cat dog you see for
cat the play my see and look is up not to dog come for you
not is play dog look my up the see for you and to cat come
see my the dog and up not look is you to cat play come for
look see and for cat dog the play up is come my to you not
for my to up and is come not cat see look play dog you the
and cat not look my the dog see to play is you come up for
cat see to play up you for look is not my come the and dog
play is my for come up to see not dog look the you cat and
cat play my to the you and for is not come look see dog up

Appendix VIII. Text stimuli of Experiment 7

Reading can be a form of enjoyment and relaxation for some people, but not for others. Why is this?

Our motivation to read is influenced by our parental figures and role-models. They decide, in the first instance, how much importance is placed on reading in the early stages of our reading development. By reading with us, they create a shared enjoyment so that certain stories will always give us a sense of nostalgia. If we have experienced this enjoyment of reading at an early age, we intrinsically feel motivated to learn to read independently. Why then do only some of us go onto become avid readers? What separates those who are drawn to reading from those who tend to avoid it?

Development of the ability to read English independently and fluently requires many skills. Teaching approaches have changed over the years. In the 1940s-1960s, the popular approach was the 'look and say' method where words were memorised. Now, phonics, where children decode words by sounds, is regarded to be the best method. However, it is important to consider that although 80% of words are phonically regular, 20% are not. These include some of the most common English words: for example, 'once, the, was, come'. These words are better learnt by sight. Also, talking about books can help comprehension, developing the skills of predicting, visualising, summarising and evaluating text. Additionally, access to a wide variety of texts can help motivation as individual interests can be pursued.

In the past, any person who was experiencing problems with reading was labelled dyslexic. Dyslexia was an umbrella term to diagnose a range of reading difficulties. However, over the years, the definition has become much narrower. This is because scientists have pinned down common traits among people with dyslexia. As many people with dyslexia experience difficulties with phonics, this has become one of the key diagnosing characteristics. In fact, if phonic intervention is applied at the right time, the likelihood of developing dyslexia, as it is currently defined, can be reduced. However, phonics is not the only problem that people with dyslexia experience, and the profile of problems associated with reading are different for each individual.

Another type of reading problem is called visual stress. This has been recorded since the 1980s. It is where people experience distortions and headaches when viewing text, despite having good eyesight. The appearance of text can distract or tire the reader, making it difficult to concentrate on the content of the text. Someone with visual stress might be able to read the words robotically, but not understand what they are reading because they are putting so much effort into making out the words. People with visual stress are said to tire as they read and find reading to be exhausting. They may be drawn to activities other than reading because they simply don't enjoy it. Very little is known about this phenomenon. Perhaps presenting the text in different ways can be beneficial to some readers. This is why we are conducting this experiment. We need you to tell us which page you find easiest to read.

Appendix IX. Handout used for Experiment 7

A1

Visual Reading Preferences Experiment: *Which text do you find easiest to read?*

Read these instructions. If you are happy to take part, please read and sign the first part of the consent form before taking part. If you are under 18, please ask your parental guardian to sign the first part of the parental consent form before taking part.

<u>Instructions</u>

1. First, the experimenter will adjust the height of the display for you.
2. You will then be asked to look at each page of text and decide the order in which you find them easiest to read.
3. The arrangement of rectangles below correspond to the arrangement of pages in front of you. You should indicate the order in which you find the text easiest to read by writing a number, from 1-5, on each rectangle below:

1	2	3	4	5
Easiest	→	→	→	Most difficult

4. Please complete the questionnaire:
   i. Do you consider yourself to be **Male/Female/other**
   ii. How old are you? ____________
   iii. Are you colour-blind? **Y/N**
   iv. Do you wear contact lenses or glasses? **Y/N** If so, are you wearing them? **Y/N**
   v. Have you ever used coloured filters for reading? **Y/N** If so, which colour?_________
   vi. Does reading become harder the longer you read? **Y/N**
   vii. Do words in a book ever: Go blurred? **Y/N** Jump around? **Y/N** Go smaller/bigger? **Y/N** Fade or disappear? **Y/N**
   viii. Are you diagnosed as dyslexic? _______________

5. Finally, are you happy for your data to be used? Your results will be anonymous, which means that your data and name are not attached. If you are happy for your data to be used, please sign the final consent (at the bottom of the consent form).

Thank you for participating.

Appendix X. Parental consent form used in Experiment 7

School of Experimental Psychology
Tel: 0117 928 8450
Jane.cowan@bristol.ac.uk

University of BRISTOL

PARENTAL CONSENT FORM

Visual reading preferences experiment

Please answer the following questions to the best of your knowledge

	YES	NO
HAVE YOU:		
• been given information explaining about the study?	☐	☐
• had an opportunity to ask questions and discuss this study?	☐	☐
• received satisfactory answers to all questions you asked?	☐	☐
• received enough information about the study for you to make a decision about your child's participation?	☐	☐
DO YOU UNDERSTAND:		
• that you are free to withdraw your consent for your child to take part		
• at any time	☐	☐
• without having to give a reason for withdrawing?	☐	☐
• that testing will stop if your child asks or appears uncomfortable?	☐	☐

If you are happy for your child to take part in the experiment, please sign here before taking part in the experiment.

I hereby fully and freely consent to my child's participation in this study

I understand the nature and purpose of the procedures involved in this study as communicated to me on the information sheet.

I understand that the investigation is designed to promote scientific knowledge and I agree that the University of Bristol can keep and use the data my family provide for research purposes only.

I understand that the data my family provide will be kept anonymous. This means that there are no links between my name or other identifying information and my study data.

Parent/Guardian signature: ________________________ Date: __________

Name in BLOCK Letters: __________________________________

Child's name ___________________________ Child's DoB: ____________

If you are happy for your child's data to be used, please sign here after taking part in the experiment.

Final consent

Having participated in this study

I agree to the University of Bristol keeping and processing the data I have provided during the course of this study. I understand that these data will be used only for the purpose(s) set out in the information sheet, and my consent is conditional upon the University complying with its duties and obligations under the Data Protection Act.

Parent/Guardian signature: ________________________ Date: __________

Name in BLOCK Letters: __________________________________

Child's name ___________________________ Child's DoB: ____________

If you have any concerns related to your participation in this study please direct them to the Faculty of Science Human Research Ethics Committee, via Liam McKervey, Research Ethics Co-ordinator (Tel: 0117 928 7841 email: Liam.McKervey@bristol.ac.uk

Appendix XI. Luminance and contrast measurements of text stimuli in Experiments 1 and 2

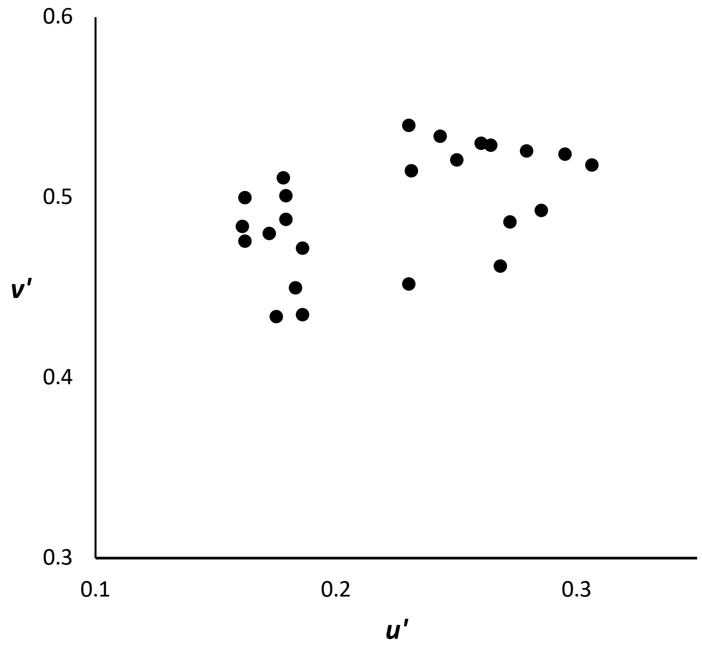
Inputted 'RGB' values were the luminance of the 'Minimum text pixel' and 'Text background' on a scale of 0-255; 'Text luminance' was the average luminance of the letters and background; 'Text background luminance' was the luminance of the screen background without the text; Weber, Michelson and RMS contrast were calculated using formula 1, 2 and 3. 'Minimum luminance' and 'Average text-pixel luminance' (the average luminance of the pixels that comprised the letters) were used to determine two values of Weber and Michelson contrast, as denoted by (M) and (A); note that Michelson text contrast is defined as Michelson contrast (A) here.

