Investigating the outcomes of adult strabismus surgery undertaken for psychosocial reasons

Gemma Elizabeth Arblaster

Division of Ophthalmology and Orthoptics, Health Sciences School

May 2021

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy

Supervised by:

Dr David Buckley

Dr Sarah Barnes

Professor Helen Davis

Abstract

Introduction

Strabismus surgery undertaken for psychosocial reasons in adults aims to improve eye alignment and health related quality of life (HRQoL). Additionally, patients can gain a larger visual field, unexpected binocular single vision (BSV), and improved binocular summation, eye movements and task performance from surgery. Despite these improvements, NHS funding for strabismus surgery, without expected visual benefit, has been withdrawn in some areas of England due to concern that not enough patient benefit from surgery is proven.

Methods

A mixed methods feasibility study was undertaken to investigate the outcomes of adult strabismus surgery undertaken specifically for psychosocial reasons. In the qualitative phase semi-structured interviews were conducted postoperatively and the findings informed the quantitative phase design. The quantitative phase prospectively recruited surgery and control group participants to undergo standard clinical measurements and additional study measurements.

Results

In the qualitative interviews participants (n=13) reported a range of improvements in their vision, task performance, physical symptoms and confidence and emotions postoperatively. Compared to the control group (n=15), the surgery group (n=12) had postoperative quantitative improvements in binocular summation at 100% contrast, coarse stereotest (CST) performance, the time to perform a touchscreen spatial localisation (TSL) task and the time to perform the clinical kinematic assessment tool (CKAT) aiming task. Improvements were also reported in vision, task performance, physical symptoms, confidence and emotions, and health related quality of life (HRQoL) using patient reported outcome measures (PROMs). Most measures were unchanged and some worsening of task performance (bead threading and grooved pegboard) was measured postoperatively.

Conclusion

Strabismus surgery undertaken for psychosocial reasons can lead to objective improvements in vision and task performance and subjective improvements in vision, task performance, physical symptoms and confidence and emotions. These improvements were in addition to the typically expected outcomes of improved eye alignment and improved HRQoL.

Acknowledgements

I would like to thank my supervisors Dr David Buckley, Dr Sarah Barnes and Professor Helen Davis. I am very grateful for their continued support, encouragement and expertise throughout my PhD study.

Thank you to the NIHR for funding this PhD and my Clinical Doctoral Research Fellowship. Throughout my fellowship I have been fortunate to have the additional support of clinical supervisors Professor Anna Horwood and Dr Helen Griffiths, and an academic advisory group including Dr Jill Carlton, Mr John Burke, Dr Simon Goodwill and Dr Charlotte Codina. I thank them all for their guidance and expert advice during my PhD.

I would like to thank the Patient and Public Involvement (PPI) group for their input into the development of this study, particularly Charlotte Watson who continued as a patient representative throughout the study. I am very grateful for her calm and reasoned perspective throughout the whole PhD.

Thank you to all the participants who took part in the study. Quite simply, without their time, enthusiasm and dedication, this study would not have been possible.

The support of the clinical and administrative staff in the Ophthalmology Department, Sheffield Teaching Hospitals NHS Foundation Trust has been invaluable. I would particularly like to thank all the Orthoptists and strabismus surgeons (Mr John Burke, Miss Jessy Choi and Miss Anamika Tandon) for their support. Thank you to the administrative staff for all their help with clinical notes and appointments, particularly Adam Gibson; and all the staff at the Clinical Research and Innovation Office (CRIO) who helped behind the scenes.

Lastly, I would like to thank my family and friends for their unwavering support throughout my PhD study. In particular, thank you to Alex, William and Daisy.

Contents

Statement of work undertaken

I successfully applied for a personal NIHR Clinical Doctoral Research Fellowship grant to support this work. During the fellowship my academic supervisors were Dr David Buckley, Dr Sarah Barnes and Prof Helen Davis. My clinical supervisors were Prof Anna Horwood and Dr Helen Griffiths. I additionally had the support of an advisory group of experts to support me during the fellowship. These included Charlotte Watson (patient representative), Dr Jill Carlton (quality of life and patient reported outcomes), Dr Simon Goodwill (sports engineer – measuring performance), Dr Charlotte Codina (measuring visual function) and Mr John Burke (Consultant Ophthalmologist and Strabismologist). All supervisors and members of the advisory group were available during the fellowship for discussion and to give advice relating their area of expertise.

The design of this research was a result of collaboration between my clinical and academic supervisors, Dr Helen Griffiths, Dr David Buckley, Prof Helen Davis and Dr Sarah Barnes and myself. It was also informed by feedback from the reviewers of a previous NIHR grant application (Doctoral Research Fellowship, 2013) and feedback from the Patient and Public Involvement (PPI) group of adults with strabismus. The PPI group was created by me for adults with strabismus to input their views into the research grant applications and the study design. All of the PPI group had strabismus, some had undergone previous surgery and some had not. The development of the PPI group was supported by a grant awarded to me by the Yorkshire and Humber Research Design Service. One member of the original PPI group (Charlotte Watson) continued as a patient representative on the advisory group throughout the study.

