OPO OPHTHALMIC & PHYSIOLOGICAL OPTICS
THE JOURNAL OF THE COLLEGE OF OPTOMETRISTS

Ophthalmic & Physiological Optics ISSN 0275-5408

INVITED REVIEW

The effect of coloured overlays and lenses on reading: a systematic review of the literature

Philip G. Griffiths[1], Robert H. Taylor[2], Lisa M. Henderson[3] and Brendan T. Barrett[4]

[1]Department of Ophthalmology, St Bernard's Hospital, Gibraltar, Gibraltar, [2]Department of Ophthalmology, York Teaching Hospital Foundation Trust, York, [3]Department of Psychology, University of York, York, and [4]School of Optometry & Vision Science, University of Bradford, Bradford, UK

Citation information: Griffiths PG, Taylor RH, Henderson LM & Barrett BT. The effect of coloured overlays and lenses on reading: a systematic review of the literature. *Ophthalmic Physiol Opt* 2016; 36: 519–544. doi: 10.1111/opo.12316

Keywords: coloured lenses, coloured overlays, Meares Irlen syndrome, reading, scotopic sensitivity syndrome, visual stress

Correspondence: Brendan T Barrett
E-mail: b.t.barrett@bradford.ac.uk

Received: 22 February 2016; Accepted: 19 July 2016

Abstract

Purpose: There are many anecdotal claims and research reports that coloured lenses and overlays improve reading performance. Here we present the results of a systematic review of this literature and examine the quality of the evidence.
Methods: We systematically reviewed the literature concerning the effect of coloured lenses or overlays on reading performance by searching the PsychInfo, Medline and Embase databases. This revealed 51 published items (containing 54 data sets). Given that different systems are in use for issuing coloured overlays or lenses, we reviewed the evidence under four separate system headings (*Intuitive, Irlen, Harris/Chromagen* and *Other*), classifying each published item using the Cochrane *Risk of Bias* tool.
Results: Although the different colour systems have been subjected to different amounts of scientific scrutiny, the results do not differ according to the system type, or whether the sample under investigation was classified as having visual stress (or a similarly defined condition), reading difficulty, or both. The majority of studies are subject to 'high' or 'uncertain' risk of bias in one or more key aspects of study design or outcome, with studies at lower risk from bias providing less support for the benefit of coloured lenses/overlays on reading ability. While many studies report improvements with coloured lenses, the effect size is generally small and/or similar to the improvement found with a placebo condition. We discuss the strengths and shortcomings of the published literature and, whilst acknowledging the difficulties associated with conducting trials of this type, offer some suggestions about how future trials might be conducted.
Conclusions: Consistent with previous reviews and advice from several professional bodies, we conclude that the use of coloured lenses or overlays to ameliorate reading difficulties cannot be endorsed and that any benefits reported by individuals in clinical settings are likely to be the result of placebo, practice or Hawthorne effects.

Introduction

In 1980, Olive Meares, a schoolteacher from New Zealand, described visual perceptual difficulties reported by some children when learning to read. These difficulties were apparently alleviated by placing sheets of coloured plastic, such as Perspex, over text.[1] Separately, American psychologist, Helen Irlen, documented the use of coloured overlays and lenses to the same effect.[2] The set of symptoms described in these publications became known as 'Irlen syndrome' (IS), 'Meares-Irlen syndrome', or 'scotopic sensitivity syndrome', the symptoms of which include (but are not restricted to) subjective reports of movement or blurring of print, doubling of letters, illusions of colour, glare from printed material, headaches, and eye-strain during reading.

In the UK, research into this area has been led by Professor Arnold Wilkins, who introduced the term 'visual stress' (used throughout the present paper* and developed one of the coloured filter systems reviewed here. It has been claimed by Wilkins[3] that visual stress may be one cause of reading difficulties†. Wilkins acknowledges the considerable overlap in symptoms between visual stress and asthenopia that arises due to other reasons including uncorrected refractive error, and oculomotor/binocular vision anomalies.[4] Once basic optometric needs have been addressed, it has been argued that the use of coloured overlays and lenses can have a positive impact on the symptoms of visual stress in affected individuals, which, in turn, may lead to better reading performance, in particular higher reading speed.[5]

Despite 35 years having elapsed since the initial description, neither the International Classification of Disease (ICD-10; World Health Organisation) nor the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association) list visual stress as a recognised disorder. Similarly, neither of these widely used diagnostic manuals makes any reference to visual-perceptual distortions as being associated with reading difficulty. The ability of coloured filters to improve reading performance in individuals who report symptoms of visual stress has been widely contested [6–11] and the practice has even been listed among 'neuromyths in education'.[12] The two narrative reviews that are broadly supportive of the use of coloured lenses/overlays adopt conflicting viewpoints. A review by Wilkins in 2002 argued that precisely the right colour is required,[4] whereas a review by Stein in 2014 argued that blue or yellow overlays suffice.[13] Taken together these narrative reviews raise doubt over whether the available evidence supports their widespread use. Indeed, it has been suggested that any improvement in reading performance seen in individuals who use coloured overlays or lenses is the result of a placebo effect, whereby belief that a product will help is enough to give the user the impression of improvement.[14]

Despite suspicions about the true effectiveness of coloured overlays and lenses, these 'reading aids' have received widespread media exposure and their use is regularly accepted in schools and higher education institutions. The use of coloured overlays and lenses continue to be endorsed by some dyslexia charity websites.[15] The issuing of coloured overlays has become embedded, to a greater or lesser extent, into the practice of a range of professionals in the UK including teachers, educational psychologists, optometrists, and NHS orthoptic departments. Furthermore, there is now an array of colour systems on offer (e.g., Irlen, Intuitive, ChromaGen/Harris), with proponents of each system claiming that their system provides an effective testing and management approach.[2,16]

What is the status of the evidence to support use of coloured overlays or lenses (spectacles or contact lenses) for the purpose of improving reading ability?

We have brought together a multidisciplinary team spanning psychology, optometry and ophthalmology to conduct a systematic search of the literature to address this question. A systematic review of the literature by Albon *et al.* [6] concluded that the available evidence was of too low quality to reach firm conclusions about the efficacy, and cost effectiveness of coloured lenses for reading disability. Similarly, a systematic review by Galuschka *et al.*[17] could not prove any positive effect of coloured lenses or overlays on literary achievement. This review is the first attempt to systematically and separately evaluate the quality of the evidence for each of the available colour systems: Irlen, Intuitive, ChromaGen/Harris and other systems that have received less attention, or are less widely used.

Method

Literature searches

We conducted our searches using Medline, PsycInfo and Embase. The searches utilised three 'concepts' identified during a preliminary search of the literature: [Concept 1] colour *(('Colour'/exp OR colour) OR (coloured OR colored) OR (lens OR lenses) OR (overlay OR overlays) OR (filter OR filters) OR (tint OR tinted))*, [Concept 2] reading *(reading OR text OR print OR printed OR words OR word OR writing OR write)*, [Concept 3] reading difficulties/visual stress terms *(Irlen OR 'Meares Irlen syndrome' OR 'Irlen syndrome' OR Meares OR 'visual stress' OR dyslexia OR 'dyslexia acquired' OR 'Learning disorders' OR 'specific learning disability' OR 'specific learning disorder' OR 'specific language disorder' OR 'specific language disability' OR 'specific language impairment' OR 'specific learning impairment' OR 'specific learning disorder' OR 'specific learning disability' OR magnocellular OR 'scotopic sensitivity syndrome' OR*

*We have used the term 'Visual Stress' throughout this review to signify the symptoms which are 'treated' using coloured lenses or overlays. Although this term was introduced by Professor Wilkins and is associated with the Intuitive colour system, here we use 'Visual Stress' (VS) more generally to describe any visual symptoms which may respond to colour intervention, regardless of the colour system (Irlen, Intuitive etc.). In the absence of precise definitions for the conditions which the various colour systems purport to treat/manage, we could have adopted one of the other terms in common usage, e.g. Meares-IS, IS or Scotopic Sensitivity Syndrome. Thus, our use of 'Visual Stress' does not signify support for the Intuitive system (or any system) over another system.

†We do not define here precisely what this term means. It means different things to different researchers. Rather, we have accepted at face value the different criteria that study authors have decided, for their sample, constitutes a 'reading difficulty' or 'reading disability'.

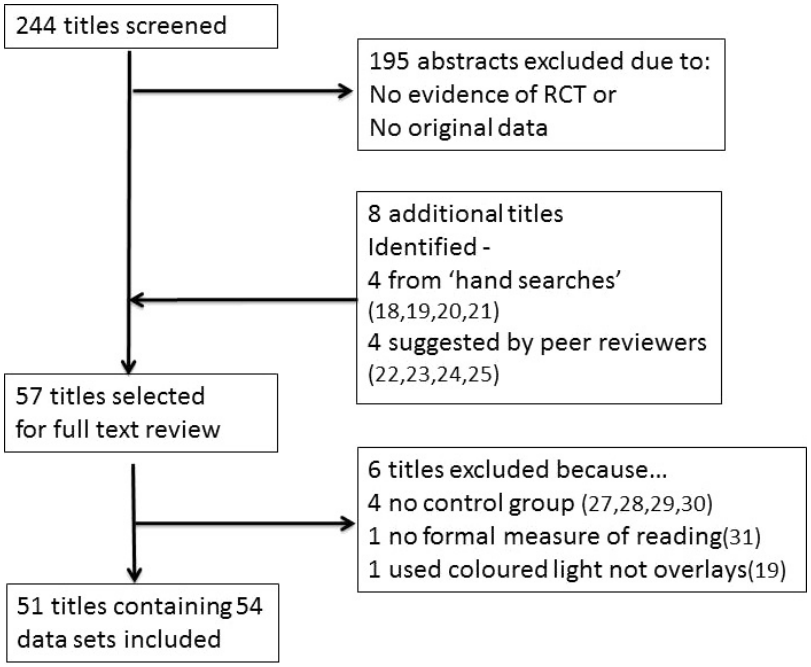

Figure 1. Flow diagram showing our search strategy and how this review came to examine 51 items of literature (but 54 data sets) concerning the impact of coloured overlays or lenses on reading.

misvis). *Figure 1* shows a Prisma flow chart of the number of published papers identified by these searches, and from other sources (e.g. through contact with published authors in the topic area and searching reference lists of the search-identified papers).

Inclusion and exclusion criteria

Figure 1 shows how the final number of studies included in this review was arrived at (n = 51). We included experimental studies, which featured primary data on at least one measure of reading ability or reading-related activity (e.g. reading unconnected words), and which described the effect on such measures when coloured filters (i.e., spectacle lenses, contact lenses or overlays) were worn or used. All of the 51 published items in our final list incorporated a control group, and employed either a crossover design (where participants undergo a number of different treatments or exposures and thus act as their own controls) or a parallel design (where the treatment group were compared to another group, for example, a group that did not receive coloured filters, or received a non-optimal colour). In searching the literature, studies of both adults and children were included and there were no restrictions on the baseline reading ability of the study sample, or on how the sample was identified. Some samples comprised 'poor' readers identified in remedial settings (e.g. classroom settings where the students were present because of known reading difficulties) whereas others comprised unselected samples (e.g., where all children in a school/year group participated). We noted whether the study looked at the 'overlap group' (i.e., individuals diagnosed with 'visual stress' *and* reading difficulties), or whether only 'visual stress' or reading difficulties were diagnosed in the sample under test. The review was not restricted to any one colour 'system' but instead included all studies where the effect of coloured lenses/overlays on reading performance had been examined, provided this was in the form of overlays or lenses. The results for each system were analysed separately. Studies were included irrespective of whether or not symptoms, or changes in symptoms, associated with use of colour were reported. The literature searches also revealed six unpublished PhD theses (i.e. from the so-called 'grey literature') that we included in our review.

Two authors (PG and LH) independently reviewed the 244 items (including two non-English papers) identified and excluded papers that did not fit the inclusion criteria; any paper that was selected by at least one reviewer was included. This exercise reduced the list to 49 items and neither of the non-English articles met the inclusion criteria (*Figure 1*). Individuals who had made a significant contribution to research on coloured lenses/overlays and visual stress or reading difficulties were invited (by e-mail) to view the list of 49 papers and to make suggestions for additional papers that should be included. Finally, we examined the reference lists of the 49 papers and identified a further

four papers which had not appeared in our search-engine results.[18–21] Peer reviewers for this review suggested a further four papers[22–25] and a total of 57 papers was therefore reviewed (*Figure 1*).