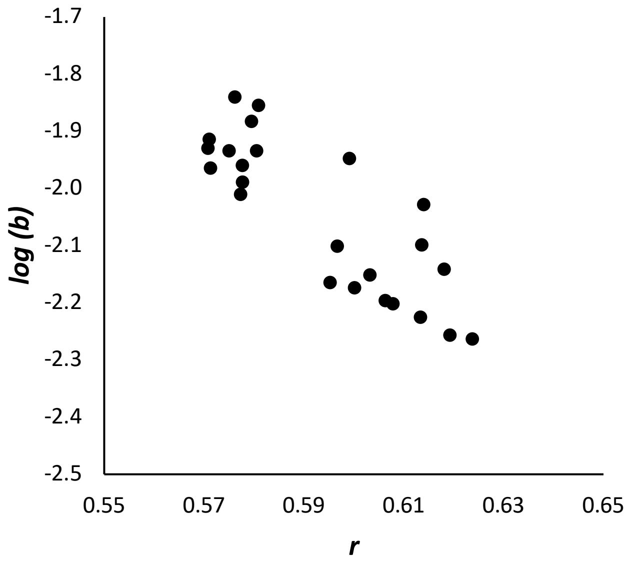
Values were calculated from a 59x59 pixel square of text (displayed in Figure 2.7). All stimuli had 1223 grey pixels and 2258 white pixels. For both experiments, text and background luminance of normal text and text with an overlay were obtained from the real world and applied to conditions 1 and 2 and the RMS contrast of condition 3 was applied to condition 2. In Experiment 1, the RMS contrast of condition 4 was applied to condition 1.

		RGB		cd/m^2				Contrast				
									Weber		Michelson	
Experiment (Condition)	Description	Minimum text pixel	Text background	Text Luminance	Text background luminance	Minimum luminance (M)	Average text-pixel luminance (A)	RMS	(M)	(A)	(M)	(A)
1 (1)	Normal text	50	255	105.772	119.625	23.729	80.196	0.163	0.802	0.330	0.669	0.197
1 (2)	Overlay text	44	115	51.621	54.135	20.923	46.979	0.059	0.614	0.132	0.442	0.071
1 (3)	Reduced contrast	180	227	105.260	106.527	84.541	102.920	0.059	0.206	0.034	0.115	0.017
1 (4)	Reduced luminance	0	228	91.575	106.995	0.340	63.106	0.163	0.997	0.472	0.994	0.258
2 (1)	Normal text	0	255	102.393	119.625	0.340	70.579	0.210	0.997	0.410	0.994	0.258
2 (2)	Overlay text	0	105	42.292	49.457	0.340	29.062	0.059	0.993	0.412	0.986	0.260
2 (3)	Reduced contrast	180	227	105.260	106.527	84.541	102.920	0.059	0.206	0.034	0.115	0.017

Appendix XII. Chromaticities of filters (Experiment 3)

CIE (1976) diagram (top) and MacLeod-Boynton (1979) diagram (bottom) displaying chromaticities of the filters used by the 24 participants in Experiment 3. Data is displayed for all 24 participants. Chromaticities and their saturations (CIE S_{uv}) are specified on the following page; Corresponding luminance and contrast measurements are displayed in Appendix XIV.



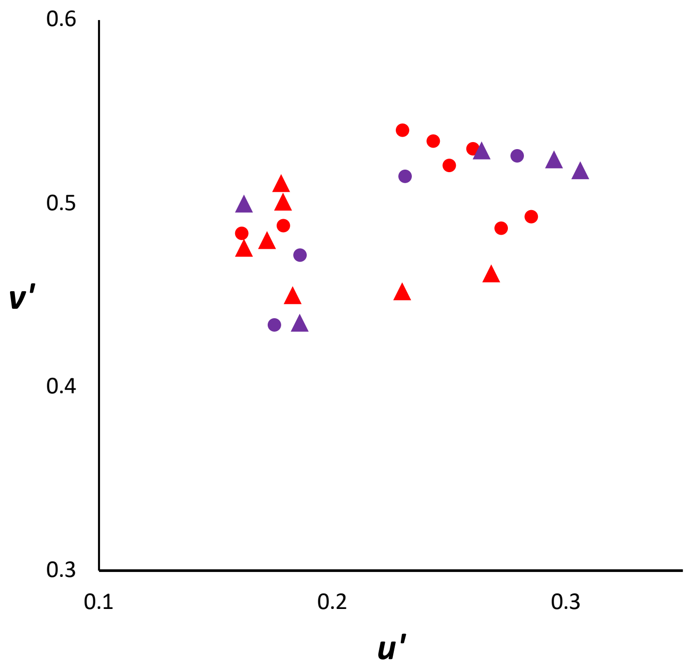


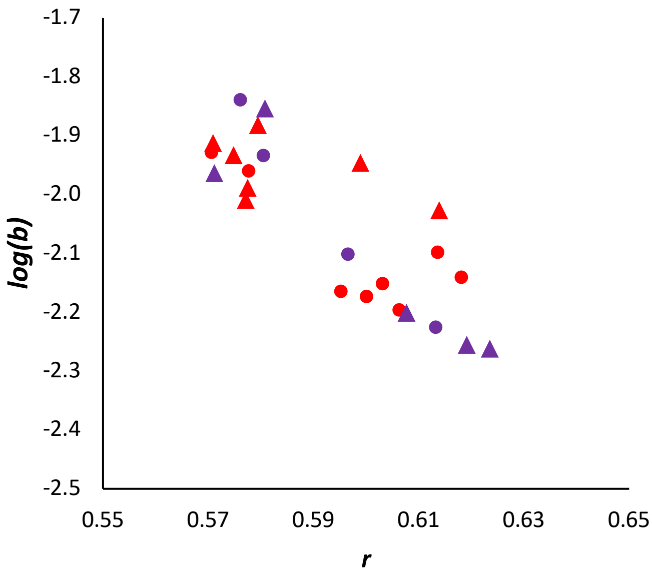
Appendix XII. continued

Saturation	CIE (1976)		MacLeod-Boyton (1979)		
S_{uv}	u'	v'	r	b	log b
0.059	0.179	0.501	0.578	0.010	-1.988
0.031	0.264	0.529	0.608	0.006	-2.201
0.059	0.178	0.511	0.577	0.010	-2.009
0.073	0.172	0.480	0.575	0.012	-1.933
0.081	0.161	0.484	0.571	0.012	-1.928
0.028	0.230	0.540	0.595	0.007	-2.164
0.059	0.295	0.524	0.619	0.006	-2.255
0.015	0.250	0.521	0.603	0.007	-2.151
0.076	0.162	0.500	0.571	0.011	-1.964
0.063	0.179	0.488	0.578	0.011	-1.960
0.065	0.186	0.472	0.580	0.012	-1.934
0.093	0.186	0.435	0.581	0.014	-1.854
0.044	0.272	0.487	0.614	0.008	-2.098
0.022	0.243	0.534	0.600	0.007	-2.173
0.062	0.230	0.452	0.599	0.011	-1.946
0.044	0.279	0.526	0.613	0.006	-2.225
0.029	0.260	0.530	0.606	0.006	-2.196
0.052	0.285	0.493	0.618	0.007	-2.140
0.069	0.306	0.518	0.624	0.005	-2.263
0.100	0.175	0.434	0.576	0.014	-1.839
0.083	0.183	0.450	0.579	0.013	-1.882
0.084	0.162	0.476	0.571	0.012	-1.913
0.060	0.268	0.462	0.614	0.009	-2.027
0.006	0.231	0.515	0.597	0.008	-2.101

Appendix XIII.Experiment 3 filter types

CIE (1976) diagram (top) and MacLeod-Boynton (1979) diagram (bottom) displaying chromaticities of the filters used by the 24 participants in Experiment 3 according to filter-type. Filters were Irlen overlays (▲), Irlen lenses (●), Intuitive overlays (▲) and Intuitive lenses (●).