During the study I applied for the required ethical and other approvals. I identified all potential participants in the Ophthalmology clinic for recruitment, with the support of the Orthoptists and Ophthalmologists in the strabismus clinic and the support of a number of Ophthalmology administrative staff, particularly Adam Gibson. I recruited all the study participants and assessed all the participants during their study visits. All the qualitative interviews were conducted and transcribed by me. At the quantitative visits I assessed all of the participants during their planned Orthoptic appointments, except a small number of patients who were seen by an Orthoptist following standard clinical procedures when I was unavailable. I conducted all of the quantitative study visits. The eye movement recording programmes were written by myself and Dr Helen Griffiths, with help from Dr Sam Hutton and the support team at SR Research. The R programme to analyse the smooth pursuit was developed by PhD student Noor Shafee and Dr Helen Griffiths. The CKAT and BIGKAT devices were loaned to me for the quantitative phase of the study by Professor Mark Mon-Williams and Dr Rachel Coates, Psychology Department, University of Leeds. The TSL task programme was written by Dr Simon Goodwill and the equipment to validate the screen used during the TSL task was borrowed from Sheffield Hallam University, under the

guidance of Dr Simon Goodwill. The balance software and hardware were selected following advice from Dr Simon Goodwill and Dr Ben Heller at Sheffield Hallam University. Dr Jill Carlton discussed the QoL measures and PROMs and advised me on the development of the additional study questions. I carried out all of the statistical analysis with supervision from my supervisor Dr David Buckley and advice from Kathleen Baster and Dr Lucy Gelder (Statistical Services Unit, University of Sheffield). I have written this thesis with advice from my supervisors and it represents my own work.

Publications

Publications arising from this work

(ARVO conference poster, abstract)

Arblaster G, Davis H, Buckley D, Barnes S. (2019) What changes do patients report after strabismus surgery for planned psychosocial benefit? Invest Ophthalmol Vis Sci. 60(9):238

Other publications during the course of this fellowship

Bjerre A, Arblaster G, Nye A, Griffiths HJ. (2018) The provision of patient information about nystagmus. Brit Ir Orthopt J 14(1): 25-29

Coughlan A, Arblaster G, Burke JP. (2018) A case report of progressive Brown syndrome? Brit Ir Orthopt J. 14(1): 30-34.

Randall D, Griffiths H, Arblaster G, Bjerre A, Fenner J. (2018) Simulation of Oscillopsia in Virtual Reality. Brit Ir Orthopt J. 14(1): 45-49.

Steel DA, Codina CJ, Arblaster GE. (2019) Amblyopia treatment and quality of life: the child's perspective on atropine versus patching. Strabismus. 27(3): 156-164.

Lijka B, Toor S, Arblaster G. (2019) The Impact of Diplopia on Reading. Brit Ir Orthopt J. 15(1): 8-14.

Porter K, Arblaster G. (2020) How does vertical reading affect reading speed? Brit Ir Orthopt J. 16(1): 38-43.

Randall D, Fox SL, Fenner JW, Arblaster GE, Bjerre A, Griffiths HJ. (2020) Using VR to Investigate the Relationship between Visual Acuity and Severity of Simulated Oscillopsia. Curr Eye Res. 45(12): 1611-1618.

List of tables

List of figures

List of abbreviations

AC/A ratio	Accommodative convergence / accommodation ratio
AHP	Abnormal head posture
Alt	Alternating
AMD	Age related macular degeneration
ANOVA	Analysis of variance
ANCOVA	Analysis of covariance
APCT	Alternate prism cover test
A&SQ	Amblyopia and Strabismus Questionnaire
AS-20	Adult Strabismus Quality of Life Questionnaire
BIGKAT	Boxed Infrared Gross Kinematic Assessment Tool (optical motion capture system, used to capture the movement of infrared emitting diodes in 3D space)
BSV	Binocular single vision
BT	Botulinum Toxin
CESD-R	Centre for Epidemiologic Studies Depression Scale - Revised
CKAT	Clinical Kinematic Assessment Tool
CL	Contact lens
CST	Coarse stereotest
CT	Cover test
DAS	Derriford Appearance Scale (DAS59 and DAS24 short form versions)
DVD	Dissociated vertical deviation
E	Esophoria
ET	Esotropia
ETs	Esotropes
EMR	Eye movement recording
EOM	Extra ocular muscles
FD2	Frisby Davis distance stereotest
FNS	Frisby Near Stereotest
Gls	Glasses
HADS	Hospital Anxiety and Depression Scale
HRA	Health Research Authority
HRQoL	Health Related Quality of Life
HT	Hypertropia
HoT	Hypotropia
IO	Inferior Oblique
IPD	Inter pupillary distance

IQR	Interquartile Range
IR	Inferior Rectus
IREDs	Infrared emitting diodes
L	Left
LCD	Liquid crystal display
LoA	Limits of agreement
LR	Lateral Rectus
L-R app	Lee-Ryan eye hand coordination application
MR	Medial Rectus
ms	Milliseconds
MSK	Musculoskeletal
NHS	National Health Service
OKN	Optokinetic nystagmus
OM	Ocular movements
OSOP	'One sheet of paper' technique
PCT	Prism Cover Test
PD	Prism Dioptres
PFR	Prism Fusion Range
PPI group	Patient and Public Involvement group
PROM	Patient Reported Outcome Measure
R	Right
RCS	Rate Correct Score (for the coarse stereotest results, RCS = number of correct responses / second)
REC	Research Ethics Committee
RMSE	Root Mean Square Error
pPA	Penalised Path Accuracy
SD	Standard deviation
SE	Standard error
SO	Superior Oblique
SPCT	Simultaneous prism cover test
SR	Superior Rectus
STAI-S	State-Trait Anxiety Inventory – state anxiety subscale
STAI-T	State-Trait Anxiety Inventory – trait anxiety subscale
STH NHS FT	Sheffield Teaching Hospitals NHS Foundation Trust
QALYs	Quality Adjusted Life Years
QoL	Quality of Life
UK	United Kingdom
USA	United States of America

VA	Visual acuity
VF-14	Visual Function Questionnaire - 14
VFQ-25	National Eye Institute Visual Functioning Questionnaire – 25
WHO	World Health Organisation
WHOQoLBref	World Health Organisation Quality of Life assessment – abbreviated version (abbreviated version of the WHOQoL-100)
X	Exophoria
XT	Exotropia
XTs	Exotropes
2D	Two dimensional
3D	Three dimensional

Chapter 1. Introduction to the thesis

This mixed methods thesis aimed to investigate the outcomes from strabismus surgery undertaken for psychosocial reasons in adults and measure whether benefit could be gained from surgery, in addition to the known psychosocial benefits.