Procedures for review

The four authors worked in pairs (LH & RT; PG & BB). For each item, the pairs completed a form which gathered the following information from each item: a brief description of the study and design; whether or not there was a control group; which colour systems had been employed and whether lenses or overlays had been used; what the independent and dependent variables were (the latter had to feature some measure of reading in order for the study to be included) and which measures of reading had been used

Each published item was evaluated according to threats of internal and external validity, in accordance with the Cochrane Collaboration's tools for assessing bias.[26] Internal validity refers to the risk of bias resulting from study design and reporting. External validity refers to the degree to which the results, even if at low risk of internal bias, can be generalised to different settings and populations. The following domains of bias were considered: selection bias (e.g., judgements on the method for random sequence generation and whether intervention allocations could have been foreseen before or during enrolment), performance bias (e.g., when participants and personnel have knowledge of the intervention, such as. experimental or placebo tint used during the study), detection bias (e.g., when assessors have knowledge of the allocated intervention), attrition bias (e.g., bias arising from loss of participants from the study) and reporting bias (e.g., when only selective outcome measures are reported; the existence of a pre-trial protocol serves as evidence of 'low' risk of reporting bias). In keeping with advice from the Cochrane Collaboration, we did not sum the risk judgments to derive a global 'risk of bias' score for each study. This is because a study may be at serious risk of bias if the bias judgements are low in all but one area.

Many of the studies we reviewed were crossover studies which are at low risk of confounding due to problems with random sequence generation, allocation concealment and similarity of groups at baseline. We recorded these studies as being at low risk of bias in these domains even if a detailed account was not given for the method of sequence generation and allocation concealment. We considered studies that used disconnected text rather than naturalistic text of the sort encountered in everyday life to have limited external validity. In addition, studies that recruited participants from specialist clinics such as those at the Institute of Optometry or Dyslexia Research Trust were recorded as having high or uncertain external bias because they may not be representative of the general population of poor readers and may have been attracted to those clinics because of a prior belief in the effectiveness of coloured lenses and overlays.

We also gathered information from each item about attrition in the use of coloured overlays/lenses over time where such information was provided. The form was initially completed by one member of each team and then reviewed by the second member. In the event that a pair was not in agreement, the paper in question was referred to the second pair for discussion and agreement. To ensure consistency between the pairs, each pair selected the three of the papers from their list which had generated the most discussion about the risk-of-bias judgments and invited the other pair to reach their own, independent judgements about the risk of bias. No systematic differences in the application of criteria for the bias judgments were identified.

From the list of 57, six items were excluded following review. The reasons for these exclusions were as follows: on closer examination four did not feature a control group,[27–30] one featured the use of coloured light rather than lenses or overlays[19] and one did not include a formal measure of reading.[31]

Where items contained several studies, we viewed them as separate data sets and included them only if they satisfied our criteria. There were two items where this occurred, Jeanes *et al.* studies 4 and 6,[32] Lightstone *et al.* studies 1 and 2[33] and Wilkins *et al.* studies 1–3.[34] Therefore, there were additional data sets for Jeanes *et al.* (+1) and Wilkins and Lewis *et al.* (+2) and Lightstone *et al.* (+1). Based on the same criteria, it was noted that two publications by Robinson & Forman[35,36] reported on the same samples, so we considered them as one for the purposes of assessing bias making one less data set. Overall, therefore, there were three additional data sets in the 51 items that we reviewed, leaving a final total of 54 data sets.

Results

A synopsis of the number of items we identified by our searches and of our reasons for excluding a proportion of these is provided in *Figure 1*. In total, 54 data sets including 2690 participants were analysed; 23 studying the Intuitive system, 15 studying Irlen, four studies of Chromagen/Harris and 12 of non-commercial filters (referred to here as 'Other'). Table S1 (available as supporting information) contains details of the following characteristics for all studies included: study design, participant description, sample size, description of the diagnosis of visual stress, the intervention used, the dependent variables (typically the reading tests used) and any pertinent notes. *Table* 1 contains 'risk of bias' judgements for the studies we reviewed, grouped according to the colour system.

Table 1. Risk of bias for included studies

Study	Random sequence generation	Allocation concealment	Similarity of groups at baseline	Blinding of personnel and participants	Blinding of outcome assessment	Attrition bias (incomplete outcome data)	Reporting bias (Selective reporting)	External bias
Intuitive Overlays and Lenses								
Wilkins 1994[49]	Low	Low	Low	low	Low	High (40% data missing).	Low	Low
Wilkins 1996[22]	Low	Low	Low	High – no placebo control filter	High	Low	Low	High. Use of WRRT
Jeanes 1997[32]	Uncertain	Uncertain	NA	High	High	High – large part of primary school cohort lost	High – no pre-trial protocol	High -use of WRRT
Lightstone 1999[33]	Low for crossover component	Low	Low	Study I: High. Study II: Uncertain	High – unmasked study	Low	Low	High -recruited from specific learning difficulties clinic at IOO. Use of WRRT
Wilkins 1999[23]	Low	Low	Low	High	High	Low	Low	High –Use of WRRT
Wilkins 2001[34]								
Study 1	Low	Low	Low	High	High	Low	High- No pre-trial protocol	High – Use of WRRT
Study 2	Low	Low	Low	High	High	Low	High – No pre trial protocol	As above
Study 3	Low	Low	Low	Low	High	High	High – no pre trial protocol	As above
Bouldoukian 2002[46]	Low	Low	Low	High	High	Low	Uncertain	High – included subjects recruited from Institute of Optometry who were already using overlays. WRRT
Evans 2002[40]	Low	Low	Low	High – no placebo control for crossover part of study	High – no placebo control	Low	Uncertain	High. Subjects recruited via advert and questionnaire
Northway 2005[24]	Low	Low	Low	High – no placebo control overlay	High	Low	Low	High. Subjects with dyslexia attending hospital eye clinic. Use of WRRT

(continued)

Table 1 (continued)

Study	Random sequence generation	Allocation concealment	Similarity of groups at baseline	Blinding of personnel and participants	Blinding of outcome assessment	Attrition bias (incomplete outcome data)	Reporting bias (Selective reporting)	External bias
Singleton 2005[47]	Uncertain	Uncertain – only 5 in each group 3:2 vs 2:3.	High – big differences in starting WRRT between high VS and low VS groups	High – No placebo control	Uncertain – No placebo control	Low	Low	High. Cases drawn from disability services of Hull University. Use of WRRT
Kriss 2005[38]	Low	Low	Low	High -unmasked, no placebo control	High – unmasked, no placebo control	Low	Low	High. Cases drawn from dyslexia clubs and may have already been exposed to overlays and their supposed benefits. Use of WRRT
Hollis 2006[42]	Low	Low	Low	High -unmasked study – no placebo control	High – unmasked study, no placebo control	Low	Unclear – no pre-trial protocol	High. How were adult volunteers randomly selected? Use of WRRT
Smith 2007[16]	Low	Low	Low	High – Compares no overlay, intuitive overlay and Eye Level Ruler	High	Low	No pre trial protocol	High. Use of WRRT. Diagnostic criteria for visual stress
Singleton 2007[43]	Low	Low	Low	High – overlay vs no overlay	High	Low	Low	High. Already attending specialist clinic
Allen 2008[25]	Uncertain	Uncertain	Low	High – no placebo control filter	High	Low	Low	High. Use of WRRT
Mitchell 2008[45]	Parallel arms	low	low	High – VS untreated control group, Low VS placebo group	High – VS untreated control group, Low VS placebo group	low	low	High. Criteria for visual stress vague. Children from remedial unit
Allen 2010[41]	unclear	low	na	Low for between group comparisons, but high for crossover comparisons	High- no placebo group	unclear	low	High. Subjects recruited by advertisement. Use of WRRT
Henderson 2013[44]	Low	low	low	High – unmasked study, no placebo control	High – unmasked study, no placebo control	low	low	Uncertain – treating poor reading rather than visual stress

(continued)

Table 1 (continued)

Study	Random sequence generation	Allocation concealment	Similarity of groups at baseline	Blinding of personnel and participants	Blinding of outcome assessment	Attrition bias (incomplete outcome data)	Reporting bias (Selective reporting)	External bias
Monger 2015[20]	Low	Low	Low	High	High	Low	Low	High, WRRT rather than naturalistic text
Irlen overlays and lenses								
O'Connor 1990[58]	Low	Low	Low	High	Low	Low	Low	Low. Subjects recruited from school setting
Blaskey 1990[56]	Uncertain	Uncertain	Low	High	High	Uncertain	Low	High. Subjects recruited following report on '60 min television programme
Cotton 1990[57]	Low	Low	Low	High	High	High – only 22 out of 38 children returned for assessment	Low	High – limited assessment of reading problems. Difficult to generalise to other poor readers
Martin 1993[59]	High	Uncertain	High	Uncertain	High	Low	Low	Low– school setting
Fletcher 1994[60]	Low	Low	Low	High – unmasked study	High- unmasked study	Low	High – only F values from ANOVA reported	Uncertain – Non standardised reading test
Tyrell 1995[50]	Low for crossover component	Low	Low	High – no placebo control group	Low assessors listened to audiotapes of reading	Low	High – multiple small groups. No evidence of pre-trial protocol	High – risk of fatigue because of amount of testing in one session. Non standardised reading test in syllables per minute
Robinson 1999a, b[35,36]	Uncertain	Uncertain	Low	Uncertain – participants saw diagnosed tint	Low – audiotaped assessments	Low for first 3–4 months	Low	Uncertain Standardised reading but recruited from spld clinic
Noble 2009[65]	Uncertain	High	Low	High	High	Uncertain	Low	Low – participants drawn from mainstream schools
Ritchie 2011[61]	Low	Low	Low	Low	Low	Low	Uncertain – 3 participants who were not masked to their chosen colour analysed separately	Uncertain – Good sample drawn from school setting. Relevance of WRRT to reading naturalistic text. However also tested with GORT
Ritchie 2012[62]	Low	Low	Low	High- participants aware of chosen colour	High	High	High – no pre-trial protocol	Low – sample drawn from school setting

(continued)

Table 1 (continued)

Study	Random sequence generation	Allocation concealment	Similarity of groups at baseline	Blinding of personnel and participants	Blinding of outcome assessment	Attrition bias (incomplete outcome data)	Reporting bias (Selective reporting)	External bias
Grey literature								
Donovan 1995[66]	Low	Low	Low	High	Uncertain	High-25% did not complete the study	Low	Low
Mason 2000[67]	Low	Low	Low	High	High	Low	Low	Uncertain – Not clear how 30 subjects were selected
Faraci 2010[64]	Uncertain	High	High	High	High	Low	Uncertain	Uncertain – recruited from main stream school but not clear how they were selected
Morrison 2012[68]	Low	Low	Low	High – chosen colour vs clear colour vs random colour	Uncertain	Low	Low	Low
Adams 2013[69]	Low	Low	Low	High – chosen colour vs clear overlay	High – chosen colour vs clear overlay	Low	Low	Uncertain – subjects recruited from classroom setting but computer reading test
ChromaGen/Harris								
Harris 1999[76]	Low	Low	Low	High – placebo lenses light blue handling tint so study not masked	Low – lens changed out of sight of the experimenter	Low	High	High Subjects recruited following article in the media. Only used WRRT
Cardona 2010[71]	Low	Low	Low	High	High	Low	Unclear	High – not clear if results can be generalised to naturalistic text
Hall 2013[74]	Uncertain-Harris filter group and DRT group in different schools	Low	Uncertain – Harris group were significantly younger	Uncertain	High	Low	Uncertain- not clear what primary outcome measure was	Low – children selected from mainstream schools
Harries 2015[75]	Uncertain – see above	Low	Uncertain – Harris group were significantly younger	Uncertain	High	Low	Low	Low – see above
Miscellaneous/'Other'								
Saint-John 1988[85]	Low	Low	Low	High – study not masked	High – study not masked	Low	Low	Uncertain – non standardised reading test

(continued)

Table 1 (continued)

Study	Random sequence generation	Allocation concealment	Similarity of groups at baseline	Blinding of personnel and participants	Blinding of outcome assessment	Attrition bias (incomplete outcome data)	Reporting bias (Selective reporting)	External bias
Gole 1989[83]	Low	Low	Low	High	High	Low	Low	Uncertain – numerous exclusion criteria
Francis 1992[87]	Unclear	High	Low	High	high	High – 23% attrition from experimental group	low	Low
Menacker 1993[84]	Low	Low	Low	Uncertain	Uncertain	Low	Low	Uncertain – participants drawn from specialist services
Sawyer 1994[82]	High	High	Uncertain	High	High	Uncertain	Low	Uncertain participants drawn from specialist services
Evans 1994[86]	Low	Low	Low	High – unmasked study	High – unmasked study	Low	Uncertain	High – unclear if SVRST is a proxy for the demands of reading
Iovino 1998[79]	Low	Low	Low	Uncertain	Uncertain	Low	Uncertain	Not clear how children where selected hence relevance to general population
Christenson 2001[77]	Low	Low	Low	High	High	Low	Low	Uncertain – little information on nature of reading difficulties
Ray 2005[18]	Uncertain	Uncertain	Uncertain	High – placebo control obviously different to experimental intervention	Low	Low	High – originally blue placebo control group that was not reported	High- subjects recruited form Dyslexia Research Trust
Vidal-Lopez 2011[80]	Low	Low	Low	High – unmasked study	High – unmasked study	Low	Low	Low
Palomo-Alvarez[78]	Low	Low	Low	High – unmasked. study	High – unmasked. study	Low	Low	Low
Rosenfield 2015[21]	Low	Low	Low	Uncertain	High	Low	Uncertain – *post hoc* subgroup analysis for group with high symptom scores	High not reading naturalistic text

VS, visual stress; WRRT, Wilkins rate of reading test; spld, specific learning difficulties; IOO, Institute of Optometry; SRVST, simulated reading visual search task.