Appendix XIV. Luminance and contrast measurements of the text stimuli in Experiments 3

Inputted 'RGB' values were 'Text background' and a minimum text pixel of 0, which outputted a Minimum pixel luminance (M) of 0.34 cd/m². 'Text background luminance' was the luminance of the screen background without the text. Weber and Michelson were calculated using Formulae 1 and 2 with minimum luminance defined by the 'Minimum pixel luminance', (M), and 'Average text-pixel luminance', (A).). Note that Michelson text contrast is defined as Michelson contrast (A) here.

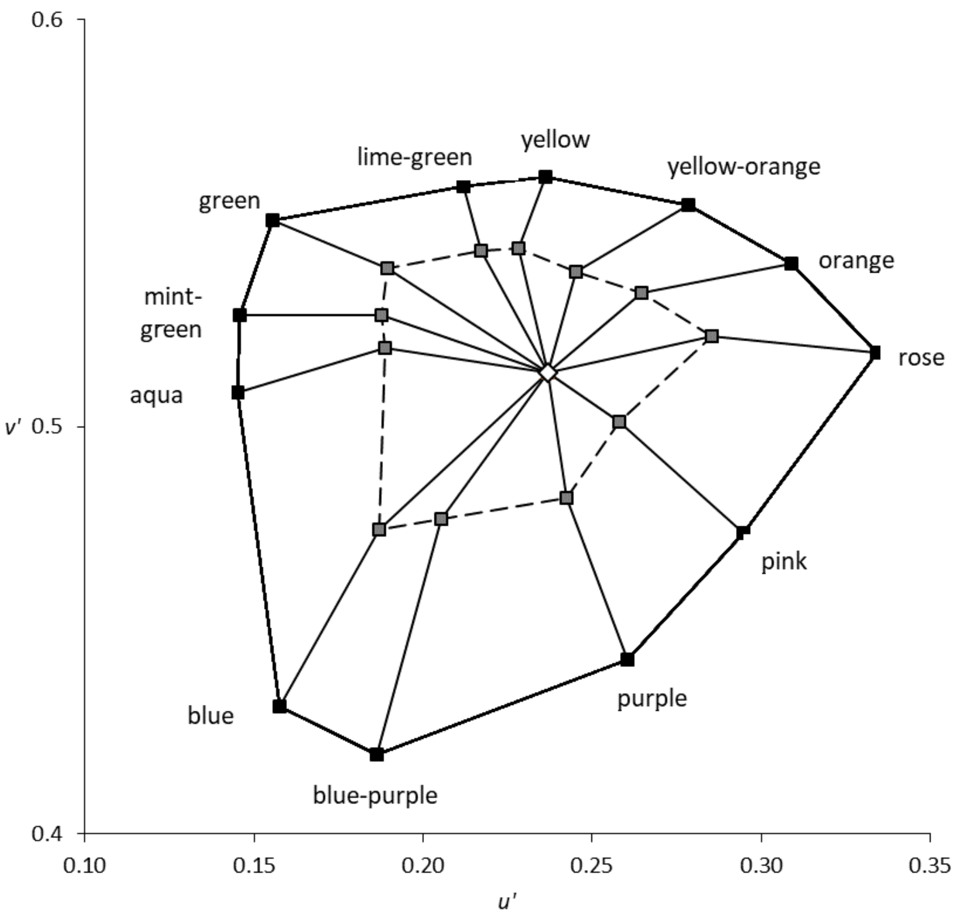
Contrast reduction from the 'no filter' condition (bold) was calculated with the Michelson formula using minimum luminance of the average text luminance (A). Data is displayed for all 24 participants. Corresponding chromaticities are displayed in Appendix XII.

Text background (Average RGB)	Text background luminance (cd/m²)	Average text-pixel luminance (A), (cd/m²)	Weber contrast using M	Weber contrast using A	Michelson contrast using M	Michelson contrast using A	Contrast reduction using Michelson with A
255	**119.625**	**70.579**	**0.997**	**0.410**	**0.994**	**0.258**	**n/a**
189	88.751	52.571	0.996	0.408	0.992	0.256	0.717
118	55.539	32.822	0.994	0.409	0.988	0.257	0.300
94	44.312	26.147	0.992	0.410	0.985	0.258	0.017
136	63.959	37.829	0.995	0.409	0.989	0.257	0.446
114	53.667	31.710	0.994	0.409	0.987	0.257	0.261
222	104.188	61.750	0.997	0.407	0.993	0.256	0.820
135	63.491	37.551	0.995	0.409	0.989	0.257	0.439
180	84.541	50.068	0.996	0.408	0.992	0.256	0.682
174	81.735	48.399	0.996	0.408	0.992	0.256	0.657
121	56.942	33.657	0.994	0.409	0.988	0.257	0.327
93	43.844	25.868	0.992	0.410	0.985	0.258	0.002
94	44.312	26.147	0.992	0.410	0.985	0.258	0.017
24	11.567	6.676	0.971	0.423	0.943	0.268	-3.977
159	74.718	44.227	0.995	0.408	0.991	0.256	0.586
81	38.231	22.531	0.991	0.411	0.982	0.258	-0.206
141	66.298	39.220	0.995	0.408	0.990	0.257	0.481
167	78.460	46.452	0.996	0.408	0.991	0.256	0.625
45	21.390	12.517	0.984	0.415	0.969	0.262	-1.488
101	47.586	28.094	0.993	0.410	0.986	0.258	0.113
113	53.200	31.431	0.994	0.409	0.987	0.257	0.251
81	38.231	22.531	0.991	0.411	0.982	0.258	-0.206
154	72.379	42.836	0.995	0.408	0.991	0.256	0.559
160	75.186	44.505	0.995	0.408	0.991	0.256	0.591
125	58.813	34.769	0.994	0.409	0.989	0.257	0.362

Appendix XV. Filter choosing chromaticities (Experiment 4 and 6)

Gamut of chromaticities, used for the filter choosing processes of Experiment 4 and 6, plotted on the CIE 1976 Lu'v' chromaticity diagram. Twenty-four chromaticities of a high ■ and low saturation setting ▣ were obtained from the colorimeter (Wilkins, 2015, personal communication). Radial lines connect approximately equal hues of low and high saturation setting. Colour names are approximations.

Experiment 4 used the chromaticities of the high saturation setting, including chromaticities represented by the outer concentric line. This was because there were six chromaticities in the low saturation setting that were too low to calculate a placebo hue of the required saturation and colour difference (Wilkins, 2015, personal communication). Experiment 6 used the chromaticities of low saturation setting. Corresponding luminance, chromaticities and saturations are displayed in the table on the following page. Colour names are approximate descriptions only.



Appendix XV. continued

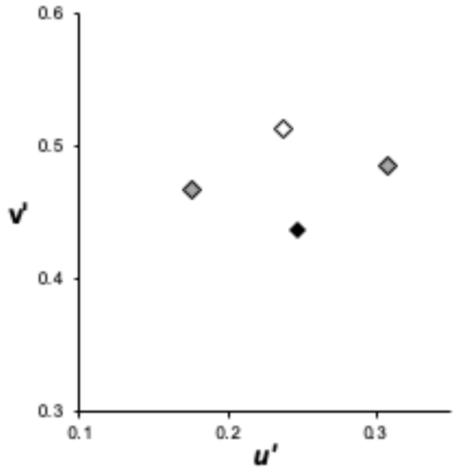
Saturation luminance and hue of the high ■ and low ▩ saturation settings of the colorimeter. * indicates the chromaticities that were too low to be used for calculating a placebo filter in Experiment 4.