The background to this study was the growing need for evidence of the outcomes from strabismus surgery, particularly in adults experiencing psychosocial problems as a result of strabismus. Some areas in England were no longer funding surgery for adults with strabismus, unless they experienced diplopia symptoms or had evidence of potential binocular single vision (BSV) preoperatively (Bristol North Somerset and South Gloucestershire Clinical Commissioning Group, 2019). Whilst the evidence describing the 'psychosocial benefits' of strabismus surgery in this subgroup of adults was considered strong by clinicians working in the field of strabismus, commissioners of National Health Service (NHS) services in some areas of England were reporting concern that not enough patient benefit was demonstrated to warrant continuing to offer strabismus surgery for psychosocial reasons. Funding for adult strabismus surgery for psychosocial reasons was considered 'at risk' by the NHS ophthalmic community. There was a clear need for more robust evidence of the outcomes of strabismus surgery for patients with psychosocial symptoms. Clinical decision making required high quality evidence and before a larger trial could be planned, the outcome measures to be included in a trial needed to be established.

The study was a mixed methods design, combining qualitative and quantitative research methods. The aim was to firstly undertake the qualitative part of the study (phase one) to interview adult patients who had already undergone strabismus surgery for psychosocial reasons. The interviews aimed to find out what the patients felt had changed, or not changed, for them postoperatively. During the semi-structured interviews participants were invited to talk about their experiences and observations following strabismus surgery, including both positive and negative changes, as well as observations of no change. The information gained from the qualitative findings, in addition to evidence from the literature, was used to inform and improve the design of the quantitative part of study (phase two). Phase two involved recruitment of adult patients schedules to have strabismus surgery for psychosocial reasons. They were invited to undergo a range of measures of their vision and task performance, and complete patient reported outcome measures (PROMs), before and after strabismus surgery. Additionally, a control group of adults with strabismus, but not having strabismus surgery, was recruited to undergo the same measures as the surgery group, at two separate visits. The control group was used to establish whether improvements could be expected from performing the measures at a second visit.

This feasibility study explored firstly whether it was feasible, and possible, to measure and quantify changes in visual function and task performance following strabismus surgery. Secondly, the study aimed to determine how feasible it was to recruit and retain adult patients to a study that required these measurements before and after surgery. Thirdly, the aim was to determine if the tests of visual function and task performance were acceptable to patients and whether the acceptable tests could be refined for a future study.

1.1 Structure of the thesis

Chapter two describes the background literature around adult strabismus, including the consequences and psychosocial impact of strabismus. Chapter three reviews the literature reporting different outcomes of adult strabismus surgery undertaken for psychosocial reasons. Chapter four describes the rationale for the overall study and summarises the study aims and objectives.

Chapter five describes the mixed methods methodology of the overall study and introduces the qualitative phase one, including the setting and methodology. Chapter six describes the qualitative findings and the themes emerging from phase one. Chapter seven describes how the qualitative findings and the literature evidence were used to select and refine the measures for the quantitative part of the study (phase two). Chapter eight describes the quantitative phase two methods and chapter nine reports the phase two quantitative results and statistical analysis. Chapter ten discusses the results of the overall study, combining both the qualitative phase one findings and the quantitative phase two results, relating the discussion to the published literature. The limitations of the study and areas for further research are discussed. The conclusions drawn and the contribution this study makes to our wider understanding of the outcomes of strabismus surgery undertaken for psychosocial reasons are also presented.

1.2 Patient and Public Involvement (PPI) group

A PPI group of adults with strabismus (n=6) was created to enable adults with strabismus to input their views into the study. All of the PPI group had strabismus and some had undergone previous surgery. One member of the original PPI group (CW) continued as a patient representative on the advisory group throughout the study. During the study design stage, the PPI group recommended that wider dissemination of the study findings to professional groups, such as GP's and Optometrists, should be considered.

This thesis also meets some of the research priorities identified by the Sight Loss and Vision Priority Setting Partnership (2013). Research priority 10: refractive error and ocular motility, rank 8

‘how can the functional effects of surgical treatment for squint best be assessed? This priority setting process was undertaken by the James Lind Alliance with high PPI input into the process.

Chapter 2. Strabismus

This chapter discusses the literature evidence around adult strabismus and the consequences of having strabismus. Treatment for adult strabismus is discussed, with particular emphasis on strabismus with psychosocial symptoms.

Strabismus is an abnormal alignment of one eye, which can have onset at any time of life. If strabismus remains following correction of any refractive error and is causing visual symptoms (for example diplopia and/or confusion) (Fawcett et al., 2004) or psychosocial symptoms for the patient (Adams et al., 2016), management options may be considered. Examples of 'psychosocial symptoms' include being more likely to suffer with anxiety and depression (McBain et al., 2014b), having low self-esteem and self-confidence (Nelson et al., 2008; Xu et al., 2012) being less likely to gain employment (Coats et al., 2000), being less likely to be promoted at work (Goff et al., 2006), being perceived negatively by others (Olitsky et al., 1999) and having problems with interpersonal relationships (Burke et al., 1997).