Intuitive overlays and lenses

Intuitive Overlays (www.ioosales.co.uk) consist of nine coloured overlays and one grey overlay. Overlays are selected by making pair-wise comparisons until the optimal single tint is found. The Intuitive Colorimeter (www.ceriumoptical.com) was subsequently developed to more precisely define the optimum colour required to ameliorate visual stress and for prescribing tinted spectacle lenses (known as Precision Tinted Lenses). The Intuitive Colorimeter consists of an illuminated chamber in which random letters arranged to resemble text can be viewed through an aperture. The hue, saturation and brightness can all be varied independently by the examiner according to the subjective responses of the person being tested. Spectacle lenses can be ordered based on this tint which may be different to that required for overlays.[33] Precision tinted lenses are generally only made up following sustained use of overlays; a detailed account of how to determine optimum tints for both Intuitive Overlays and Precision Tinted Lenses is available.[37]

The Wilkins Rate of Reading Test (WRRT) is frequently used to assess the benefit of coloured overlays and lenses.[22] The test comprises passages of randomly ordered, high-frequency words, printed in a small font. The text is designed to be crowded and visually aversive. The outcome measure is the number of words correctly read per minute (wpm). It is not intended to be a test of reading ability *per se* but rather a measure of the extent to which colour can influence reading rate. Reading the WRRT 5%, 8% or 10% faster with the chosen overlay has been used as the criterion for a clinically significant improvement with colour,[38] although there appears to be no basis for various cut-offs, and even the most stringent criteria remain arbitrary. A recent article has suggested that an increase in reading speed of 15% or more may be required to be confident that there is a genuine improvement with coloured overlays,[5] although we are unaware of any studies that use this diagnostic criterion for visual stress. These criteria also need to viewed in the context of the test-retest variability of the WRRT. For example, one study reported that 5% of children read more than 25% faster using Intuitive Overlays.[34] However the same paper also reports the test-retest variability of the WRRT. Reading from figure 2 in that paper, it appears that just by repeating the test (without any overlay), 3.8% of participants read more than 25% faster (and 12.3% read more than 15% faster) when tested on the second occasion. As a result, even a more rigorous criterion of reading the WRRT 15% faster may still produce large numbers of false positives. It is notable that in any given population (including those with and without visual stress), there is substantial variability in baseline reading performance prior to intervention (e.g. 20–145 words per minute in children aged 7–12[38]; see also[39] in Table S1). Wilkins *et al.*[5] acknowledge this variability, indicating that the highest rate of reading (>160 wpm) can be more than four times the slowest rate (<40 wpm) in children with similar scholastic reading ability. Consequently, it is not clear what the normal reading rate is for a particular age group and it is not possible to determine what a 'normal' range of improvement would be with an overlay. In line with this evaluation Wilkins *et al.* state 'a confusing array of criteria have been applied and further analysis of WRRT data is required'.[5]

It has been argued that changes in reading should be seen first with the WRRT and only later using naturalistic text, which depends on higher order reading skills. We found no evidence to support this claim. Furthermore, it could be argued using the same logic that changes would also be seen immediately in psychophysical tests that use gratings at the spatial frequencies that are considered to be aversive in visual stress, however, this has not been observed.[39]

Nineteen papers including 23 trials of Intuitive Overlays or Precision Tinted Lenses fit our selection criteria. In most cases these were exploratory studies that contained a cross-over trial as part of wider study investigating the use of overlays, where reading with a chosen overlay was compared to reading with a clear overlay or without an overlay[20,22–25,32–34,40–44] and were therefore at a high risk of bias due to the lack of a placebo control condition (*Table* 1). The diagnostic criteria for visual stress were not consistent, even in papers by the same authors. Criteria included subjective reports of perceptual distortions while reading[45]; immediate subjective benefit from overlays[46]; reported distortions on viewing a three cycles per degree square wave grating[41,42]; subjective reports on the Visual Processing Problems Inventory (VPPI)[47]; computerised visual search test performance under visually stressful conditions,[48] voluntary sustained use of overlays,[32,33,49] or reading rate on the WRRT that was 5, 8 or 10% faster with a chosen overlay than without.[38] In those studies that used the voluntary sustained use criterion, the duration of overlay use ranged from 3 to 12 weeks.[24,49] It should be noted that for those studies using the criterion of faster reading with the WRRT, it is not clear that small improvements in the rate of reading on the WRRT generalise to better or faster reading of naturalistic text.

In general, studies using disconnected or naturalistic texts comparing a chosen coloured overlay/lens with a placebo coloured overlay/lens reported improvements in reading for both conditions compared to baseline, but crucially no significant difference between the placebo and selected overlay.[45,49] One argument for this, based on a study by Tyrell *et al.*,[50] is that improvements in reading naturalistic text are only observed following prolonged periods of reading. Another argument is that some studies compared the

chosen colour with a closely related colour and it is for this that reason no difference was found.[49] However this calls into doubt the need for precision tinting claimed by some study authors.[5] Furthermore studies comparing a chosen tint with a complementary colour did not find any significant improvement in reading naturalistic text.[32,45]

Although there are claims for the superiority of Intuitive overlays over other tinting systems[5] there is only one head to head study that compares Intuitive Overlays with another system (Reading Rulers, Crossbow Education, UK) and this study is at high risk of bias (Table 1).[16] There are no head to head studies comparing Intuitive Overlays and Precision Tinted Lenses with the Irlen system though one study plots the chromaticity co-ordinates of Intuitive Overlays against those of Irlen overlays but on different scales that make comparison difficult.[32]

Specific studies

There are too many studies of Intuitive Overlays to describe them all in depth, however three studies were primarily designed as clinical trials and had some features of masked randomised control trials (RCTs) and are thus discussed in detail. In a double-masked, placebo-controlled trial using a crossover design,[49] sixty-eight reading-impaired children with an average age at recruitment of 12.2 years were recruited from two state schools, one private school for boys and the Dyslexia Institute in Leeds. The children recruited from schools were judged to be failing in reading by their teachers. The diagnostic criterion for visual stress was voluntary sustained use of overlays for at least 3 weeks. The children were issued with tinted spectacle lenses prescribed using the optimum colorimeter setting and with spectacles with a placebo colour that was just outside the range reported to improve perception. Inside the colorimeter, the participants did not see the precise tint of the lenses they were to be prescribed. For this reason, and because an interval of 1 month was left between testing and receiving the experimental or placebo lens, effective masking was maintained. Participants wore each set of tinted lenses for 1 month and were tested at the end of each period using naturalistic text. Participants also kept symptom diaries throughout the study. The study was limited by a high drop-out rate and a failure to analyse the data on an intention-to-treat basis.[51,52] There was no improvement in reading rate, accuracy or comprehension for 45 out of the 68 participants (66%) for whom data were available. Although analysis of the symptom diaries appeared to show a small benefit in favour of the optimum tint, data were only available for 36 out of 68 (53%) participants, meaning that this study was at high risk of bias (*Table* 1). The study also contains some contrary evidence where 22 participants preferred the precision tinted lenses, while 26 preferred the placebo control lenses. There was also evidence for novelty effects because 31 children preferred the first pair of glasses, but only 17 preferred the second pair (four expressed no preference).

Bouldoukian *et al.* adopted a different method, this time using overlays.[46] The optimum tint was compared with a pale yellow, placebo filter that was manifestly different from the Intuitive Overlays, thus neither participants nor experimenters were masked to the intervention being used. The placebo overlay was described to the participants as 'a new filter from the United States where it was thought to be a wonderful discovery' and marked with the words '*Research Model A16 Anti UV IR Filter. Made in the USA*'. A 4% increase in reading speed was reported with the prescribed overlay as compared to the placebo (a small increase of 4 wpm). However, the assumption that it is possible to match the placebo effect of the experimental intervention with an enhanced placebo is unfounded and as a result this study is at high risk of bias. The problems associated with enhanced placebos are discussed in more detail in the general discussion.

Mitchell *et al.* used a parallel-groups design.[45] Participants had dyslexia and reported visuo-perceptual distortions. Seventeen children received no lenses; 17 children received lenses based on the optimum intuitive colorimeter setting; 15 received lenses of a colour complementary to the chosen colour. All groups showed improvements in reading but there was no significant difference between the placebo and experimental lenses for reading speed, accuracy or comprehension.

Summary of intuitive studies

The results of the exploratory studies of the intuitive system have to be viewed with caution.[32,34] Multiple statistical comparisons and a flexible *post hoc* approach to data interpretation leaves these studies at high risk of producing false positive results.[53,54] As a result they should be seen as generating- rather than testing hypotheses. Furthermore, the three studies with features of masked RCTs[41,49] were each prone to bias. Improvements have been reported with prescribed overlays/lenses, but similar improvements are also found with placebo colours, questioning the need for precision in tinting lenses and overlays. The reduction in the use of overlays over time [32,34] also raises questions about the practicability of a technique that requires the assessment of all poor readers and the issuing of overlays to as many as 60% in order to identify a subset who may or may not benefit in terms of reading naturalistic text.

Irlen

The Irlen testing procedure involves a series of questions to probe for perceptual difficulties during reading (e.g., distortions and movement of text, light sensitivity, headaches,

eyestrain, tiredness, loss of concentration etc.), followed by a series of visual tasks involving counting lines/symbols within high-contrast pictures. Finally, one of ten (or a combination of) overlays is selected by the individual, following a series of pair-wise comparisons between coloured overlays placed over text. In *The Irlen Revolution* it is argued that only coloured overlays and lenses provided by the Irlen Institute are effective in treating Irlen syndrome (IS),[55] but no scientific evidence supports this claim.