	Low saturation setting ▩				High saturation setting ■			
	Saturation	Luminance	Hue		Saturation	Luminance	Hue	
Colour descriptor	S_{uv}	L	u'	v'	Suv	L	u'	v'
rose	0.049	16.15	0.285	0.522	0.098	15.00	0.334	0.518
orange	*0.034	15.53	0.265	0.533	0.077	12.23	0.309	0.540
yellow-orange	*0.0262	16.10	0.245	0.538	0.059	11.98	0.278	0.555
yellow	*0.0317	17.08	0.228	0.544	0.048	16.66	0.236	0.562
lime-green	*0.036	16.60	0.217	0.543	0.052	13.98	0.212	0.559
green	0.054	14.84	0.189	0.539	0.089	12.05	0.156	0.551
mint-green	0.051	14.27	0.188	0.528	0.092	10.52	0.146	0.528
aqua	0.048	14.85	0.189	0.519	0.092	10.62	0.145	0.509
blue	0.063	14.16	0.187	0.475	0.114	10.99	0.158	0.431
blue-purple	0.048	15.41	0.206	0.477	0.106	11.73	0.187	0.419
purple	*0.0312	16.67	0.242	0.483	0.074	14.56	0.260	0.443
pink	*0.0243	16.20	0.258	0.501	0.070	13.37	0.295	0.474

Appendix XVI. Acquiring placebo hues (Experiment 4)

The placebo hue was automatically generated by random selection from two possibilities ◇,that had a CIE 1976 LUV colour difference of 78 from the chosen hue and were equidistant from the chosen hue and whitepoint (Wilkins et al., 2002; Wilkins, 2015, personal communication).

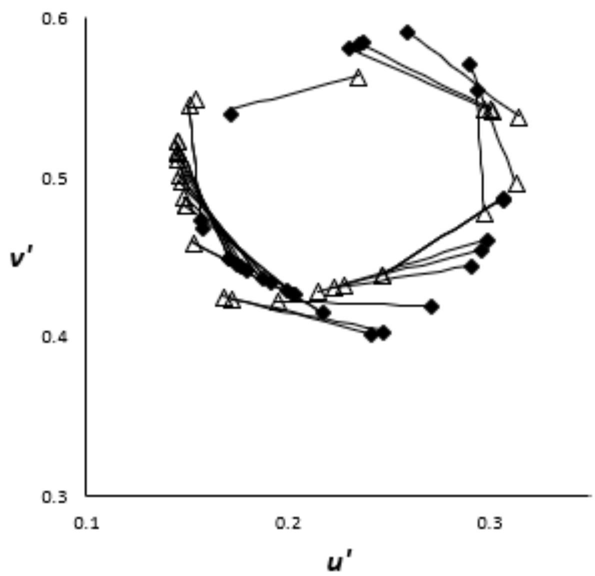
The table (below) displays chromaticities of the chosen filter and the two possible placebo hues for 28 participants in Experiment 4. One set of chosen and placebo hues are highlighted and plotted in the CIE 1976 LUV chromaticity diagram (right). The placebo hues that were used in Experiment 4 are displayed in Appendix XVII.

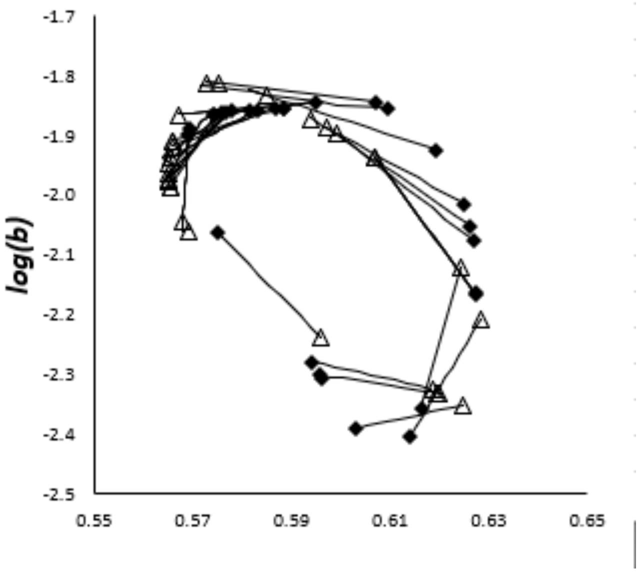


Chosen hue ◆		Placebo 1 ◇		Placebo 2 ◇	
u’	v’	u’	v’	u’	v’
0.145	0.522	0.184	0.588	0.171	0.449
0.247	0.438	0.176	0.468	0.307	0.486
0.145	0.515	0.178	0.584	0.176	0.444
0.149	0.482	0.156	0.559	0.203	0.427
0.313	0.497	0.262	0.439	0.290	0.570
0.298	0.478	0.229	0.443	0.294	0.554
0.173	0.423	0.129	0.487	0.247	0.403
0.149	0.482	0.156	0.559	0.203	0.427
0.145	0.522	0.184	0.588	0.171	0.449
0.145	0.517	0.180	0.585	0.175	0.446
0.168	0.425	0.127	0.490	0.241	0.401
0.146	0.502	0.169	0.575	0.187	0.436
0.147	0.497	0.166	0.571	0.191	0.434
0.314	0.537	0.300	0.462	0.259	0.591
0.154	0.549	0.212	0.600	0.157	0.472
0.297	0.543	0.286	0.467	0.230	0.580
0.301	0.542	0.289	0.466	0.236	0.583
0.195	0.422	0.142	0.478	0.271	0.419
0.228	0.432	0.164	0.475	0.299	0.461
0.215	0.428	0.156	0.477	0.291	0.444
0.235	0.583	0.300	0.542	0.172	0.539
0.149	0.487	0.159	0.563	0.199	0.429
0.223	0.430	0.161	0.476	0.296	0.454
0.151	0.544	0.206	0.599	0.158	0.468
0.153	0.458	0.139	0.534	0.217	0.415
0.145	0.511	0.175	0.581	0.179	0.442
0.302	0.542	0.289	0.466	0.238	0.584
0.247	0.438	0.176	0.468	0.307	0.486

Appendix XVII. Chromaticities of filters (Experiment 4)

CIE (1976) diagram (top) and MacLeod-Boynton (1979) diagram (bottom) displaying chromaticities of each participants' chosen ◆ and placebo Δ filter in Experiment 4. The connecting lines represent the chromaticities that were displayed during the colour change. Data is displayed for 28 of the 29 participants. A corresponding table displaying the chromaticity values and their saturations (CIE S_{uv}) is on the following page; measurements of luminance and contrast are displayed in Appendix XVIII.





Appendix XVII continued

Saturation	Chosen hue ◆					Placebo hue Δ				
	CIE (1976)		MacLeod-Boyton (1979)			CIE (1976)		MacLeod-Boyton (1979)		
S_{uv}	u'	v'	r	b	log b	u'	v'	r	b	log b
0.092	0.145	0.522	0.565	0.010	-1.989	0.171	0.449	0.574	0.014	-1.867
0.076	0.247	0.438	0.607	0.012	-1.937	0.307	0.486	0.627	0.007	-2.164
0.092	0.145	0.515	0.565	0.011	-1.973	0.176	0.444	0.577	0.014	-1.862
0.093	0.149	0.482	0.566	0.012	-1.910	0.203	0.427	0.588	0.014	-1.855
0.078	0.313	0.497	0.628	0.006	-2.210	0.290	0.570	0.614	0.004	-2.404
0.070	0.298	0.478	0.624	0.008	-2.121	0.294	0.554	0.616	0.004	-2.357
0.111	0.173	0.423	0.575	0.015	-1.813	0.247	0.403	0.610	0.014	-1.855
0.093	0.149	0.482	0.566	0.012	-1.910	0.203	0.427	0.588	0.014	-1.855
0.092	0.145	0.522	0.565	0.010	-1.989	0.171	0.449	0.574	0.014	-1.867
0.092	0.145	0.517	0.565	0.011	-1.978	0.175	0.446	0.576	0.014	-1.863
0.112	0.168	0.425	0.573	0.015	-1.812	0.241	0.401	0.607	0.014	-1.844
0.091	0.146	0.502	0.565	0.011	-1.947	0.187	0.436	0.581	0.014	-1.858
0.091	0.147	0.497	0.565	0.012	-1.938	0.191	0.434	0.583	0.014	-1.857
0.081	0.314	0.537	0.625	0.004	-2.353	0.259	0.591	0.603	0.004	-2.391
0.090	0.154	0.549	0.569	0.009	-2.062	0.157	0.472	0.569	0.013	-1.899
0.067	0.297	0.543	0.618	0.005	-2.325	0.230	0.580	0.594	0.005	-2.279
0.070	0.301	0.542	0.620	0.005	-2.331	0.236	0.583	0.596	0.005	-2.302
0.101	0.195	0.422	0.585	0.015	-1.835	0.271	0.419	0.619	0.012	-1.924
0.082	0.228	0.432	0.599	0.013	-1.898	0.299	0.461	0.627	0.008	-2.077
0.088	0.215	0.428	0.594	0.013	-1.873	0.291	0.444	0.625	0.010	-2.017
0.070	0.235	0.563	0.596	0.006	-2.239	0.172	0.539	0.575	0.009	-2.064
0.092	0.149	0.487	0.566	0.012	-1.920	0.199	0.429	0.587	0.014	-1.856
0.084	0.223	0.430	0.597	0.013	-1.887	0.296	0.454	0.626	0.009	-2.052
0.091	0.151	0.544	0.568	0.009	-2.046	0.158	0.468	0.569	0.013	-1.890
0.100	0.153	0.458	0.567	0.014	-1.865	0.217	0.415	0.595	0.014	-1.846
0.092	0.145	0.511	0.565	0.011	-1.965	0.179	0.442	0.578	0.014	-1.860
0.071	0.302	0.542	0.620	0.005	-2.332	0.238	0.584	0.596	0.005	-2.309
0.076	0.247	0.438	0.607	0.012	-1.938	0.307	0.486	0.627	0.007	-2.165