2.1 Prevalence of strabismus

The incidence of adult strabismus (constant and intermittent tropias, excluding phorias) is widely reported to be 4-5% (Beauchamp et al., 2003; Goseki & Ishikawa, 2017; Hashemi et al., 2017). However, paediatric evidence of strabismus prevalence is often cited (Bruce & Santorelli, 2016; Drover et al., 2008; Graham, 1974; Stayte et al., 1993). The incidence in adults has been reported to be as low as 1% (Hashemi et al., 2019). Whilst many paediatric studies have the advantage of large cohorts, for example vision screening populations (Bruce & Santorelli, 2016), there is little evidence of how paediatric strabismus prevalence compares in adulthood. Increasing prevalence of strabismus with increasing age has been reported (Hashemi et al., 2017). This is likely to be due to the majority of paediatric strabismus remaining into adulthood, with a small proportion resolving. This in combination with other strabismus types acquired later in life has led to the commonly cited prevalence of 4% being considered an underestimate of adult strabismus prevalence (Hertle, 1998).

2.2 Consequences of strabismus

Strabismus can cause a range of symptoms and complaints. Patients with strabismus may report diplopia, confusion, loss of BSV, a change in their field of binocular vision (typically a reduction in esotropia (ET) and an expansion, or panoramic vision, in exotropia (XT)), difficulty performing a specific activity (for example reading), asthenopic symptoms (for example eyestrain), abnormal head posture (AHP), psychosocial problems and the actual misalignment of the eyes (Beauchamp et al., 2003). Childhood onset strabismus is associated with strabismic amblyopia. In addition to

these commonly recognised consequences of strabismus, patients may also report other visual symptoms or display other visual behaviours, such as closing one eye, difficulties in busy environments, subconsciously using information from their strabismic eye and difficulties performing some tasks. These are described in the following sections.

2.2.1 Binocular summation and inhibition

Binocular summation is the improved performance of a visual task binocularly compared to monocularly, using the better eye. Binocular inhibition is the improved performance of a visual task monocularly compared to binocularly. Clinical low contrast visual acuity (VA) charts in dim lighting can be used to measure binocular summation and inhibition (Pineles et al., 2014) as well as computerised presentations of different contrast stimuli (Dorr et al., 2019). Binocular summation and low contrast VA charts may represent a more realistic visual task compared to high contrast VA charts, particularly when trying to measure the everyday visual difficulties patients with strabismus can report (Pineles et al., 2013). Patients closing one eye when performing visual tasks, despite not experiencing diplopia has been suggested to be the result of strabismus, rather than amblyopia (Baker et al., 2007; Chang et al., 2017). Strabismus is thought to cause functional deficits in binocular vision, either by decreasing binocular summation or causing binocular inhibition (Pineles et al., 2013; Tandon et al., 2014). Binocular summation has also been found to reduce with increasing age (Pineles et al., 2014). In adults with strabismus binocular inhibition has been shown to increase as the level of contrast is decreased, with lower VFQ-25 visual function scores and lower AS-20 health related quality of life (HRQoL) scores being associated with binocular inhibition (Tandon et al., 2014). Binocular summation was found to be lower in patients with strabismus without BSV compared to intermittent strabismus with BSV (Dorr et al., 2019).

2.2.2 Complex visual scenes

Patients with strabismus may complain of finding some environments more visually challenging than others, often when the scene is busy, cluttered or complex. In a laboratory setting, significantly less binocular summation and more binocular inhibition has been measured against background noise in patients with strabismus compared to normal controls. Exotropes (XTs) were more affected by background noise, compared to esotropes (ETs) and controls, suggesting the complexity of visual scenes may not affect all types of strabismus uniformly (Pineles et al., 2014).

2.2.3 Contribution of the suppressed eye

In the presence of manifest strabismus, suppression occurs to prevent, or eliminate, symptoms of confusion and diplopia. The patient is typically unable to perceive an image with their deviating eye, despite having both eyes open. Experimental evidence has shown attention plays a role in suppression (Economides et al., 2012) and it is increasingly recognised that the supressed eye

contributes to visual performance, rather than being completely 'switched off' or ignored. Suppressed eyes can subconsciously contribute to the detection of stimuli in some areas of the visual field (Barrett et al., 2013) and to the planning of saccades (Griffiths et al., 2011).

2.2.4 Task performance

As well as strabismus causing a loss of BSV and suppression, difficulties performing tasks such as reading and driving (Beauchamp et al., 2003) have been reported and impaired eye-hand coordination during reaching tasks has been measured (Niechwiej-Szwedo et al., 2017), even though the suppressed eye may contribute to visual performance (Barrett et al., 2013; Economides et al., 2012; Griffiths et al., 2011).

2.3 Psychosocial impact of strabismus

The term 'psychosocial' has been used to describe the impact strabismus has on all aspects of life. Due to an increased need to demonstrate both the effects of strabismus and treatment outcomes, the psychosocial impact of strabismus has received clinical and research attention in recent years. The terms 'psychosocial' and 'psychosocial symptoms' will be used throughout this thesis to broadly encompass the psychological and social symptoms that an individual with strabismus may report affects them and their life. This term is used with the acknowledgement that not all people with strabismus suffer with or report psychosocial symptoms. However, even patients with reduced VA who found it difficult to see their own strabismus have reported a negative effect of strabismus on their life, in particular how others treated them (Dawson et al., 2013).

2.3.1 Strabismus and quality of life

The World Health Organisation (WHO) described quality of life (QoL) as the broad concept of an individual's perception of their position in life, which may be influenced by how and where they live, as well as their own perceptions and expectations of life (Whoqol Group, 1998). HRQoL typically refers to how health affects QoL, although there may be a lack of consistency and overlap in the literature for definitions of HRQoL, QoL and health (Karimi & Brazier, 2016).