Nine studies were identified that had features of RCTs[35,36,47,56–62]; two publications which report on the same trial are counted as a single study.[35,36] A consistent definition of IS was applied across these studies and the diagnostic procedures involved the use of the Irlen proprietary testing materials delivered by Irlen trained staff. The Irlen Reading Perceptual Scale and Irlen Differential Perceptual Schedule are frequently used to assess for IS, and consist of three sections: (1) the Irlen Reading Strategies Questionnaire (32 questions, 16 of which are related to reading strategies and reading behaviours such as skipping/re-reading lines, misreading words, losing place in the text, poor comprehension and slow reading; the remaining 16 questions relate to eye strain and fatigue whilst reading, including headaches, dry/itchy/burning eyes, blinking and squinting); (2) a series of visual tasks (e.g., counting squares in gridded rows; answering questions about distortions whilst observing visual images): and (3) an assessment of the extent to which performance on these visual tasks and on reading is improved by the use of coloured overlays. It is claimed that the subsections of this assessment have high internal validity and reliability although the evidence to support such claims comes from the unpublished literature on the Irlen website.[63]

Where available, the classification rates of IS in various samples of participants are displayed in Table S1. These rates ranged from 46%[30] to 96%[64] with a median of 66%. Two publications contained information about how long the overlays were consistently used by participants. Cotton *et al.* studied 60 participants, of whom 38 chose an overlay and 22 (37% of the original cohort) were still using their overlay 6 weeks later.[57] Ritchie *et al.* studied 61 children, 47 were diagnosed with IS (77%), and of these, 22 (36% of the original cohort) were still using their filter a year later.[62]

Specific studies

Because there has been a lesser volume of research of Irlen filters for reading, we have discussed all of the studies that we identified via our literature searches. In a 'pilot' RCT, 30 participants (aged 9–51 years) who tested positive for scotopic sensitivity syndrome and vision problems identified by standard optometric testing were randomly allocated to an Irlen filter treatment group (n = 11), a vision therapy group (n = 11, with three dropping out) or an untreated control group (n = 8, with five dropping out).[56] The Irlen filter group tried both a prescribed set of coloured lenses and a 'placebo tint' for 2 weeks, and then selected the lenses that made them feel 'most comfortable and subjectively improved their reading ability' for a further two weeks. Either three participants (their main text) or eight (their table 3) chose the placebo lenses. Neither the participants nor the experimenters were masked to group status, although for the Irlen group, the experimenters did not know during testing when the placebo or true Irlen filter was being worn. In the Irlen group, the symptom scores and scotopic sensitivity screening scores improved following the wearing of the tinted lenses, though only in the case of the Irlen lenses. However, there was no convincing evidence for improvements in any of the reading measures and optometric examination revealed that all subjects in this group still had significant visual issues despite wearing the filters. In contrast, for the vision therapy group, visual problems were resolved in seven of the eight cases and significant improvements were found in symptom scores, in scotopic sensitivity screening scores and for one test of reading which tested word recognition in context and speed). In the control subjects there was no change from pre- to post-testing for any of the outcome measures. On the basis of these pilot data, the authors concluded that Irlen lenses have negligible effects on reading.

Noble *et al.* examined the effects of Irlen overlays on reading rate, accuracy, fluency and comprehension, via teacher-led screening and assessment.[65] Participants were screened from two Grade 3 mainstream schools, rather than being referred or self-selected. Seventy-one participants were identified as having IS and competent reading ability, although reading ability varied considerably within the treatment and control groups. Children from one school (n = 31) were provided with coloured overlays for 3 months; a waiting list control group children from the other school (n = 40) received overlays after 3 months. Three children dropped out of the treatment group, however it is unclear at what point this dropout occurred. Furthermore, '..there were a small number of students in the study who did not use their overlays consistently', but no details are provided regarding the number or group status of these children. Significant improvements in reading rate, accuracy, fluency and comprehension were reported after 3 months of use (grade-equivalent score increases of 14–19 months), but no further improvements were reported at a six-month follow-up. The waiting control group showed no significant improvements during the first 3 months of the study (without overlays), but showed significant gains (grade-equivalent score increases of between 20–32 months) during the second 3 months of the study (with their overlays). Improvements in symptoms were also

reported when overlays were used. However, these results are subject to high levels of bias as a consequence of no group allocation concealment, no masking of participants and experimenters, no masking of the outcome assessment, and no placebo or control intervention group (*Table* 1).

Martin *et al.*[59] examined the impact of coloured overlays on 'process variables' (i.e., visual processing, phonological processing and working memory) as well as reading accuracy, reading comprehension, non-word reading, sentence comprehension. Three-hundred Tasmanian children, aged 11–12 years, were screened; 58 were selected based on reading difficulties and 62 were selected as reading at a level consistent with their age and IQ. Sixty-two percent of the poor readers were diagnosed with IS and prescribed coloured lenses. Children were tested prior to intervention, at a six-month post-test and at a follow-up session after 12 months for the following groups: those with reading difficulties who were diagnosed with IS and who were prescribed tinted lenses ($n = 20$), children with reading difficulties who were not diagnosed with IS ($n = 20$), and children without reading difficulties or IS ($n = 20$). No significant between-group differences were reported for any of the outcome measures. Study limitations include small sample sizes which reduces the power of the study and the long list of dependent variables; no IS control group and no placebo lens group; no random allocation to groups and no blinding to group status.

O'Connor *et al.*[58] studied the use of coloured lenses in children with poor reading ability. Out of a total of 600 mainstream school children (aged 8–12 years), teachers selected 105 children who were 'reading at least 18 months below grade level and whom they considered to have reading ability well below their abilities in other areas'. Students who displayed symptoms on the IDPS and 'a marked improvement' in reading with an overlay were classified as 'scotopic'. Consistent with the study above,[59] the prevalence of IS amongst their sample of poor readers was 64% (67/105). Twenty-five students (24%) were classified as 'non-scotopic' and the remaining 13 children (12%) showed scotopic signs but no improvement with an overlay and were excluded. Scotopic children were randomly assigned to one of four groups via a stratified randomisation procedure to ensure similarity at baseline: Group A ($n = 17$) received the prescribed coloured overlay; Groups B ($n = 17$) and D ($n = 16$) received a transparent overlay; Group C ($n = 17$) received a randomly coloured (non-prescribed) overlay. Non-scotopic children were randomly assigned to one of two groups: Group E ($n = 12$) were given transparent overlays and Group F ($n = 13$) were given a random (i.e. a non-prescribed) colour. Children were told that the overlays would 'make reading a little easier for them' and pre- and post-testing were separated by one week. The pre-test was omitted for Group D, to control for repeated testing effects. Children who received their prescribed colour (Group A) showed significant improvements of 6.6 months in reading rate, 6.9 months in reading accuracy, and a substantial 19.4 months in reading comprehension, while reading performance appeared to decline in the other groups over the week of the study. This study is, however, prone to bias because participants were not blinded to group status and there is also question as to how representative the 'poor readers' were, given the small group sizes.

A double-masked RCT was carried out by Ritchie *et al.*, following a 'dyslexia friendly schools' initiative by Inverclyde Council.[61] Ritchie *et al.* examined 75 children with below-average reading ability who were screened by an Irlen-trained practitioner. Fourteen children were unable to complete the Irlen screening tasks and were excluded. Of the remaining 61 children, 47 were diagnosed with IS (77%), again suggesting an extremely high prevalence rate. The study used a crossover design and compared reading rate (as measured via the WRRT) in 60 of the 61 children; 43 of the 60 children had IS, 14 did not. Three of the children with IS who were not masked because they were aware that their optimum colour was for assisting with reading were excluded from the main analysis. Also, in two of these three children it emerged that they had been using their filter for several days before the study commenced. Children were tested using the prescribed overlay, a placebo overlay of a complementary colour and a clear overlay. For both the IS and the non-IS group, overlays had no significant effect on reading rate. The three children who were non-masked showed significant improvements in the WRRT (78%, 58% and 15%), indicative of a placebo effect. However, this latter analysis has been criticised as selective and non-representative.[63] Forty-four children with IS were enrolled into a parallel-groups study where 22 were randomised to receive their optimum overlay while 22 received a colourless overlay. There was no significant difference between the two groups for any of the measures of reading or reading comprehension. Ritchie *et al.* followed up the same cohort one year later and found that 22 (30%) were still using their Irlen overlays or lenses, and of those available for follow up,[62] a deterioration in reading was evident. This is one of the strongest RCTs on Irlen overlays published to date, but is not without limitations. For example, the null result could be a consequence of an inadequately powered design. Furthermore, given the exclusion of three children who showed positive effects, this study may be subject to reporting bias.

Robinson and Foreman are frequently cited as supporting the use coloured overlays as a treatment for reading difficulties[35,36]; however, the results of this study do not support this claim. One hundred and thirteen subjects aged 9–13 years with poor reading and IS recruited from the

'Special Education Centre' at the University of Newcastle, Australia were randomly allocated to one of three experimental groups: an optimum (diagnosed) tint group ($n = 38$), a blue tint group ($n = 41$), or a placebo tint group ($n = 34$) (i.e., a similar colour to the optimum tint but which did not ameliorate visual symptoms). A no-treatment control group ($n = 35$) with poor reading but without IS was recruited from two local schools, introducing a potential recruitment bias. Although this study is described as a 'long-term, placebo-controlled study' lasting for 20 months, it was only placebo-controlled for the first 3–4 months, after which all participants used their optimum tint. At the start of the study there was no significant difference between the groups for any measure of reading although scores in the optimum tint group were lower than in the other groups. At the end of the initial 3–4 month period all groups had improved but crucially, there was still no significant difference between the groups. Although attention has been drawn to the bigger improvement in comprehension in the optimum tint group, it is not statistically correct to make within-group comparisons in a parallel groups study.[69]

Finally, Tyrell *et al.*[50] published an exploratory study of 60 children aged 8–16 years. The sample consisted of a series of small subgroups, based on ability, ranging from above average to well-below average. However subsequent statistical analysis focussed on comparing children who 'chose coloured overlays' against children who 'chose clear overlays' rather than the subgroups defined in the methods section. Participants were tested reading aloud naturalistic text of their own choosing for 15 min. There was no immediate effect of coloured overlays; however, after 10 min, the children who chose a coloured overlay read significantly more syllables per minute, with their overlay than without and reported more symptoms of visual discomfort and tiredness when reading without their overlay. However, these differences were very small and it is debatable whether this was clinically significant. There was no placebo control group and it is not clear if the division of the reading task into 5-min segments was a *post hoc* decision or a predefined means of analysis. At best, therefore, this study should be seen as hypothesis-generating rather than hypothesis-confirming.

Unpublished data

An additional six unpublished postdoctoral theses were identified via our searches[30,64,66–69] (*Figure 1*). Anderson is not considered here as it was purely observational and there was no control group.[30] The results from the other five 'grey' items are now considered.

Donovan studied 83 children with ages from 10–15 years (average ~12 years) who were diagnosed as reading disabled.[66] Each child tested positive for IS. The pattern of results obtained was extremely mixed with the prescribed overlay improving some performance measures but reducing others, or having beneficial effects in readers of a certain level but negative effects in readers of a different level. Interestingly there was no significant interaction between IS-level (mild, moderate or severe) and the effect of the overlay, even for those variables for which the overlay appeared to have an effect.

Mason[67] studied 30 university students who demonstrated low reading ability and whose symptoms of IS were in the severe range. The participants were self-referrals to the University's learning assistance centre. Participants were divided into three groups; ten received coloured overlays, ten received reading instruction and the remainder received no treatment. There was no between-groups difference in relation to the 'change in reading rate' exhibited from pre- to post-treatment testing indicating that coloured overlays were no more beneficial than reading instruction and no better than no intervention at all. However, the author did acknowledge that the study was underpowered given the small number of participants per group.

Faraci[64] examined 26 children with an average age of 9 years who tested positive for scotopic sensitivity syndrome. The children were divided into two groups who either did or did not receive overlays, although both groups received the same instruction in reading. The overlay group were asked to use their overlays for all school- and home-based reading and homework activities. After 3 months, reading fluency was significantly higher amongst the overlay group, but no statistically significant difference was evident for phonics or for reading accuracy or comprehension. A major drawback of this study is that the author assumed that the two groups (overlay and no overlay) were matched in their baseline reading performance.

Morrison examined whether individuals diagnosed with IS showed differences in reading fluency and eye movements when they read with and without coloured overlays.[68] Participants ($n = 24$) were mainly undergraduate psychology students who did not report a reading problem or a reading disability but who were IS-positive. The results revealed no difference in reading fluency of curriculum-based material, or associated eye movements when the optimum coloured overlay was compared to a clear or randomly coloured overlay.

Adams compared reading on a computer screen in 32 children (aged 12 to 14 years of age) with a clear overlay compared to with a chosen overlay.[69] The children were selected for the study on the basis that they reported perceptual distortions and that the chosen overlay removed the distortions. The differences in scores for the clear versus chosen overlay conditions were not statistically significant.

Overall, the 'grey' literature does not support the use of coloured overlays and lenses to improve reading

performance. Generally the study quality was found to be acceptable but many of the studies were underpowered owing to too few participants being recruited or dividing the participants into too many groups.