Appendix XVIII. Luminance and contrast measurements of text stimuli in Experiments 4

Inputted 'RGB' values were 'Text background' and a minimum text pixel of 0, which outputted a Minimum pixel luminance (M) of 0.34 cd/m^2. 'Text background luminance' was the luminance of the screen background without the text. Weber and Michelson were calculated using Formulae 1 and 2 with minimum luminance defined by the 'Minimum pixel luminance', (M), and 'Average text-pixel luminance', (A).); note that Michelson text contrast is defined as Michelson contrast (A) here.

Data is displayed for 28 of the 29 participants. Corresponding chromaticities are displayed in Appendix XVII.

Text background (Average RGB)	Text background luminance (cd/m^2)	Average text-pixel luminance (A) (cd/m^2)	Weber contrast using M	Weber contrast using A	Michelson contrast using M	Michelson contrast using A
38	18.116	10.570	0.981	0.417	0.828	0.926
50	23.729	13.908	0.986	0.414	0.867	0.943
38	18.116	10.570	0.981	0.417	0.828	0.925
19	9.228	5.285	0.963	0.427	0.683	0.856
25	12.035	6.954	0.972	0.422	0.750	0.889
48	22.794	13.351	0.985	0.414	0.861	0.941
40	19.051	11.126	0.982	0.416	0.836	0.929
19	9.228	5.285	0.963	0.427	0.683	0.857
38	18.116	10.570	0.981	0.417	0.828	0.926
19	9.228	5.285	0.963	0.427	0.683	0.856
20	9.696	5.563	0.965	0.426	0.697	0.863
38	18.116	10.570	0.981	0.417	0.828	0.925
38	18.116	10.570	0.981	0.417	0.832	0.923
22	10.631	6.119	0.968	0.424	0.720	0.875
43	20.455	11.961	0.983	0.415	0.847	0.934
43	20.455	11.961	0.983	0.415	0.846	0.934
44	20.923	12.239	0.984	0.415	0.849	0.935
43	20.455	11.961	0.983	0.415	0.848	0.933
47	22.326	13.073	0.985	0.414	0.859	0.939
45	21.390	12.517	0.984	0.415	0.853	0.937
53	25.133	14.742	0.986	0.413	0.874	0.946
38	18.116	10.570	0.981	0.417	0.828	0.926
23	11.099	6.398	0.969	0.424	0.731	0.880
41	19.519	11.404	0.983	0.416	0.840	0.931
39	18.584	10.848	0.982	0.416	1.000	1.000
38	18.116	10.570	0.981	0.417	1.000	1.000
21	10.164	5.841	0.967	0.425	1.000	1.000
40	19.051	11.126	0.982	0.416	1.000	1.000

Appendix XIX.Filter users' chosen colours and cloze reading speed improvements (Experiment 3 and 4)

Number of participants who chose each category of colour of filter in Experiment 3 and 4. Chromaticities were categorised to the nearest hue angle. Colour descriptors correspond with those used in Appendix XV. For each colour, average baseline (no filter) reading speed are displayed for Experiment 3; average improved reading speed (with filter) are displayed for Experiment 3 and 4.

Experiment 3

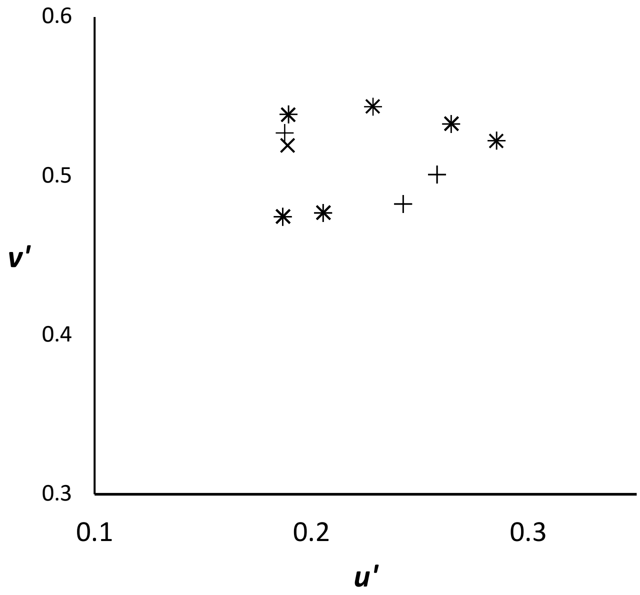
Average hue angle (huv)	Colour descriptor	Number of participants	Cloze baseline (words/min)	Cloze improvement (%)
0	rose	2	128.943	-1.811
30	orange	4	102.227	3.018
60	yellow-orange	1	73.064	-23.037
90	yellow	1	116.516	-1.083
120	lime-green	0		
150	green	1	100.212	-3.698
180	mint-green	3	139.464	-15.843
210	aqua	5	121.383	-14.999
240	blue	3	122.600	2.401
270	blue-purple	1	152.420	-17.975
300	purple	1	111.575	33.307
330	pink	2	74.593	21.246

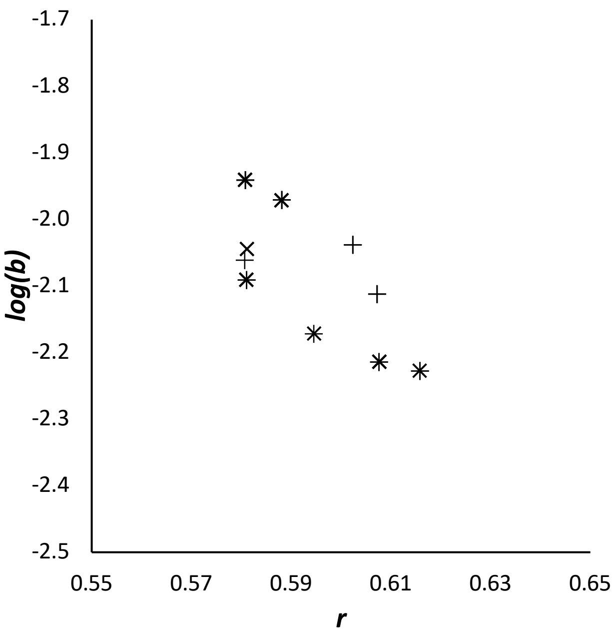
Experiment 4

Average hue angle (h_{uv})	Colour descriptor	Number of participants	Cloze improvement (%)
0	rose	1	3.707
30	orange	4	-4.171
60	yellow-orange	0	
90	yellow	1	-19.958
120	lime-green	0	
150	green	2	-19.445
180	mint-green	8	-8.009
210	aqua	3	9.873
240	blue	3	2.989
270	blue-purple	5	-16.247
300	purple	0	
330	pink	1	-25.710

Appendix XX. Chromaticities of filters (Experiment 6)

CIE (1976) diagram (top) and MacLeod-Boynton (1979) diagram (bottom) displaying chromaticities of each participants' chosen filter in the leading (+) and non-leading (x) conditions of Experiment 6. Data is displayed for all 40 participants, with many data points overlapping due to there being only 12 filter options available. A corresponding table displaying the chromaticity values and their saturations (S_{uv}) is on the following page; measurements of luminance and contrast are displayed in Appendix XXI.