Patients with strabismus without diplopia tend to have more psychosocial concerns compared to those with diplopia, yet not all studies differentiate between those with and without diplopia (McBain et al., 2014a). Even fewer studies report whether patients have intermittent strabismus or BSV, making it challenging to draw specific conclusions about those with strabismus and psychosocial symptoms only. Questionnaires asking patients to rate how much strabismus has affected their lives and their QoL have been used. Patients retrospectively report strabismus has affected them by interfering with friendships, causing a cosmetic problem, leading to them feeling different and having poor self-image (Satterfield et al., 1993). Receiving ridicule because of

strabismus was also reported to be a lifelong problem. Using the Hopkins Symptom Checklist to measure psychological symptoms, patients with strabismus fell between 'normal' and patients being treated for anxiety and depression disorders (Satterfield et al., 1993). Xu et al. (2012) reported a group of patients, age 16 years and above, due to undergo strabismus surgery (n=56). The patients had a range of strabismus types, but none had diplopia preoperatively. All patients reported their strabismus caused some extent of psychosocial impact. The majority of patients reported strabismus negatively affected their relationships with others and caused a lack of self-confidence and avoidance of social situations.

2.3.2 Strabismus and health related quality of life

Self-reported visual functioning in adults with strabismus has been measured with the VFQ-25 (Mangione et al., 2001; Mangione et al., 1998), with the results used to infer a measurement of how much strabismus affects QoL. Lower vision related QoL (VFQ-25 score) was reported in the 2.9% of German adults with manifest strabismus (Fieß et al., 2020). VFQ-25 results in adults with strabismus have been compared to other eye disorders. Strabismus patients reported significantly worse, or the same, visual function as patients with ocular diseases like diabetic retinopathy, age related macular degeneration (AMD) and glaucoma. Only low vision was associated with worse visual function than strabismus (Chang et al., 2015). Using the Psychological Impact questionnaire, strabismus has been shown to impact on patients regardless of their size or type of deviation, age, gender, diplopia symptoms or vision in their worse eye (Ritchie et al., 2013).

The Adult Strabismus Quality of Life Questionnaire (AS-20) is a HRQoL measure specifically developed for strabismus, however it was designed for use in all types and aetiologies of strabismus (Hatt et al., 2009b; Hatt et al., 2007). The AS-20 is thought to measure similar psychosocial properties to the Amblyopia and Strabismus Questionnaire (A&SQ) but includes more measures of function problems (van de Graaf et al., 2017). The AS-20 has psychosocial and functional subscales to capture the different ways strabismus can affect HRQoL, with the functional subscale being more sensitive to diplopia (Hatt et al., 2009b). The AS-20 had greater sensitivity to strabismus than the more general VFQ-25 (Hatt et al., 2009a) and is considered repeatable (Leske et al., 2010). Compared to the Derriford Appearance Scale (DAS59), the AS-20 remained specific to strabismus and not affected by other body appearance factors (Durnian et al., 2009).

In a study of young adults with strabismus in India, Sah et al. (2017) found that females and ETs reported worse HRQoL (AS-20) than males and XTs. No difference between the psychosocial and functional AS-20 subscales was measured, demonstrating that strabismus affected psychosocial and functional aspects equally. In a study of adult strabismus in the UK, female gender and having lower socioeconomic status have been associated with significantly worse HRQoL (AS-20).

However, age and type and direction of deviation were not found to be factors affecting HRQoL (Durnian et al., 2010).

The AS-20, VFQ-25 and A&SQ (Chinese versions) were used by Wang et al. (2014) to investigate HRQoL in adults presenting with strabismus, 67% of whom had no diplopia. The A&SQ and AS-20 were found to be comparable, but the AS-20 had better sensitivity. Patients with diplopia reported lower functional AS-20 scores compared to those without. All adults with strabismus combined reported lower AS-20 scores than visually normal adults and adults with other eye diseases. In a later mixed methods study, Wang et al. (2018) used qualitative interviews to explore symptoms and QoL in adults with strabismus (n=30) and a questionnaire which was completed by 437 patients. The interviewees and questionnaire respondents were grouped into diplopia or no diplopia, however it is unclear whether any had BSV or intermittent deviations. Strabismus affected QoL by impacting appearance, daily activities, personal development, social interaction and emotions. The interview findings and questionnaire results showed that patients with diplopia reported more symptoms than those without diplopia. However, those without diplopia did report symptoms, particularly blurred vision, monocular vision, physical discomfort and eye fatigue.

2.3.3 Strabismus and mental health

In a prospective study of patients with strabismus, with and without diplopia, (McBain et al., 2014b) a range of questionnaires were used preoperatively to measure HRQoL, including the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) and the AS-20 (Hatt et al., 2009b). Poorer HRQoL was reported by 68% of patients on the psychosocial subscale and 83% on the functional subscale (AS-20). Patient beliefs and understanding of strabismus, as well as their social support network, predicted their QoL and mood, rather than clinical or demographic features (Adams et al., 2016; McBain et al., 2014b). Patients with strabismus were also ten times more likely to suffer with clinical depression or anxiety, compared to the general population, which was similar to patients with other eye conditions or facial disfigurement, but was higher than patients with lifelong health conditions, such as diabetes (McBain et al., 2014b).

In adults with strabismus, having worse depression (CESD-R depression screening questionnaire) was associated with reduced HRQoL, particularly for social interactions, tasks like reading and general function (AS-20) (Hatt et al., 2014). Worse HRQoL was associated with larger deviations and being younger. However, it was unclear whether depressive symptoms develop because of strabismus or the poorer HRQoL associated with strabismus; or whether depression was related to poorer HRQoL independently of strabismus (Hatt et al., 2014). In China, strabismus has been linked to anxiety, depression and drinking alcohol in children aged 10-17 years (Lin et al., 2014).