Summary of Irlen studies

The use of Irlen lenses and overlays to improve reading in individuals with IS cannot be endorsed on the basis of the studies in the peer-reviewed or 'grey' literature. The two trials at lowest risk of bias failed to show any improvement in reading outcomes when using prescribed coloured overlays and lenses.[35,36] Importantly, the use of Irlen procedures has led to a high percentage of both normal readers and poor readers being diagnosed with IS and related perceptual phenomena across studies. Although these prevalence rates are consistent with Irlen's original predictions[2] they have been criticised as being vastly over-inclusive.[14]

ChromaGen/Harris lenses

ChromaGen$_{TM}$ spectacles or contact lenses[70] were developed by David Harris as a treatment for congenital colour vision disorders to allow the subjective appreciation of a wider range of colours. On the basis of anecdotal reports from patients with colour vision deficiency that the lenses improved the clarity of text, and that colour improves reading performance of individuals with visual stress,* they were applied to the treatment of dyslexia. When used in colour vision deficiency, one lens (usually a contact lens) is worn on the non-dominant eye. In reading difficulties, the right and left eyes are assessed independently so that subjects may receive different coloured lenses (contact lenses or spectacles) for each eye. Thus, although the original set of spectacle lenses comprised eight colours (substantially less than Intuitive and Irlen systems), there is obviously a much larger number of combinations because the optimal colour for the two eyes may differ. Indeed, it appears that around 50% of individuals fitted with ChromaGen lenses are prescribed different colours for each eye.[71] Harris Lenses are similar to ChromaGen lenses (i.e., they involve the same number of colours and are prescribed via the same procedures) but Harris Lenses have a surface mirror coating that reflects light more evenly across the spectrum. Consequently, they appear more natural to an outside viewer while preserving transmission qualities.[72] The mechanism by which ChromaGen or Harris filters work to help reading is not well established, nor is the reason why the two eyes may require a different colour. In relation to the latter, one suggestion is that different coloured lenses may differentially affect the rate of neurological transmission in the two eyes, akin to the use of neutral density filters in the Pulfrich phenomenon.[73]

By comparison with other colour systems, the ChromaGen/Harris system has not been subjected to the same volume of scientific scrutiny; only four papers were identified in the peer-reviewed literature that assessed some measure of reading. The studies in this area have compared the ChromaGen system to placebo lenses, where participants are typically told there is an invisible tint, or to control (no lens) conditions. In head-to-head trials published in the peer-reviewed literature, the ChromaGen system has only been compared to the Dyslexia Research Trust (DRT) system which comprises blue and yellow lenses.[74,75]

Specific studies

Only four studies have examined ChromaGen/Harris lenses and all are described here. Following a pilot investigation of ten participants (published in the Optical Press, rather than in the peer-reviewed literature) in which it was claimed that ChromaGen lenses out-performed coloured lenses from the Intuitive Colorimeter, Harris and MacRow-Hill compared the ability of ChomaGen contact lenses and placebo contact lenses carrying a 'light blue' handling tint to improve reading fluency in adults with dyslexia.[76] The study was described as a 'double-masked' trial; however, it is highly likely that participants (who had responded to media interest) were aware of the difference between the two types of contact lenses. Nevertheless, the research team who carried out the outcome assessments were masked to group status. Participants had a formal diagnosis of dyslexia from an educational psychologist and were willing to wear contact lenses, although it is not clear if participants suffered from visual stress. Fifty-three participants started the trial but six failed to complete the study because they were either unable to tolerate contact lenses, fulfil the minimum reading requirement or were unwilling to complete the testing which took place on the same day. Across all participants, there was a statistically significant increase of 12 wpm (a 15% increase) relative to the baseline reading rate, compared to a significant increase of 7 wpm with the placebo lens (an 8% increase). The improvement in reading rate with the Chromagen lenses was statistically significant relative to both the baseline reading rate and the improvement seen with the placebo lenses, however, the improvement seen with the placebo lenses was also statistically significant. One important aspect of the results was that participants who received the ChromaGen lenses before the placebo lenses showed a statistically larger improvement in reading rate compared to those who received the placebo lenses first. This suggests that novelty effects may have exerted an influence on the results. There are other serious issues of external validity. For example, participants were recruited in response to publicity in the media about the possible benefits of colour.

Cardona *et al.*[71] compared ChomaGen spectacle lenses with placebo lenses in 56 teenage children. The placebo lenses were clear but, as in the study by Harris and MacRow-Hill,[76] children were informed that lenses had a new invisible tint that provided the same effect as coloured lenses. It seems unlikely that this measure would have controlled for placebo effects associated with coloured lenses. Overall placebo lenses and ChromaGen lenses improved reading rate relative to the control condition where no lenses were worn. However, there was no improvement in reading speed with ChromaGen lenses over that seen with the placebo lenses. The results for reading accuracy showed a borderline significant benefit of the ChromaGen lenses over the placebo, although the magnitude of the effect size was not stated.

Two studies compared Harris lenses with blue or yellow lenses from the DRT in a head-to-head fashion.[74,75] These studies included the same group of participants comprising 73 delayed readers who reported that a filter (Harris or DRT) helped them see text more clearly. A positive feature of these studies was the lack of external bias because subjects were recruited from mainstream state primary schools. Unfortunately, because of a prior assumption that the treatment works, there was no placebo control group. Treatment fidelity is also questionable, as there was no mention of whether all of the children who chose a filter continued to use it for the full three-month period. The groups were well matched on spelling and reading at baseline. After 3 months, both groups had improved their reading and, to a lesser extent, their spelling: importantly, there was no significant difference between the two groups in the improvement in reading or spelling scores. The DRT group did improve their speed of reading non-words more than the Harris lens group, but neither group improved in their ability to read irregular words. Although the conclusions of these papers was that both systems improved reading ability in children with reading delay, the added time and effort required to decide upon the optimal tint using the Harris filters compared to the DRT filters (where either a yellow or blue tint is issued) was considered to give the DRT system an advantage. Because no control group or placebo lenses were used in these studies, it is impossible to know the role that placebo effects played in the improvements seen.

Summary of ChromaGen/Harris lens studies

The results from the small number of studies assessing the effectiveness of the ChromaGen/Harris filter system for patients with reading difficulty collectively suggest that the system may deliver better reading performance than placebo lenses. However, this evidence is rather weak principally because it is unlikely that the participants in these studies were well masked to treatment groups and so they would have known the difference between the two types of lenses (i.e., tinted vs clear). In head-to-head studies with the DRT lens system, Chromagen/Harris lenses showed comparable changes in reading performance to the DRT system. However, the value of the 'benefits' that were claimed in these studies is difficult to establish because of the absence of placebo-lens or no lens (control) groups.

Non-mainstream (other) studies of colour

This section evaluates studies of less well known or non-commercial coloured overlays or filters. Because of the diverse nature of these studies it is difficult to create a coherent narrative so we have commented on each study individually. Three publications used one colour: one blue[77]; two yellow,[18,78] one used two colours (blue and red)[79] and the remaining used less than 10 colours except one which used 15 colours.[80] Two publications that compared blue or yellow overlays with ChromaGen are considered under the Chromagen section.[74,75]

Of the 12 publications, seven utilised a crossover design, and five used a between-subject, parallel design. All were identified as having at least one design aspect that was judged to be at 'high' risk of bias (*Table* 1). Christenson *et al.* examined the effect of a blue filter in 16 children with dyslexia[81] who were randomly allocated to a blue lens group or a no lens group. The children initially assigned to the no lens group received blue lenses 2–5 weeks later. The wearing of a blue filter did not significantly affect the reading level or speed. Palomo-Alverez and Puell studied the use of a yellow filter.[78] Poor readers aged 9 to 11 years old, (defined as 'poor readers without dyslexia and minimal refractive error') were randomly assigned to a yellow filter group worn for 3 months for school and homework[45] or a no treatment, control group.[36] There was no statistically significant difference between the two groups in accommodation, symptom scores or reading speed with the filter.

The study by Ray, Fowler and Stein is difficult to evaluate because the publication appears in conference proceedings and insufficient information is presented to assess the risk of bias, in particular in relation to random sequence generation and allocation concealment.[18] The study examined the use of a single, yellow overlay on word reading accuracy in 38 'severely disabled readers' aged 7–14 years. Some of the 38 'severely disabled readers' used the yellow filter and some used a cardboard sheet with a rectangle cut out that revealed one line of text only. There was also a blue overlay placebo control that was not reported. Reading was assessed using naturalistic text after 3 months. Detailed quantitative data such as means, standard deviations, confidence intervals and effect sizes are lacking. Although this study is sometimes described as being 'double masked'[13,75] it was at high risk of bias primarily because children would

have been aware of which intervention they were receiving (*Table* 1). There are also issues of external validity because children were recruited from the Dyslexia Research Trust, which promotes the use of blue and yellow overlays as an intervention for reading impairment.

Iovino *et al.* studied 60 children in total, comprising four groups of 15, categorised as reading/spelling/arithmetic disabled, reading/spelling disabled, arithmetic disabled and those with ADHD.[79] Each group viewed text through a blue overlay, red overlay and no overlay in a single session. While a significant improvement in comprehension accuracy was reported, there was no difference in reading rate or accuracy between the three environments and there was no significant group-to-colour interaction in reading rate.

Sawyer *et al.*[82] studied 86 students from 7–15 years of age from their caseload of those with specific learning disorders. From a cohort of approximately 300, 110 reported a positive reaction to four coloured overlays (red, green, blue and yellow). One-hundred and eighty five similar students from a nearby town served as a control group. After one and half school terms there was no significant between group improvements in confidence in reading, interest in reading, or in the amount read.

Gole *et al.*[83] recruited 24 students with 'dyslexia'. Thirteen were allocated to the treatment group on the basis of their positive subjective response to six coloured lenses presented in random order. The remaining 11, all of whom had a negative subjective response, acted as controls and received a clear lens for one term followed by a randomly selected tinted lens for two terms. Reading was assessed at the start of the study and at the end of each school term. There was no statistical difference between the absolute value or change in reading ages for rate, comprehension, or accuracy of reading in the treatment and control groups.

Menacker *et al.*[84] studied 24 children (8–12 years) with dyslexia. All children read passages of naturalistic text using four coloured lenses. Half of the children used 0.12 log unit density lenses and the other half 0.30 log unit density lenses. All children read similar passages using the four coloured lenses, a neutral density filter and with no filter. There was no significant change in reading error or rates attributable to lens colour or density.

Saint-John and White[85] studied 11 children (aged 11–12 years) with specific reading difficulty and 11 controls who had no difficulty. The children chose one of six coloured overlays, which was mounted into spectacles. The dependent variables were reading accuracy and speed on four passages of non-standardised text. All children read with their selected colour, with the polaroid and with no lens. Colour transparencies did not improve reading any more than a polaroid or an empty frame.

Evans *et al.*[86] explored the hypothesis that the effect of coloured overlays was mediated by treating pattern glare. They described two studies. In the first, they asked 151 optometry students to look at a pattern glare stimulus consisting of a high-contrast striped pattern with a spatial frequency of 4 cycles/degree. Symptoms such as bending of lines, blur, diamond shaped lattices, fading, flickering, shimmering or wobbling, glare or dazzle, or colours were present in 149. Five of these individuals with high scores and six individuals with low scores were assessed with and without eight coloured overlays, however, there were no significant differences in the search time on this task regardless of the overlays used.

Vidal-Lopez studied 54 children aged 12–14 years.[80] Twenty-seven were diagnosed with visual stress according to the Wilkins' pattern glare test and assessment questions based on the Irlen Questionnaire. The remainder acted as a control group. The visual stress group selected the coloured filter that ameliorated their perceptual distortions while the control group was given a filter chosen randomly from among the 15 supplied by Panoptica (Delt Orgaz, S.L., Barcelona, Spain; www.panoptica.es/). All subjects read single Spanish words and a Spanish equivalent of the WRRT whilst wearing the coloured lens and a clear lens. Both groups read slightly faster with the coloured filter but the difference was only statistically significant in the non-visual stress group. The authors explored whether this improvement might be due to improved motivation by measuring response criterion to a psychophysical test, again with a coloured and a clear lens. They found that increases in reading speed were associated with changes in response criterion suggesting that participants had become less conservative observers, however, this difference was not statistically significant. The authors argue that because the improvement in reading was greater in the group without visual stress and because it was associated with changes in response criterion, placebo effects were the most likely explanation. According to the visual stress hypothesis, greater improvements would be expected in the group with visual stress and these would not be associated with changes in response criterion; however, this was the opposite of what was observed. The study could usefully be repeated with a larger sample size.

Francis *et al.*[87] used blue, red, green and yellow overlays in 35 children with reading difficulties (10–14 years) and compared them to 27 children who received no intervention. Of the 35 children, 23 (66%) continued to use them for a whole term. The other 12 were resistant to using the device or the teachers felt it was having an adverse effect. There were no significant differences between the groups in the improvement in reading age or reading speed.