Appendix XX. continued

Leading condition (+)

Saturation	CIE (1976)		MacLeod-Boyton (1979)		
S_{uv}	u'	v'	r	b	log b
0.051	0.188	0.527	0.581	0.009	-2.062
0.032	0.228	0.544	0.595	0.007	-2.172
0.024	0.258	0.501	0.607	0.008	-2.112
0.031	0.242	0.483	0.602	0.009	-2.038
0.054	0.189	0.539	0.581	0.008	-2.091
0.031	0.242	0.483	0.602	0.009	-2.038
0.034	0.265	0.533	0.608	0.006	-2.214
0.024	0.258	0.501	0.607	0.008	-2.112
0.034	0.265	0.533	0.608	0.006	-2.214
0.054	0.189	0.539	0.581	0.008	-2.091
0.063	0.187	0.475	0.581	0.011	-1.941
0.048	0.205	0.477	0.588	0.011	-1.971
0.049	0.285	0.522	0.616	0.006	-2.227
0.031	0.242	0.483	0.602	0.009	-2.038
0.048	0.205	0.477	0.588	0.011	-1.971
0.063	0.187	0.475	0.581	0.011	-1.941
0.048	0.205	0.477	0.588	0.011	-1.971
0.063	0.187	0.475	0.581	0.011	-1.941
0.049	0.285	0.522	0.616	0.006	-2.227
0.024	0.258	0.501	0.607	0.008	-2.112

Non-leading condition (x)

Saturation	CIE (1976)		MacLeod-Boyton (1979)		
S_{uv}	u'	v'	r	b	log b
0.034	0.265	0.533	0.608	0.006	-2.214
0.063	0.187	0.475	0.581	0.011	-1.941
0.054	0.189	0.539	0.581	0.008	-2.091
0.048	0.189	0.519	0.581	0.009	-2.044
0.032	0.228	0.544	0.595	0.007	-2.172
0.048	0.205	0.477	0.588	0.011	-1.971
0.034	0.265	0.533	0.608	0.006	-2.214
0.063	0.187	0.475	0.581	0.011	-1.941
0.054	0.189	0.539	0.581	0.008	-2.091
0.048	0.189	0.519	0.581	0.009	-2.044
0.063	0.187	0.475	0.581	0.011	-1.941
0.034	0.265	0.533	0.608	0.006	-2.214
0.049	0.285	0.522	0.616	0.006	-2.227
0.048	0.205	0.477	0.588	0.011	-1.971
0.034	0.265	0.533	0.608	0.006	-2.214
0.048	0.205	0.477	0.588	0.011	-1.971
0.063	0.187	0.475	0.581	0.011	-1.941
0.054	0.189	0.539	0.581	0.008	-2.091
0.048	0.205	0.477	0.588	0.011	-1.971
0.063	0.187	0.475	0.581	0.011	-1.941

Appendix XXI. Luminance and contrast measurements of text stimuli in Experiments 6

Inputted 'RGB' values were 'Text background' and a minimum text pixel of 0, which outputted a Minimum pixel luminance (M) of 0.34 cd/m^2. 'Text background luminance' was the luminance of the screen background without the text. Weber and Michelson were calculated using Formulae 1 and 2 with minimum luminance defined by the 'Minimum pixel luminance', (M), and 'Average text-pixel luminance', (A).); note that Michelson text contrast is defined as Michelson contrast (A) here.

Contrast reduction from the 'no filter' condition (bold) was calculated with the Michelson formula using minimum luminance of the average text luminance (A). Data is displayed for the 40 participants who chose a filter in Experiment 6: the 'leading condition' is below and the 'non-leading condition' is on the following page. Corresponding chromaticities are displayed in Appendix XX.

Leading condition

Text background (Average RGB)	Text background luminance (cd/m^2)	Average text-pixel luminance (A) (cd/m^2)	Weber contrast using M	Weber contrast using A	Michelson contrast using M	Michelson contrast using A	Contrast reduction using Michelson with A
255	**119.625**	**70.579**	**0.997**	**0.410**	**0.994**	**0.258**	**n/a**
51.306	24.197	14.186	0.986	0.414	0.972	0.261	-1.149
61.550	29.343	17.246	0.988	0.412	0.977	0.260	-0.698
58.345	27.472	16.133	0.988	0.413	0.976	0.260	-0.843
60.056	28.407	16.689	0.988	0.412	0.976	0.260	-0.768
53.383	25.133	14.742	0.986	0.413	0.973	0.261	-1.054
60.056	28.407	16.689	0.988	0.412	0.976	0.260	-0.768
55.898	26.536	15.577	0.987	0.413	0.975	0.260	-0.922
58.345	27.472	16.133	0.988	0.413	0.976	0.260	-0.843
55.898	26.536	15.577	0.987	0.413	0.975	0.260	-0.922
53.383	25.133	14.742	0.986	0.413	0.973	0.261	-1.054
50.905	24.197	14.186	0.986	0.414	0.972	0.261	-1.149
55.467	26.068	15.299	0.987	0.413	0.974	0.260	-0.964
58.159	27.472	16.133	0.988	0.413	0.976	0.260	-0.843
60.056	28.407	16.689	0.988	0.412	0.976	0.260	-0.768
55.467	26.068	15.299	0.987	0.413	0.974	0.260	-0.964
50.905	24.197	14.186	0.986	0.414	0.972	0.261	-1.149
55.467	26.068	15.299	0.987	0.413	0.974	0.260	-0.964
50.905	24.197	14.186	0.986	0.414	0.972	0.261	-1.149
58.159	27.472	16.133	0.988	0.413	0.976	0.260	-0.843
58.345	27.472	16.133	0.988	0.413	0.976	0.260	-0.843

Appendix XXI continued

Non-leading condition

Text background (Average RGB)	Text background luminance (cd/m^2)	Average text-pixel luminance (A) (cd/m^2)	Weber contrast using M	Weber contrast using A	Michelson contrast using M	Michelson contrast using A	Contrast reduction using Michelson with A
55.898	26.536	15.577	0.987	0.413	0.881	0.948	-0.922
50.905	24.197	14.186	0.986	0.414	0.870	0.943	-1.149
53.383	25.133	14.742	0.986	0.413	0.875	0.945	-1.054
53.419	25.133	14.742	0.986	0.413	0.875	0.945	-1.054
61.550	29.343	17.246	0.988	0.412	0.892	0.953	-0.698
55.467	26.068	15.299	0.987	0.413	0.879	0.947	-0.964
55.898	26.536	15.577	0.987	0.413	0.881	0.948	-0.922
50.905	24.197	14.186	0.986	0.414	0.870	0.943	-1.149
53.383	25.133	14.742	0.986	0.413	0.875	0.945	-1.054
53.419	25.133	14.742	0.986	0.413	0.875	0.945	-1.054
50.905	24.197	14.186	0.986	0.414	0.870	0.943	-1.149
55.898	26.536	15.577	0.987	0.413	0.881	0.948	-0.922
58.159	27.472	16.133	0.988	0.413	0.885	0.950	-0.843
55.467	26.068	15.299	0.987	0.413	0.879	0.947	-0.964
55.898	26.536	15.577	0.987	0.413	0.881	0.948	-0.922
55.467	26.068	15.299	0.987	0.413	0.879	0.947	-0.964
50.905	24.197	14.186	0.986	0.414	0.870	0.943	-1.149
53.383	25.133	14.742	0.986	0.413	0.875	0.945	-1.054
55.467	26.068	15.299	0.987	0.413	0.879	0.947	-0.964
50.905	24.197	14.186	0.986	0.414	0.870	0.943	-1.149

Appendix XXII. Experiment 6: chosen colours, Cloze and WRRT reading speeds and reading speed improvements

Number of participants who chose each category of colour of filter in Experiment 6. Colour descriptors correspond with those used in Appendix XV. Cloze and WRRT Average baseline (no filter) reading speeds and Average improved reading speed (with filter) are displayed for each colour.