In the USA, children (less than 19 years old) with XT, particularly intermittent XT, (but not ET) were three times more likely to develop mental illness by early adulthood (Mohney et al., 2008). Divergence insufficiency has been associated with hospital admission for mental health reasons and being more likely to use prescribed psychiatric medications and have an unspecified anxiety disorder compared to control subjects. Hypertropia (HT) has been associated with generalised anxiety disorders (Hassan et al., 2015), although this retrospective study did not report the presence or absence of visual symptoms, including diplopia, which could have affected health and mental illness.

Patients with strabismus greater than 30 prism dioptres (PD) (age 15-65 years old) reported a higher rate of social phobia (53.1%) compared to age and sex matched eye clinic patients without strabismus (17.4%) when using a range of QoL measures. Patients with strabismus also had significantly worse measures of social fear and avoidance as well as scores showing their strabismus caused them 'disability' in different aspects of their life (Bez et al., 2009). However, selection of larger deviations may have biased the results to show higher levels of social phobia, as larger deviations have been associated with a more negative impact on self-esteem (Nelson et al., 2008).

2.3.4 Eye contact

Patients with strabismus often report difficulties making eye contact and being in situations where communication and interaction with others is required. The AS-20 psychosocial subscale includes 10 questions that specifically ask about interactions with others and the way a person perceives they are treated because of their strabismus (Hatt et al., 2009a). Visual analogue scales have also been used for patients to score how noticeable and severe they perceived their strabismus to be. Patients without diplopia scored their strabismus as more noticeable and more severe than those with diplopia (Jackson et al., 2006). In a study of 15-25 year-olds with strabismus in India, 50% reported problems making eye contact with people and 50% reported hiding their strabismus, for example by using an AHP or wearing dark glasses (Menon et al., 2002). Xu et al. (2012) reported 74% of their cohort had problems making eye contact and avoided public activities. Ghiasi et al. (2013) also reported high rates (greater than 63%) suffered with embarrassment, negative self-esteem, avoiding activities and hiding their strabismus from others. Thirty nine percent reported problems making eye contact.

2.3.5 Simulations of strabismus

Each patient's strabismus in combination with their facial features, personality and behaviours are different, making studies of comparable patients, deviations and matched groups difficult. Strabismus simulations using dolls and photographs have been created for standardised experiments and negative perceptions about strabismus have been shown to begin at a young

age. Dolls with simulated strabismus have shown children under 4.5 years-old were not aware of strabismus, but children older than 5.75 years-old were increasingly negative about strabismus (Paysse et al., 2001). Using photographs of children with 50PD strabismus, ET was perceived to be worse than XT and right strabismus was perceived to be worse than left (Mojon-Azzi et al., 2011).

Photographs of simulated strabismus have also been used to explore adult perceptions of strabismus. Adults with 50PD strabismus were perceived as significantly less intelligent and worse at communication (Olitsky et al., 1999). Females with strabismus were perceived as less able in the workplace compared to orthotropia (Coats et al., 2000; Mojon-Azzi & Mojon, 2009). The presence of ET led to military personnel being rated as significantly less suitable for promotion compared to without strabismus (Goff et al., 2006). ET was perceived more negatively than XT and HT, and females judged the images more negatively than males, regardless of the gender of the individual in the image (Kothari & Joshi, 2014). However, others have found XT to be more negatively perceived than ET (Mojon-Azzi & Mojon, 2009). Strabismus, particularly XT, was perceived to negatively affect employment and dating opportunities, more so than having a very large nose or facial scarring. Only a missing tooth and bad acne were perceived as worse than strabismus (Mojon-Azzi & Mojon, 2009; Mojon-Azzi et al., 2008). Among Optometry students XT was perceived to be more noticeable than ET, however eye contact was perceived to be better in ET compared to XT (Dolven et al., 2011).

2.4 Management options for strabismus

Patients may seek surgery due to the presence of strabismus (the misalignment of their eyes), due to experiencing psychosocial symptoms caused by having strabismus, or both (Beauchamp et al., 2003). The aim of strabismus management is to reduce or eliminate the visual and/or psychosocial symptoms caused by strabismus by realigning the eyes into a straighter position. Management options include strabismus surgery and Botulinum Toxin (BT). Surgery is typically considered a long-term management option. BT wears off and is therefore a short-term management option unless repeated or used to preoperatively investigate diplopia risk. Surgical decision-making depends on individual patient factors and clinical investigations of the strabismus. Postoperative restoration of BSV or improved diplopia (or confusion) symptoms are considered functional aims of surgery that give the patient visual benefit. If no visual symptoms were present and no potential BSV was predicted, surgery may still be considered if the strabismus caused psychosocial symptoms for the patient.

2.4.1 Strabismus surgery

Strabismus surgery involves strengthening or weakening the action of one or more of the extraocular muscles (EOM), to realign the eyes into a straighter position. Procedures are usually performed under general anaesthesia as a day case in the NHS. Recovery from the strabismus surgery is affected by both the anaesthesia and the procedure itself. Conjunctival redness after strabismus surgery in adults is expected to resolve after 10 weeks (Escardó-Paton & Harrad, 2009).

2.4.2 Risks of strabismus surgery

Strabismus surgery is not without risk. During a consultation discussing strabismus management options, surgery risks are explained, as well as the potential benefits, to enable the patient to make an informed decision about undergoing strabismus surgery. Mild complications include dellen, ocular surface problems, suture granuloma, ocular redness, chemosis and haemorrhage (see Table 3 in Bradbury and Taylor (2013)). Severe complications include globe perforation, suspected slipped muscle, infection (House et al., 2019), intraoperative lost or slipped muscles, endophthalmitis, surgical induced necrotising scleritis, retinal detachment, adherence syndrome and anterior segment ischaemia (Bradbury & Taylor, 2013). Muscle slippage is uncommon and can usually be reattached satisfactorily during surgery (Al-Haddad & Abdul Fattah, 2017). Globe perforation is considered rare and serious consequences resulting from perforation are even rarer (Awad et al., 2000). In a UK prospective observational study of strabismus surgery complications occurring over a 2-year period, severe complications were reported in 1 in 400 operations, and 1 in 2,400 (16%) had a poor outcome from that complication (Bradbury & Taylor, 2013). A later 5-year prospective UK study confirmed a similar rate of severe complications of strabismus surgery (1 in 455 cases) (Ritchie & Ali, 2019).