Summary of non-mainstream colour studies

None of these studies contain strong evidence that the use of coloured overlays or lenses leads to benefits in the in

measures of reading in individuals with reading difficulties and/or visual stress.

General discussion

Randomised controlled trials and systematic reviews of those trials are considered the best available form of evidence for therapeutic interventions. The key feature of a systematic review is that all studies are appraised according to the same template for assessing the risk of bias but only those at low risk of bias are included in the final analysis. In general, the studies reviewed here were at high or uncertain risk of bias but in order to appraise the literature as it currently stands we adopted an inclusive approach. Consequently, we have included numerous studies excluded by other systematic reviews. For example the systematic review by Galushka *et al.*[17] only included two studies.[35,45]

Approaches to reviewing literature frequently involve a meta-analysis, which combines data derived from a systematic-review. Thus, every meta-analysis should be based on an underlying systematic review, but not every systematic review leads to a meta-analysis.[88] Based upon our view that a large majority of the literature we reviewed is at high risk of bias it is not clear that this field of research is ready for meta-analysis.

Many of the studies we reviewed relied on *p*-values to support their outcomes; for example, claiming a result was statistically significant if p was <0.05. However, it is important to remember that the *p*-value is just the final step in the design and execution of a study. In practice, decisions made earlier in experimental design or in the analysis of the data are more important to the outcome and the idea of the 'risk of bias' tool is to give greater weight to the behaviour and practices that lead to the statistics. If there are problems with those behaviours and practices (which there were with almost all of the studies we reviewed) a *p*-value with an arbitrary value of 0.05 or less adds nothing useful.[89,90] For these reasons we have not quoted *p*-values.

This systematic review of the literature leads to the conclusion that there is little evidence to support the use of coloured filters (overlays, spectacle lenses or contact lenses) to improve reading. Although each of the colour systems has been subject to varying levels of empirical evaluation, the results of this review do not differ according to system-type, and the majority of the limitations identified apply to all systems. Our results are consistent with the results of previous literature reviews, including a recent review by Albon *et al.*,[6] which concluded that the available evidence was too low in quality to reach firm conclusions about the effectiveness of coloured filters for reading disability. Similarly, Uccula *et al.*[9] concluded that the issue remains 'controversial' and 'not settled'. Handler and Fierson found a 'continued lack of definitive evidence of the effectiveness' of coloured lenses and filters.[8] A 2009 review prepared for the New Zealand Ministry for Health concluded that 'there is not a sufficient body of evidence as yet to state that coloured filters or overlays improve the reading ability of those with reading difficulties'.[91] An evidence and consensus based clinical practice guideline recently published in the German literature concluded that Irlen lenses should not be used in the treatment of reading and spelling disorders in children or adolescents.[92]

Previous reviews of this literature have been criticised for considering the literature without taking into account which particular colour system was under investigation and whether it was reading difficulty itself that was being treated or the co-morbid condition visual stress.[5] In this review, we have considered the studies separately for each of the main colour systems and we find no evidence to support the use of any system to aid reading. Our analysis also shows that even if only published research using the inconsistent diagnostic criteria for visual stress are selected, there are no studies at low risk of bias to support the use of coloured overlays or lenses to aid reading.

It has been argued that coloured lenses and overlays are also being used to treat the symptom complex of visual stress. Although we specifically only searched the literature for studies of whether colour overlays or lenses impact on reading outcomes, some of the studies we reviewed also made reference to changes in symptoms. For example Wilkins and colleagues looked at symptom diaries but data were only available for 53% of study participants thus precluding any meaningful analysis of the data.[49] Also, Mitchell and colleagues used the IDPS which contains questions about symptoms. In this parallel groups study there were significant improvements in the IDPS scores in both the chosen lens and the placebo lens group but no difference between the groups.[45] While our review focused on colour and reading, we failed to we see evidence that colour impacts positively on symptoms. Ultimately, a large scale RCT using a validated symptom questionnaire is required to answer this question.

It is important to stress that the lack of evidence which we and previous reviewers have identified does not in itself prove that colour has no effect on reading; lack of evidence is not evidence for a lack of effectiveness. On the surface, this statement could be interpreted as tacit support for a continuation of the practice of issuing coloured filters and lenses while the necessary evidence is being gathered. However, our main finding, consistent with several previous reviewers is that the quality of the available evidence is sufficiently low such that, despite the many anecdotal claims and often powerful testimony of patients, we have serious reservations about this practice. Published studies on the topic first appeared over 20 years ago so this field of research is not new. We believe the onus is on the

proponents to increase their efforts to gather the evidence to support this clinical practice. Below we draw attention to the main limitations of previous studies and we make some suggestions about how future studies might be conducted to avoid or counter such criticisms.

Most of the studies we reviewed were not well designed, there was little evidence of a pre-study protocol, studies were often under-powered and all had areas of bias that were either 'high' or 'uncertain' (*Table* 1). Many publications contained errors between tables and text, and suspect statistical analysis was frequently observed including absence of a pre-trial specified statistical approach, uncorrected multiple tests on the same data sets, effect sizes not reported and missing descriptive statistics.

There was a common failure to consider how research participation effects might influence results.[93] For instance, many studies showed evidence of novelty effects. The latter refers to an intervention that is new and exciting and which, consequently, may improve motivation and produce initial positive effects that diminish over time.[94] For example, in[49] 31 participants preferred their first filter whereas only 17 preferred the filter they received second. Similarly in a crossover trial of ChromaGen lenses, participants who received the experimental lenses before the placebo lenses showed a bigger improvement in reading rate.[76] Furthermore, an uncontrolled field trial of Irlen lenses showed improvements in reading during the first 3 months of use but no improvements in the subsequent 3 months.[65] Although this result was attributed to participants reaching their grade level, another explanation is that participants became less impressed by their overlay with time. Arguably, the high rate of attrition observed in many studies may also reflect novelty effects.

The act of being observed by the experimenter may also enhance performance; this is known as the Hawthorne effect.[95,96] When participants in both the experimental group and placebo group improve more than would be expected due to normal maturation, Hawthorne effects may be the most plausible explanation. Related to this, it was striking that trials that were well masked showed no statistically significant improvement in reading with a chosen colour compared to a placebo condition.[35,45,49,61] On the other hand unmasked studies that compared chosen colour with no overlay or clear overlay,[32–34,40,43,47,50,65] or a card with a rectangular slot cut out[18] often reported significant effects on reading. This difference in outcome between masked and unmasked studies points strongly to placebo effects.

Three studies attempted to control for placebo effects using enhanced placebos rather than trying to mask participants and experimenters to the intervention.[46,71,76] In such studies steps were taken to enhance the placebo effect of the control filter by describing it as a 'special' or 'wonderful discovery'. Implicit in this, is the assumption that the placebo effect of the experimental intervention can be accurately quantitated and that the placebo effect of the control intervention can be precisely modified to match it. The placebo effect is not sufficiently well understood to allow this. Indeed selecting the chosen tint involves a more prolonged relationship with the practitioner and a richer therapeutic ritual, both of which are powerful drivers of the placebo effect.[97–99] Hence, such 'enhanced' placebos are not recommended for future research. It is important to acknowledge that incorporating a well-masked placebo control condition that comprises identical diagnostic and therapeutic rituals is particularly difficult in trials of coloured lenses and overlays. Nevertheless, the use of the Intuitive Colorimeter has the potential to allow masking because during the assessment, participants do not see the actual lens they will ultimately receive.[49]

Some researchers have claimed that it would be unethical to include a placebo control group.[74,75] This reflects a prior assumption that treatment with coloured overlays/lenses is effective. Our review shows that in those studies that were well masked there was as much improvement in the placebo control group as the experimental group.[35,45,49,61] As a result we do not consider it unethical to include a placebo control group. Indeed, the ethics of organising further trials that are at high risk of bias because of the lack of a placebo control group also needs to be considered.

One perspective is that, even if the benefit of colour upon reading stems purely from the placebo effect, the most important aspect is that reading has improved and the source of that improvement is of lesser importance. While we understand this logic, we disagree given that coloured overlay or lens therapy can have a substantial financial cost for the patient or their parents, and may delay identification of the real reason(s) for reduced reading ability, hence stalling appropriate remediation or management approaches.

Most studies adopted a crossover or within-subject design. Since participants act as their own controls, such studies are less prone to confounding at baseline[100] and the paired data they produce add to the statistical power. Studies of this type are generally considered suitable for assessing short acting or temporary interventions for chronic conditions.[100,101] The principle drawback of crossover studies is their vulnerability to attrition because there has to be sufficient time to allow all participants to receive both treatments. For this reason, a longer-term, parallel-groups study might be more suitable for assessing the effect of coloured lenses and overlays on reading performance. Parallel-arm designs require a substantially larger number of participants because the different groups need to be carefully matched on variables such as age, gender, reading-skill at baseline and the rate of attrition. Nonetheless, the

advantages of parallel arms studies far outweigh these practical disadvantages. Based upon pilot data using the reading test selected and the test-retest variation of the test, power calculations can be conducted in advance of the study to establish the appropriate sample size, taking account the likely attrition.

The ability to generalise the data to the wider population (in other words the external validity) should also be considered. Participants recruited from specialist clinics may not be representative of the general population of poor readers and furthermore participants recruited from clinics such as those at the Institute of Optometry[46] or Dyslexia Research Trust[18] may have specifically sought treatment with overlays because of their prior belief that treatment with colour is an effective therapy. An additional problem is that they may know their preferred colour, making masking difficult. In Ritchie *et al.* it was striking that the two participants who showed biggest improvements in reading were aware of their chosen colour.[61] Ideally, to ascertain the effect of colour on reading, participants should have no prior exposure to the use of coloured lenses or overlays, and be drawn from unselected samples of children and adults.

In terms of applicability of the results to real-world reading, the external validity of the reading tests themselves also needs to be considered. Even if the WRRT is a useful diagnostic test for visual stress, unless it can be shown that improvements in reading the WRRT translate to reading naturalistic text of the sort that is encountered in everyday life, it is, by itself, an unsuitable outcome measure. Age-appropriate normative data should be available for each reading test. The test-retest repeatability of the test should also be well established so that the impact of any change following the use of coloured overlays/lenses can be considered alongside the normal variation in baseline measures of reading performance. Also, since there are different measures of reading skill, outcome measures of reading should not be restricted to any one particular aspect (e.g. speed, accuracy or comprehension); rather it is suggested that all of these aspects should be represented in the outcome measures used.

Since poor reading may have a variety of causes, studies of the impact of coloured lenses or filters on reading performance should rule out other possible causes in study participants.[102] Many of the studies we reviewed included eye examinations of their participants prior to the issuing of any colour intervention. This could be considered good scientific practice in order that the impact of coloured overlays/lenses on only the 'target' condition is assessed. Although there is ongoing debate about the frequency with which refractive and oculomotor anomalies account for poor reading performance[103–105] they are likely to be important confounding variables that must be controlled for.

Many of the studies we reviewed sought to establish if coloured overlays/lenses make an immediate difference to reading. We suggest that the effect of coloured overlays/lenses should be examined over a period of time that is not less than 3 months, and ideally for up to one year. This will enable researchers to observe any decay in the frequency with which the coloured overlay/lenses is/are used, and to assess any longitudinal changes in reading performance relative to age-appropriate norms.[106] Follow-up over longer periods will also help to eliminate the impact of novelty effect upon study outcome.

The proportion of cases diagnosed with visual stress ranged from 46–96% for Irlen and 13–88% for Intuitive (Table S1). One would expect that prevalence rates to be reasonably consistent between unselected populations. Furthermore, based on the rapid discontinuation of use seen in many studies, it can be argued that the diagnostic procedures currently in use would appear to produce a large number of false positives. The lack of constant diagnostic criteria makes it difficult to be sure that the same condition is being investigated and treated across different studies. Researchers with an interest in visual stress need to agree on the diagnostic criteria for the condition which will not only enable robust epidemiological studies to ascertain the prevalence but also examination of whether coloured lenses or filters have an impact on reading performance in individuals who test positive for that condition.

To avoid claims that statistical analyses were conducted *post hoc*, in addition to establishing, *a priori*, the outcome measures used to determine whether colour has aided reading, researchers should set out in advance which statistical tests will be applied. This measure alone has been shown to reduce the number of trials for which a positive effect is reported.[107] Although this statistical approach does not preclude the reporting of exploratory analyses *post hoc*, the results of such analyses should be seen only as hypothesis-generating rather than hypothesis-confirming.