Average hue angle (huv)	Colour descriptor	Number of participants	Cloze baseline (words/min)	Cloze Improvement (%)	WRRT Baseline (words/min)	WRRT Improvement (%)
0	rose	3	140.492	-24.573	168.333	-2.799
30	orange	6	137.626	15.768	161.667	5.987
60	yellow-orange	0				
90	yellow	0				
120	lime-green	2	135.514	8.748	176.500	-3.226
150	green	6	143.405	8.591	160.600	4.876
180	mint-green	2	150.398	-5.142	155.000	-1.633
210	aqua	8	143.179	-2.760	161.375	-0.759
240	blue	7	152.967	-2.551	165.000	1.377
270	blue-purple	3	129.561	9.288	164.000	1.217
300	purple	0				
330	pink	3	140.171	-7.816	150.667	6.448

Appendix XXIII. Ratings of text (Experiment 7)

Participants were asked to rate their preference of text when printed on five different paper conditions: White (W), Light grey (LG), Light blue (LB), Dark grey (DG) and Dark blue (DB). A rating of 'one' indicated highest preference and 'five' indicated lowest preference.

Here, ratings are categorised according to participants 'VS Score' (number of symptoms of visual stress) within two participant groups: the 'VS group' (below) categorised participants with a VS score of two or more; the 'no VS group' (right) categorised participants with a VS score of one or less.

VS group

VS	Rating				
Score	W	LG	LB	DG	DB
2	1	3	2	4	5
2	1	4	3	5	2
2	4	2	3	1	5
2	3	1	2	4	5
2	5	3	4	1	2
2	3	1	2	4	5
2	4	1	2	3	5
2	2	1	3	4	5
2	2	1	5	4	3
3	4	1	5	2	3
3	3	1	2	5	4
3	4	1	2	5	3
3	4	2	1	3	5
3	5	4	1	2	3
4	4	2	1	3	5
4	5	4	2	3	1
4	4	1	2	3	5
4	5	1	2	3	4
4	4	1	2	3	5
4	5	4	1	2	3

No VS group

VS	Rating				
Score	W	LG	LB	DG	DB
0	1	3	2	4	5
0	2	3	1	5	4
0	3	5	2	1	4
0	2	1	4	5	3
0	2	1	3	5	4
0	4	1	3	2	5
0	3	4	5	1	2
0	3	1	2	4	5
0	3	5	1	4	2
0	1	3	2	4	5
0	1	3	4	2	5
0	1	5	3	2	4
0	1	5	4	2	3
0	1	5	3	2	4
0	1	3	2	4	5
0	1	2	3	4	5
0	1	3	2	4	5
0	1	3	2	5	4
0	1	3	5	4	2
0	3	4	2	1	5
0	2	1	4	3	5
0	2	3	1	4	5
0	5	1	4	2	3
0	3	2	4	1	5
0	1	2	3	4	5
0	1	3	2	4	5
0	1	2	3	5	4
0	1	2	3	4	5
0	2	1	4	5	3
0	1	2	3	5	4
0	1	2	3	5	4
0	1	4	2	5	3
0	3	2	1	4	5
1	3	4	2	5	1
1	2	3	1	5	4
1	1	3	2	4	5
1	2	1	3	4	5
1	5	3	4	1	2
1	3	1	4	2	5
1	2	4	3	5	1
1	2	3	4	1	5
1	2	3	1	5	4
1	1	5	2	3	4
1	2	3	1	4	5
1	1	2	3	4	5
1	1	2	5	3	4
1	1	2	3	5	4
1	1	2	3	4	5
1	1	2	5	3	4
1	4	1	3	2	5
1	2	4	3	5	1
1	1	3	2	4	5
1	1	2	3	4	5
1	1	3	2	4	5
1	1	2	3	4	5
1	2	5	3	1	4
1	3	1	4	2	5
1	4	1	2	5	3
1	2	4	3	5	1

Appendix XXIV. Luminance and contrast measurements of text stimuli in Experiment 8

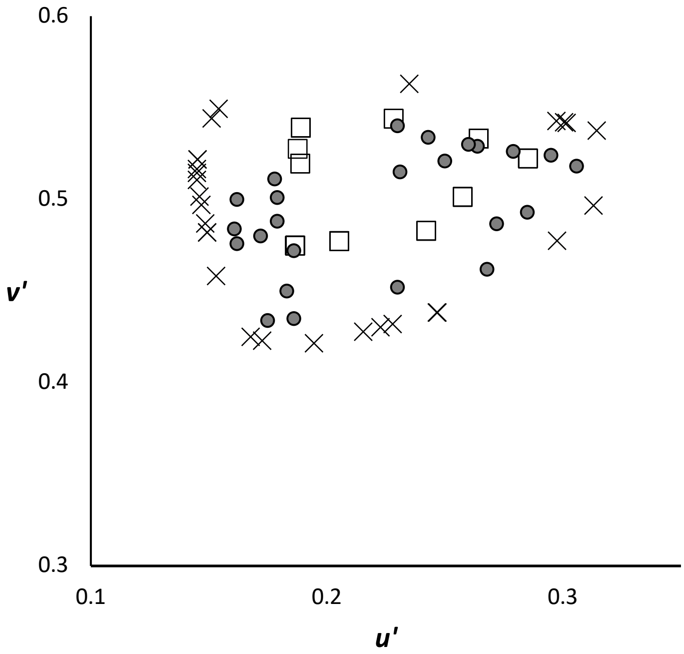
Inputted 'RGB' values were the luminance of the 'Minimum text pixel' and 'Text background' on a scale of 0-255; 'Text luminance' was the average luminance of the letters and background; 'Text background luminance' was the luminance of the screen background without the text; 'Weber, Michelson and RMS contrast were calculated using formula 1, 2 and 3; 'Minimum luminance' and 'Average text-pixel luminance' (the average luminance of the pixels that comprised the letters) were used to determine two values of Weber and Michelson contrast, as denoted by (M) and (A);); note that Michelson text contrast is defined as Michelson contrast (A) here.

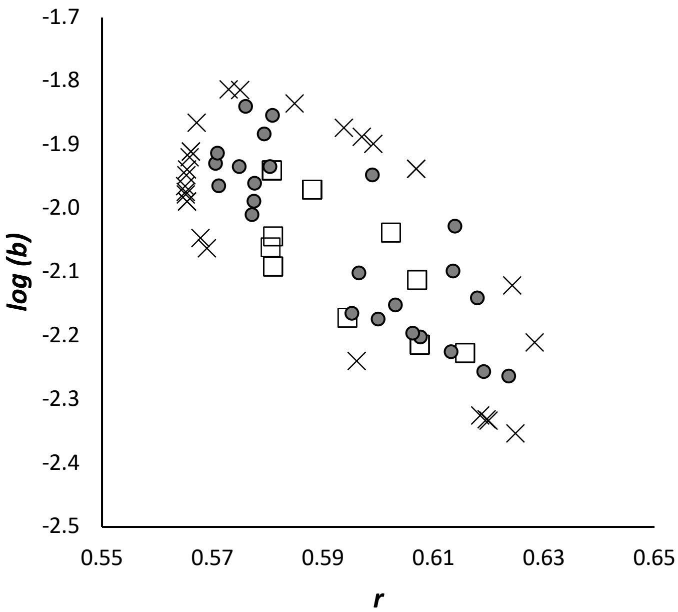
Values were calculated from a 59x59 pixel square of text (displayed in Figure 2.7). The stimuli of condition 1 and 2 stimuli had 1223 grey pixels and 2258 white pixels. The stimuli of condition 3, had 976 grey-scale pixels and 2505 white pixels. The text luminance of the 'Normal text' condition reached a close approximation to normal printed text (97.9cd/m^2) by using a screen that was not gamma corrected. However, similar to condition 3 in Experiments 1 and 2 (Appendix XI), condition 2 was calculated by applying the RMS contrast of printed text covered by an overlay to a stimuli with text luminance of normal text, when Michelson contrast should have been used.