The additional risk of strabismus surgery is that the desired postoperative outcome is not achieved, for example an over or undercorrection. Further treatment or reoperation may be required to achieve the desired result. In a retrospective study reporting reoperation data in the USA, approximately 1 in 15 patients had a strabismus reoperation within a year of surgery and this rate increased with age (Repka et al., 2018).

2.4.3 Numbers undergoing strabismus surgery

Between 2014 and 2020, a mean of 11,776 surgical procedures (range 11,214-11,987) were performed on the EOM each year in the NHS in England. Typically, 44% of these surgical procedures were performed on adults and 56% on children (Health and Social Care Information Centre) (Appendix A). The number of adults undergoing strabismus surgery in the UK has been increasing (Astle et al., 2016; Das et al., 2017). Children and younger adults who initially delayed

strabismus surgery may be choosing to have surgery later in life (Astle et al., 2016). Delaying strabismus surgery until adulthood is also common in the USA. Surgery not previously being offered, surgery offered but declined in the past, previously successful non-surgical management options, previous negative experiences with surgery, and previously receiving information stating surgery was not possible or carried high risk of failure have all been cited as reasons for later presentation (Coats et al., 2005). The number of patients over 65 years old undergoing strabismus surgery in the USA is stable (Repka et al., 2012), but expected to increase over time (Repka, 1997). There are differences in surgery rates in different populations in the USA (Repka et al., 2013) that may reflect varying strabismus prevalence, awareness of surgery, or access to healthcare. In a retrospective study of strabismus surgery rates in the USA, 1 in 20 patients with strabismus was reported to have undergone strabismus surgery over a 4-year period (Repka et al., 2018).

2.4.4 Cost of strabismus surgery

Strabismus surgery costs in England are typically quoted as the NHS tariff for an ocular motility procedure, yet this excludes all pre and postoperative outpatient care. In adults in 2017, the average tariff for an ocular motility procedure was reported to be £1,343 (Das et al., 2017). In 2019-2020 the tariff for an ocular motility procedure ranged from £1,420 (complex) to £1,277 (major, adult) (NHS Improvement, 2019-2020) (Appendix A). Using an approximate value of £1,350 per procedure, the cost to the NHS in England for 4,933 adult procedures in 2019-2020 was £6,659,550.

2.4.5 Utility analysis of strabismus surgery

The effect of strabismus on QoL can be quantified by a utility value, which is calculated using the patient perceived burden of having strabismus (1 representing perfect health, 0 representing death). The difficulty determining a utility value for strabismus is due to its association with amblyopia (van de Graaf et al., 2010). Using a time trade-off method, QoL of life in strabismus has been reported to be 0.93 (IQR 0.83-1.0) (Beauchamp et al., 2005c), 0.85 (Beauchamp et al., 2006) and 0.76 ± 0.31 (Fujiike et al., 2011). Lower utility values (strabismus having a greater effect on QoL, leading to worse QoL) were reported in patients with worse visual symptoms (diplopia and asthenopia) and those seeking surgery. Most patients interviewed were willing to trade part of their life expectancy in exchange for being rid of strabismus, particularly if they were seeking surgery (Beauchamp et al., 2005c).

The cost utility of a medical procedure can be used to evaluate its cost effectiveness. The cost (of strabismus surgery), utility value (of strabismus) and length of time patients have experienced reduced QoL (because of strabismus) in ‘quality-adjusted life years’ (QALYs) are used in the calculation. The cost utility of strabismus surgery therefore reflects the cost to achieve the gain in

utility (postoperative outcome). The cost utility of strabismus surgery, in a range of different types of strabismus, has been reported to be $1,632/QALY (Beauchamp et al., 2006) and $1,303/QALY (Fujiike et al., 2011), with strabismus surgery leading to a mean value gain of 0.99 QALYs (Fujiike et al., 2011). However, these figures are likely to underestimate the cost utility due to the costs of healthcare in the USA and Japan. Despite this likely underestimation of cost-effectiveness, strabismus surgery is considered 'highly cost effective' due to the cost utility being less than $50,000/QALY (Beauchamp et al., 2006; Fujiike et al., 2011). In the UK, interventions costing the NHS up to £20,000 per QALY gained are considered cost effective. Strabismus surgery is therefore described as highly cost effective in the UK (Das et al., 2017).

2.4.6 Funding for strabismus surgery

Financial pressure on the NHS and funding of NHS procedures are emotive issues that spark media interest (BBC News, 2018). NHS funding decisions are made based on both clinical and cost effectiveness. Whilst the NHS has not withdrawn funding for strabismus surgery, some areas of England are no longer funding strabismus surgery, unless the patient has visual symptoms (such as diplopia) or proven visual benefit from treatment (such as regaining BSV). There was concern that not enough patient benefit was proven in those without expected functional visual gains from surgery (Bristol North Somerset and South Gloucestershire Clinical Commissioning Group, 2019). The concern that strabismus surgery for psychosocial symptoms does not provide enough patient benefit was in contrast with the belief of clinicians working in the area of strabismus, who consider strabismus surgery for adults with psychosocial problems to be highly beneficial for patients (Das et al., 2017; Royal College of Ophthalmologists, 2016). Of note in the debate, was that some hospitals have resumed strabismus surgery in England following presentation of additional evidence of patient reported QoL benefits postoperatively (Billington, 2018). This study is focussed on strabismus causing psychosocial symptoms in adults, rather than visual symptoms, due to the need to increase the evidence of treatment outcomes specifically in patients with strabismus and psychosocial symptoms (Durnian et al., 2011).