Proponents of the use of coloured lenses or filters to treat visual stress have attached importance to the results of fMRI studies.[5,108] A detailed account of these studies is beyond the scope of this review. However, ignoring the problems of interpretation of this kind of study which can be at high risk of producing false positive results,[109,110] ultimately it has to be shown that colour improves the behaviour in question, in this case reading.[111] For this reason we have not considered the results of neuro-imaging studies here.

Conclusion

This systematic review suggests that whilst many studies report improvements with coloured lenses or filters, the effect size is generally small and/or similar to

the improvement found with a placebo condition. The vast majority of studies in each area are subject to high or uncertain risk of bias in one or more key aspects of study design or outcome. Studies which are less at risk from bias generally offered less support for the benefit of colour on reading ability. For these reasons, in common with previous reviews of the literature, we conclude that the use of coloured overlays and lenses to ameliorate reading difficulties cannot be endorsed. From the evidence reviewed, placebo, Hawthorne and novelty effects provide the most likely explanation for the benefit which many individuals report.

Disclosure

The authors declare that this review was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Meares O. Figure/background, brightness/contrast and reading disabilities. *Visible Lang* 1980; 14: 13–29.
2. Irlen H. *Reading by the Colors: Overcoming Dyslexia and Other Reading Disabilities Through the Irlen Method*. Avery Pub. Group: Garden City Park, NY, 1991; p. 195.
3. Wilkins AJ. *Reading Through Colour: How Coloured Filters can Reduce Reading Difficulty, eye Strain, and Headaches*. John Wiley & Sons: Chichester, West Sussex, UK; Hoboken, NJ, 2003. p. 156.
4. Wilkins A. Coloured overlays and their effects on reading speed: a review. *Ophthalmic Physiol Opt* 2002; 22: 448–454.
5. Wilkins AJ, Allen PM, Monger LJ & Gilchrist JM. Visual stress and dyslexia for the practicing optometrist. *Optom Pract* 2016; 17: 103–112.
6. Albon E, Adi Y & Hyde C, West Midlands Health Technology Assessment Collaboration. *The Effectiveness and Cost-Effectiveness of Coloured Filters for Reading Disability: A Systematic Review*. Studley: West Midlands Health Technology Assessment Collaboration, Department of Public Health and Epidemiology, University of Birmingham, 2008.
7. Fletcher JM & Currie D. *Vision Efficiency Interventions and Reading Disability [Internet]*. Perspectives on Language — Winter Edition 2011. Available from: http://www.onlinedigeditions.com/display_article.php?id=625330.
8. Handler SM, Fierson WM, Ruben JB *et al.* Joint technical report – Learning disabilities, dyslexia, and vision. *Pediatrics* 2011; 127: e818–e856.
9. Uccula A, Enna M & Mulatti C. Colors, colored overlays, and reading skills. *Front Psychol* 2014; 5: 833.
10. Aleci C. Colored filters and DYSLEXIA. A quick gliding over myth and (possible) reality. *Neuro Ophthalmol Vis Neurosci* 2014; 1: 18–21.
11. Hyatt KJ, Stephenson J & Carter M. A review of three controversial educational practices: perceptual motor programs, sensory integration, and tinted lenses. *Educ Treat Child* 2009; 32: 313–342.
12. Howard-Jones PA. Neuroscience and education: myths and messages. *Nat Rev Neurosci* 2014; 15: 817–824.
13. Stein J. Dyslexia: the role of vision and visual attention. *Curr Dev Disord Rep* 2014; 1: 267–280.
14. McIntosh RD & Ritchie SJ. Rose-tinted: The use of coloured filters to treat reading difficulties. In: *Neuroscience in Education: The Good, the bad, and the Ugly* (Della Sala S & Anderson M, editors), Oxford University Press: Oxford, New York, 2012; p. 385.
15. Henderson LM, Taylor RH, Barrett B & Griffiths PG. Treating reading difficulties with colour. *BMJ* 2014; 349: g5160.
16. Smith L & Wilkins A. How many colours are necessary to increase the reading speed of children with visual stress? A comparison of two systems. *J Res Read* 2007; 30: 332–343.
17. Galuschka K, Ise E, Krick K & Schulte-Körne G. Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLoS ONE* 2014; 9: e89900.
18. Ray NJ, Fowler S & Stein JF. Yellow filters can improve magnocellular function: motion sensitivity, convergence, accommodation, and reading. In: *Clinical and Basic Oculomotor Research: In Honor of David S Zee [Internet]*. New York, NY, USA: New York Academy of Sciences, 2005; pp. 283–293 (Annals of the New York Academy of Sciences, Vol. 1039). Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2005-07863-028&site=ehost-live.
19. Loew SJ, Rodríguez C, Marsh NV, Jones GL, Núñez JC & Watson K. Levels of visual stress in proficient readers: effects of spectral filtering of fluorescent lighting on reading discomfort. *Span J Psychol* [Internet]. 2015 [cited 2015 Aug 12]; 18. Available from: http://www.journals.cambridge.org/abstract_S1138741615000591.
20. Monger LJ, Wilkins AJ & Allen PM. Pattern glare: the effects of contrast and color. *Front Psychol* 2015; 6: 1651.
21. Rosenfield M, Rosenlicht T, Jones YH & Simon J. Effect of coloured overlays on Computer Vision Syndrome. *Optom Vis Perform* 2015; 3: 133 –137.
22. Wilkins AJ, Jeanes RJ, Pumfrey PD & Laskier M. Rate of Reading Test: its reliability, and its validity in the assessment of the effects of coloured overlays. *Ophthalmic Physiol Opt* 1996; 16: 491–497.
23. Wilkins A & Lewis E. Coloured overlays, text, and texture. *Perception* 1999; 28: 641–650.
24. Northway N. Predicting the continued use of overlays in school children–a comparison of the Developmental Eye Movement test and the Rate of Reading test. *Ophthalmic Physiol Opt* 2003; 23: 457–464.
25. Allen PM, Gilchrist JM & Hollis J. Use of visual search in the assessment of Pattern-Related Visual Stress (PRVS) and

its alleviation by colored filters. *Invest Opthalmol Vis Sci* 2008; 49: 4210–4218.
26. Higgins JPT, Altman DG, Gotzsche PC *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; 343: d5928.
27. Cotton MM & Evans KM. Parents' and children's expectations about Irlen (tinted) lenses. *Percept Mot Skills* 1994; 78: 1387–1390.
28. Robinson GL & Conway RN. The effects of Irlen colored lenses on students' specific reading skills and their perception of ability: a 12-month validity study. *J Learn Disabil* 1990; 23: 589–596.
29. MacLachlan A, Yale S & Wilkins A. Open trial of subjective precision tinting: a follow-up of 55 patients. *Ophthalmic Physiol Opt* 1993; 13: 175–178.
30. Anderson LL. *Effects of Irlen Syndrome and visual dysfunctions on reading comprehension*. [Internet]. [US]: ProQuest Information & Learning; 1997. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1997-95001-146&site=ehost-live.
31. Wilkins A & Neary C. Some visual, optometric and perceptual effects of coloured glasses. *Ophthalmic Physiol Opt* 1991; 11: 163–171.
32. Jeanes R, Busby A, Martin J *et al.* Prolonged use of coloured overlays for classroom reading. *Br J Psychol* 1997; 88: 541–548.
33. Lightstone A, Lightstone T & Wilkins A. Both coloured overlays and coloured lenses can improve reading fluency, but their optimal chromaticities differ. *Ophthalmic Physiol Opt* 1999; 19: 279–285.
34. Wilkins A, Lewis E, Smith F, Rowland E & Tweedie W. Coloured overlays and their benefit for reading. *J Res Read* 2001; 24: 41–64.
35. Robinson GL & Foreman PJ. 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. *Percept Mot Skills* 1999; 89: 83–113.
36. Robinson GL & Foreman PJ. Scotopic sensitivity/Irlen syndrome and the use of coloured filters: a long-term placebo-controlled study of reading strategies using analysis of miscue. *Percept Mot Skills* 1999; 88: 35–52.
37. Allen P, Evans BJW & Wilkins AJ. Vision and reading difficulties part 5: Clinical protocol and the role of the eye care practitioner. *Optom Today* 2009; 30–36.
38. Kriss I & Evans BJW. The relationship between dyslexia and Meares-Irlen Syndrome. *J Res Read* 2005; 28: 350–364.
39. Simmers AJ, Bex PJ, Smith FK & Wilkins AJ. Spatiotemporal visual function in tinted lens wearers. *Invest Ophthalmol Vis Sci* 2001; 42: 879–884.
40. Evans BJW & Joseph F. The effect of coloured filters on the rate of reading in an adult student population. *Ophthalmic Physiol Opt* 2002; 22: 535–545.
41. Allen PM, Hussain A, Usherwood C & Wilkins AJ. Pattern-related visual stress, chromaticity, and accommodation. *Invest Ophthalmol Vis Sci* 2010; 51: 6843–6849.
42. Hollis J & Allen PM. Screening for Meares-Irlen sensitivity in adults: can assessment methods predict changes in reading speed? *Ophthalmic Physiol Opt* 2006; 26: 566–571.
43. Singleton C & Henderson L-M. Computerized screening for visual stress in children with dyslexia. *Dyslexia Chichester Engl* 2007; 13: 130–151.
44. Henderson LM, Tsogka N & Snowling MJ. Questioning the benefits that coloured overlays can have for reading in students with and without dyslexia. *J Res Spec Educ Needs* 2013; 13: 57–65.
45. Mitchell C, Mansfield D & Rautenbach S. Coloured filters and reading accuracy, comprehension and rate: a placebo-controlled study. *Percept Mot Skills* 2008; 106: 517–532.
46. Bouldoukian J, Wilkins AJ & Evans BJW. Randomised controlled trial of the effect of coloured overlays on the rate of reading of people with specific learning difficulties. *Ophthalmic Physiol Opt* 2002; 22: 55–60.
47. Singleton C & Trotter S. Visual stress in adults with and without dyslexia. *J Res Read* 2005; 28: 365–378.
48. Singleton C & Henderson L-M. Computerised screening for visual stress in reading. *J Res Read* 2007; 30: 316–331.
49. Wilkins AJ, Evans BJ, Brown JA *et al.* Double-masked placebo-controlled trial of precision spectral filters in children who use coloured overlays. *Ophthalmic Physiol Opt* 1994; 14: 365–370.
50. Tyrrell R, Holland K, Dennis D & Wilkins A. Coloured overlays, visual discomfort, visual search and classroom reading. *J Res Read* 1995; 18: 10–23.
51. Fergusson D, Aaron SD, Guyatt G & Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002; 325: 652–654.
52. Sedgwick P. Analysis by intention to treat. *BMJ* 2011; 342: d2212.
53. Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005; 2: e124.
54. Simmons JP, Nelson LD & Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 2011; 22: 1359–1366.
55. Irlen H. *The Irlen Revolution: A Guide to Changing Your Perception and Your Life*. Square One Publishers: Garden City Park, NY, 2010; p. 212.
56. Blaskey P, Scheiman M, Parisi M, Ciner EB, Gallaway M & Selznick R. The effectiveness of Irlen filters for improving reading performance: a pilot study. *J Learn Disabil* 1990 Dec; 23: 604–612.
57. Cotton MM & Evans KM. An evaluation of the Irlen lenses as a treatment for specific reading disorders. *Aust J Psychol* 1990; 42: 1–12.
58. O'Connor PD, Sofo F, Kendall L & Olsen G. Reading disabilities and the effects of colored filters. *J Learn Disabil* 1990; 23: 597–603.
59. Martin F, Mackenzie B, Lovegrove W & McNicol D. Irlen lenses in the treatment of specific reading disability: an