Condition	Description	RGB		cd/m^2				Contrast				
									Weber		Michelson	
		Minimum text pixel	Text background	Text Luminance	Text background luminance	Minimum luminance (M)	Average text-pixel luminance (A)	RMS	(M)	(A)	(M)	(A)
1	Normal text	0	255	96.552	120.375	0.110	47.767	0.346	0.999	0.603	0.998	0.432
2	Reduced contrast	103	249	100.995	114.872	13.772	77.543	0.237	0.880	0.325	0.786	0.194
3	Increased linewidth	0	255	99.957	120.375	0.110	47.767	0.236	0.999	0.603	0.998	0.432

Appendix XXV. All chosen hues

CIE (1976) Lu'v' chromaticity diagram (top) and MacLeod-Boynton (1979) r log (b) diagram (bottom) displaying the chosen hues of all participants in Experiment 3(●), 4(×), and 6(□).

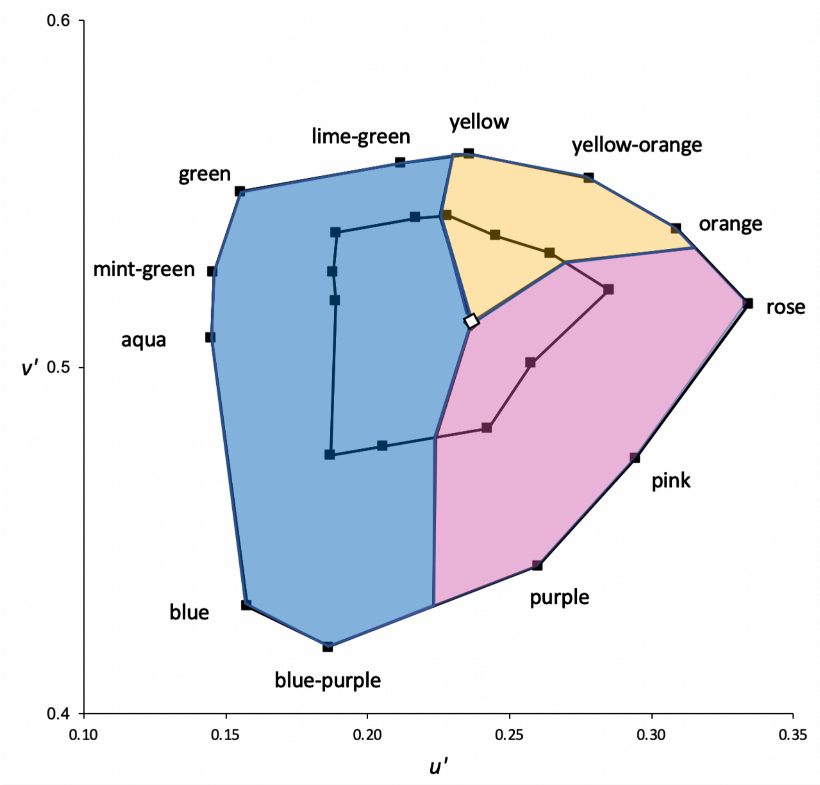




Appendix XXVI. Colour choices and gender

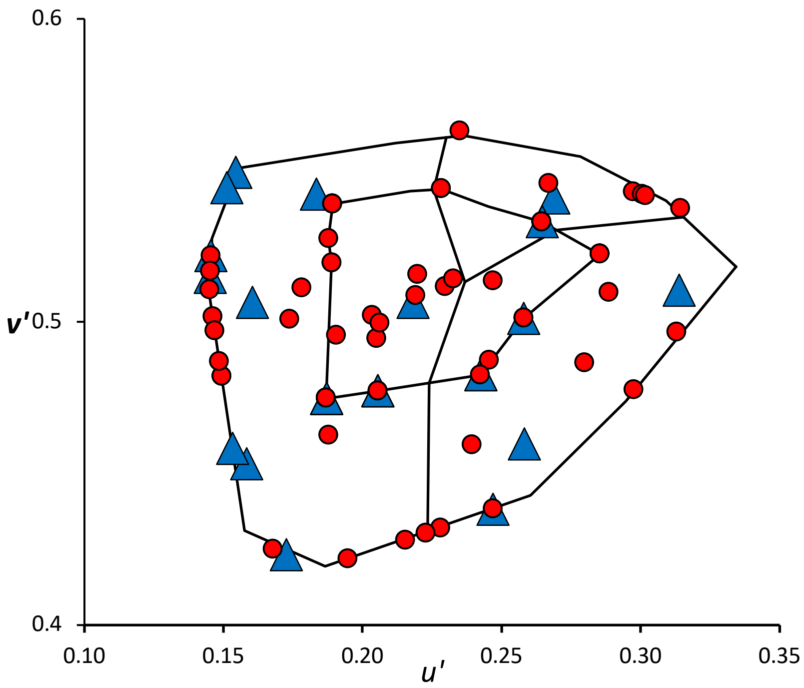
Conway et al.'s (2016) gender stereotypical categorisation of filters (left) that informed the boundaries used to separate the CIE (1976) chromaticity diagram (right) into gender stereotypical zones of male, female and neutral colours. Chosen filters of Experiment 3, 4 and 6 are categorised according to this gender stereotypical categorisation on the following page.

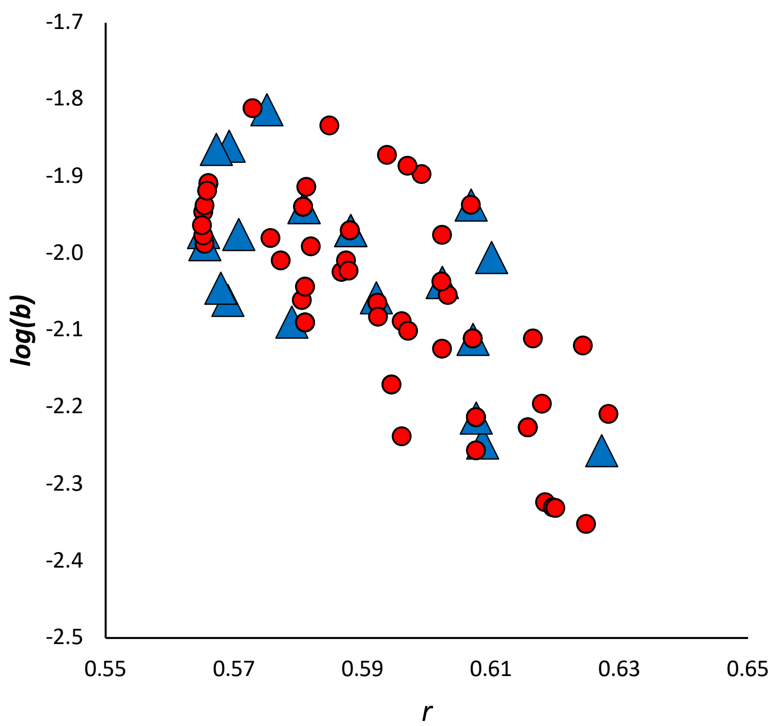
Colour of Overlay	Categorisation
Yellow	neutral
Yellow + Yellow	neutral
Yellow + Orange	neutral
Orange	neutral
Orange + Orange	neutral
Rose	female
Rose + Orange	female
Rose +Rose	female
Pink	female
Pink +Rose	female
Pink + Pink	female
Purple	female
Purple + Pink	female
Purple + Purple	female
Blue	male
Blue + Purple	male
Blue + Blue	male
Aqua	male
Aqua + Blue	male
Aqua + Aqua	male
Mint Green	male
Mint Green + Aqua	male
Mint Green + Mint Green	male
Lime Green	male
Lime Green + Mint Green	male
Lime Green + Lime Green	male
Lime Green + Yellow	male



Appendix XXVI. continued

CIE (1976) chromaticity diagram (top) and MacLeod-Boynton (1979) diagram (bottom) displaying the chosen hues of males ▲ and females ● in Experiment 3, 4 and 6. Boundaries of the CIE (1976) chromaticity diagram correspond to Conway et al.'s (2016) gender stereotypical categorisation of filters (see previous page).





Appendix XXVII. Saturation

Box plot displaying Saturations of chosen filters in Experiment 3 (Exp 3), 4 (Exp 4) and 6 (Exp 6).