evaluation of outcomes and processes. *Aust J Psychol* 1993; 45: 141–150.
60. Fletcher J & Martinez G. An eye-movement analysis of the effects of Scotopic Sensitivity correction on parsing and comprehension. *J Learn Disabil* 1994; 27: 67–70.
61. Ritchie SJ, Della Sala S & McIntosh RD. Irlen colored overlays do not alleviate reading difficulties. *Pediatrics* 2011; 128: e932–e938.
62. Ritchie SJ, Della Sala S & McIntosh RD. Irlen colored filters in the classroom: a 1-year follow-up. *Mind Brain Educ* 2012; 6: 74–80.
63. Research Papers | Coloured Lenses and Visual Stress [Internet]. [cited 2016 Jan 20]. Available from: http://colouredlensesandvisualstress.com/research-papers/
64. Faraci ME. *Does the Irlenrtm method bring about an increase in reading scores on a specific test of reading for students found to have Scotopic Sensitivity Syndrome?* [Internet]. [US]: ProQuest Information & Learning; 2010. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2010-99130-231&site=ehost-live.
65. Noble J, Orton M, Irlen S & Robinson G. A controlled field study of the use of coloured overlays on reading achievement. *Aust J Learn Disabil* 2004 Jun; 9: 14–22.
66. Donovan MA. *Reading effectiveness of Irlen Overlays*. [Internet]. [US]: ProQuest Information & Learning; 1995. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1995-95024-094&site=ehost-live.
67. Mason CA. *Scotopic Sensitivity/Irlen Syndrome and reading in college-level students*. [Internet]. [US]: ProQuest Information & Learning; 2000. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2000-95003-188&site=ehost-live.
68. Morrison RF. *The effect of color overlays on reading efficiency*. [Internet]. [US]: ProQuest Information & Learning; 2012. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2012-99170-193&site=ehost-live.
69. Adams TA. *Improving scores on computerized reading assessments: The effects of colored overlay use*. [Internet]. [US]: ProQuest Information & Learning; 2013. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2013-99090-005&site=ehost-live.
70. ChromaGen | Cantor+Nissel [Internet]. [cited 2015 Nov 14]. Available from: http://dyslexia-help.co.uk/.
71. Cardona G, Borràs R, Peris E & Castañé M. A placebo-controlled trial of tinted lenses in adolescents with good and poor academic performance: reading accuracy and speed. *J Optom* 2010; 3: 94–101.
72. Hodd, Nigel Burnett. *The Cromagen method for colour vision deficiency and special learning difficulties. Optom Today*. 2000;November 17. http://burnetthoddoptometry.com/burnetthoddoptometry/wp-content/uploads/2014/04/The-Chromagen-method-for-colour-deficiency-and-specific-learning-difficulties.pdf (accessed 19/7/2016).
73. Morgan MJ & Thompson P. Apparent motion and the Pulfrich effect. *Perception* 1975; 4: 3–18.
74. Hall R, Ray N, Harries P & Stein J. A comparison of two-coloured filter systems for treating visual reading difficulties. *Disabil Rehabil* 2013; 35: 2221–2226.
75. Harries P, Hall R, Ray N & Stein J. Using coloured filters to reduce the symptoms of visual stress in children with reading delay. *Scand J Occup Ther* 2015; 22: 153–160.
76. Harris D & MacRow-Hill SJ. Application of ChromaGen haploscopic lenses to patients with dyslexia: a double-masked, placebo-controlled trial. *J Am Optom Assoc* 1999; 70: 629–640.
77. Christenson GN, Griffin JR & Taylor M. Failure of blue-tinted lenses to change reading scores of dyslexic individuals. *Optom J Am Optom Assoc* 2001; 72: 627–633.
78. Palomo-Álvarez C & Puell MC. Effects of wearing yellow spectacles on visual skills, reading speed, and visual symptoms in children with reading difficulties. *Graefes Arch Clin Exp Ophthalmol* 2013; 251: 945–951.
79. Iovino I, Fletcher JM, Breitmeyer BG & Foorman BR. Colored overlays for visual perceptual deficits in children with reading disability and attention deficit/hyperactivity disorder: are they differentially effective? *J Clin Exp Neuropsychol* 1998; 20: 791–806.
80. Vidal-López J. The role of attributional bias and visual stress on the improvement of reading speed using colored filters. *Percept Mot Skills* 2011; 112: 770–782.
81. Dyslexia Research Trust – Vision & Coloured Filters [Internet]. [cited 2015 Dec 27]. Available from: http://www.dyslexic.org.uk/research/vision-coloured-filters.
82. Sawyer C, Taylor S & Willcocks S. Transparent coloured overlays and specific learning difficulties. *Educ Psychol Pract* 1994; 9: 217–220.
83. Gole GA, Dibden SN, Pearson CC *et al.* Tinted lenses and dyslexics–a controlled study. SPELD (S.A.) Tinted Lenses Study Group. *Aust N Z J Ophthalmol* 1989; 17: 137–141.
84. Menacker SJ, Breton ME, Breton ML, Radcliffe J & Gole GA. Do tinted lenses improve the reading performance of dyslexic children? A cohort study *Arch Ophthalmol* 1993; 111: 213–218.
85. Saint-John LM & White MA. The effect of coloured transparencies on the reading performance of reading-disabled children. *Aust J Psychol* 1988; 40: 403–411.
86. Evans BJ, Cook A, Richards IL & Drasdo N. Effect of pattern glare and colored overlays on a stimulated-reading task in dyslexics and normal readers. *Optom Vis Sci* 1994; 71: 619–628.
87. Francis M, Taylor S & Sawyer C. Coloured lenses and the Dex frame: New issues. *Support Learn* 1992; 7: 25–27.
88. Rudnicka AR & Owen CG. An introduction to systematic reviews and meta-analyses in health care: systematic reviews and meta-analyses. *Ophthalmic Physiol Opt* 2012; 32: 174–183.
89. Leek JT & Peng RD. Statistics: P values are just the tip of the iceberg. *Nature* 2015; 520: 612.

90. Nuzzo R. Scientific method: Statistical errors. *Nature* 2014; 506: 150–152.
91. Malins C, Mind Matters Psychology, New Zealand, Ministry of Health, Irlen Screening Research Project. *The use of coloured filters and lenses in the management of children with reading difficulties a literature review commissioned for the Irlen Screening Research Project, Ministry of Health* [Internet]. [Wellington, NZ]: Mind Matters Psychology; 2009 [cited 2016 Feb 15]. Available from: http://www.moh.govt.nz/moh.nsf/pagesmh/10456/$File/use-of-coloured-filters-and-lenses-children-learning-difficulties-nov2010.pdf.
92. Galuschka K & Schulte-Körne G. The diagnosis and treatment of reading and/or spelling disorders in children and adolescents. *Dtsch Ärztebl Int* 2016; 113: 279–286.
93. McCambridge J, Kypri K & Elbourne D. Research participation effects: a skeleton in the methodological cupboard. *J Clin Epidemiol* 2014; 67: 845–849.
94. Love A, Kavazis AN, Morse A & Mayer KC Jr. Soccer-specific stadiums and attendance in major league soccer: investigating the novelty effect. *J Appl Sport Manag Res Matters* Vol 5 No 2 2013 [Internet]. 2013; Available from: http://js.sagamorepub.com/jasm/article/view/3294.
95. Sedgwick P. The Hawthorne effect. *BMJ* 2011; 344: d8262.
96. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M & Fisher P. The Hawthorne effect: a randomised, controlled trial. *BMC Med Res Methodol* 2007; 7: 30.
97. Kaptchuk TJ, Kelley JM, Conboy LA *et al.* Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ* 2008; 336: 999–1003.
98. Meissner K, Fässler M, Rücker G *et al.* Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. *JAMA Intern Med* 2013; 173: 1941–1951.
99. Benedetti F & Dogue S. Different placebos, different mechanisms, different outcomes: lessons for clinical trials. *PLoS ONE* 2015; 10: e0140967.
100. Sedgwick P. Bias in randomised controlled trials: comparison of crossover group and parallel group designs. *BMJ* 2015; 351: h4283.
101. Ding H, Hu GL, Zheng XY, Chen Q, Threapleton DE & Zhou ZH. The method quality of cross-over studies involved in cochrane systematic reviews. *PLoS ONE* 2015; 10: e0120519.
102. Principles of examining patients with specific learning difficulties – The College of Optometrists [Internet]. [cited 2016 Feb 11]. Available from: http://guidance.college-optometrists.org/guidance-contents/knowledge-skills-and-performance-domain/examining-patients-with-specific-learning-difficulties/principles-of-examining-patients-with-specific-learning/?searchtoken=A85.
103. Creavin AL, Lingam R, Steer C & Williams C. Ophthalmic abnormalities and reading impairment. *Pediatrics* 2015; 135: 1057–1065.
104. Motsch S & Mühlendyck H. Frequency of reading disability caused by ocular problems in 9- and 10-year-old children in a small town. *Strabismus* 2000; 8: 283–285.
105. Lambert SR. Should glasses be prescribed for all children with moderate hyperopia? *Ophthalmology* 2016; 123: 676–678.
106. Duff FJ & Clarke PJ. Practitioner Review: Reading disorders: what are the effective interventions and how should they be implemented and evaluated?: Effective reading interventions. *J Child Psychol Psychiatry* 2011; 52: 3–12.
107. Kaplan RM & Irvin VL. Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS ONE* 2015; 10: e0132382.
108. Society for Coloured Lens Prescribers [Internet]. [cited 2015 Nov 22]. Available from: http://www.ioo.org.uk/s4-clp/code-of-conduct.htm.
109. David SP, Ware JJ, Chu IM *et al.* Potential reporting bias in fMRI studies of the brain. *PLoS ONE* 2013; 8: e70104.
110. Vul E, Harris C, Winkielman P & Pashler H. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci J Assoc Psychol Sci* 2009; 4: 274–290.
111. Bishop DVM. Research Review: Emanuel Miller Memorial Lecture 2012 – Neuroscientific studies of intervention for language impairment in children: interpretive and methodological problems. *J Child Psychol Psychiatry* 2013; 54: 247–259.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Study characteristics. For exploratory studies containing multiple experiments only those with reading as a dependent variable are tabulated.

Philip G. Griffiths was formerly a Consultant Ophthalmologist and Clinical Senior Lecturer in Ophthalmology in Newcastle upon Tyne with a special interest in neuro-ophthalmology. He was a member of the mitochondrial diseases research group and was involved with phenotype genotype correlations of these disorders which frequently affect the eye. Philip has been a co-author of two Cochrane Collaboration systematic reviews on the management of traumatic optic neuropathies and with his colleagues designed the first treatment trial of Leber's hereditary optic neuropathy. More recently Philip has been working as Consultant Ophthalmologist in Gibraltar where he has developed an interest in visual stress and its possible relationship with reading difficulties.

Robert H. Taylor qualified in 1985 from Guy's hospital, London University and worked in a variety of locations including Birmingham, Sheffield and Leeds before taking up a consultant post in York where he has worked for 20 years, specialising in general, paediatric ophthalmology and adult motility. Robert has published 53 peer review papers, made 65 oral/poster presentations at scientific meetings, and given 84 lectures, in York, nationally and internationally. In addition, he has published two books (*Key Topics in Ophthalmology*) plus 1 chapter in a Taylor/Hoyt's *Paediatric Ophthalmology* text book (on vision and reading). Robert has been head examiner for FRCOphth Part 2 (exit exam) 2011-2016 and head examiner for Refraction certificate 2008 to 2011 and is currently the Chair of the examinations committee of the Royal College of Ophthalmology. Research interests include refraction accuracy, reading and maximising visual potential in cataract patients. Robert is also interested in various models of healthcare delivery.

Lisa M. Henderson is currently a Lecturer in Psychology at the University of York. Her main research interests are in the development and disorders of reading and language. She is particularly interested in vocabulary learning and the role that sleep plays in consolidating newly learned words across typical and atypical development. Prior to completing a PhD with Professor Maggie Snowling and Dr Paula Clarke on vocabulary impairments in children with and without autism and language difficulties, Lisa worked as a research assistant for Dr Chris Singleton at the University of Hull, which initiated her interest in visual stress and it's connection to dyslexia. They carried out research to develop and validate an objective screener for symptoms of visual stress in dyslexic and non-dyslexic populations.

Brendan T. Barrett is an Optometrist who trained at the Dublin Institute of Technology. Having completed his PhD, he became a lecturer in the Optometry school at Glasgow Caledonian University before moving to the University of Bradford, where he has been a member of academic staff ever since. In 2008, he completed a BSc in Psychology with the Open University. His current title is 'Professor of Visual Development'. He is an active researcher in the binocular vision area (in particular, in amblyopia) and in the area of vision and sport. In vision and sport, his current research is aimed at examining if elite-level athletes have better vision than less able sportspeople, and if so, which tests reveal the differences. In amblyopia, he is interested in the natural history of how eyes become amblyopic and in the everyday impact of living with amblyopia. He is also interested in examining the evidence base for a variety of aspects of current optometric practice.