2.5 Benefits of strabismus surgery

2.5.1 Strabismus surgery for visual benefit

Restoring BSV or eliminating (or improving) diplopia are considered functional aims of strabismus surgery, as they improve vision (Fawcett et al., 2004). Even alignment of childhood onset strabismus in adulthood can improve vision by restoring sensory fusion (Murray et al., 2007), stereopsis (Mets et al., 2004; Morris et al., 1993) and an expanded field of binocular vision (Murray et al., 2007). Achieving BSV postoperatively has been associated with greater long term stability in

the postoperative deviation (Kushner & Morton, 1992) and improved QoL (SF-36 and A&SQ) (Dickmann et al., 2013).

2.5.2 Strabismus surgery for psychosocial benefit

In patients with psychosocial symptoms, strabismus surgery is described as corrective, correcting an abnormality of eye alignment (Olitsky et al., 1999); or reconstructive, aiming to restore a straighter, 'normal' position of the eyes (Marsh, 2015). The term 'cosmetic' strabismus surgery is avoided, due to cosmetic surgery aiming to enhance or beautify (Olitsky et al., 1999). The aim of 'reconstructive' strabismus surgery is therefore to align the eyes in a straighter position and reduce the negative impact of having strabismus.

The psychosocial symptoms that can be caused by strabismus and beliefs about strabismus are increasingly recognised as important factors affecting a person's life (Adams et al., 2016). Strabismus surgery has been reported to significantly improve all aspects of patients' lives (Beauchamp et al., 2005a; Merrill et al., 2010; Xu et al., 2016) both from the patients' and surgeons' perspectives (Beauchamp et al., 2005b). Even patients with very poor VA can gain significant improvements in confidence and less negative interactions and perceptions from other people following treatment to improve eye alignment (Dawson et al., 2013).

Whilst all patients can gain improvements in measures of anxiety, depression, HRQoL, social avoidance, daily functioning and psychological adjustment after strabismus surgery, it is particularly those with psychosocial symptoms that gain most improvement in these areas (Jackson et al., 2006). Improved measures of social phobia, anxiety, depression, QoL and disability measures of how much strabismus affected work, family and social life have all been reported after successful strabismus surgery (Alpak et al., 2014). The definition of 'successful' strabismus surgery and the postoperative improvements in QoL will be discussed in chapter 3. Ozates et al. (2019) compared a range of psychological measures in patients with constant and intermittent XT, pre and postoperatively. Those with constant XT reported significantly worse psychological function than intermittent XT (and BSV) preoperatively. Surgery for constant XT significantly improved all of the psychological measures, yet surgery for intermittent XT improved only some (STAI-S and STAI-T). Postoperatively there was no difference between the groups, showing that surgery for a constant XT can improve psychological measures to the same level as those with BSV who underwent surgery for visual benefit.

2.6 Chapter 2 summary

From the evidence presented in chapter 2, it is evident that having strabismus in adulthood can cause visual and psychosocial symptoms, as well as problems performing everyday tasks.

Psychosocial symptoms were associated with mental health problems, poorer QoL and HRQoL. Strabismus surgery to realign the eyes into a straighter position can lead to visual and psychosocial benefit for the patient, yet some areas in England are no longer funding strabismus surgery in patients with psychosocial symptoms due to concern not enough patient benefit is proven. Chapter 3 will present a systematic search and review of the literature reporting outcomes from strabismus surgery undertaken for psychosocial reasons.

Chapter 3. Outcomes of strabismus surgery undertaken for psychosocial reasons

The preceding chapter introduced the issue that there is concern among some commissioners of NHS services in England, that not enough benefit from surgery is demonstrated in patients with strabismus and psychosocial symptoms. This chapter presents a review of the literature, following a systematic search, reporting the outcomes of strabismus surgery undertaken for psychosocial reasons. The evaluation of surgical 'success' and the timing of postoperative outcome are discussed.

Not all patients with strabismus report symptoms and/or seek surgery (section 2.3). The term 'strabismus surgery for psychosocial reasons' will be used with the acknowledgement that patients may seek surgery due to the presence of strabismus, due to experiencing psychosocial symptoms caused by having strabismus, or both (Beauchamp et al., 2003). For some patients these reasons may be inter-related and difficult to separate.

3.1 Systematic literature search

A systematic search of the literature was undertaken to identify evidence of adult strabismus surgery outcomes in surgery undertaken for psychosocial reasons (search strategy described in Appendix B). As the literature review was not a formal systematic review, a review protocol was not registered with PROSPERO. Of specific interest were the treatment outcomes in strabismus in patients with psychosocial symptoms, but no diplopia and no demonstrable BSV. The search was purposely broadened to include larger strabismus cohorts, where surgery for psychosocial reasons may have been a subgroup. Whilst there is a current need to increase the evidence in adults with strabismus and psychosocial symptoms, this group have less commonly been studied as a separate cohort. One hundred and sixty-three papers were identified in the search and reviewed. Sixty-four papers were included in the literature review reporting the outcomes of strabismus surgery undertaken for psychosocial reasons.

3.2 Strabismus surgery outcomes

3.2.1 Eye alignment

Overwhelmingly the most commonly reported strabismus surgery outcome was the primary position angle of deviation, usually in the distance, measured by the prism cover test (PCT) and reported in PD. Additionally, stating criteria for 'success' based on the strabismus size postoperatively was common. These had the advantage of allowing comparison between the

percentage successfully aligned with surgery, even when different procedures or techniques were compared. Typically, a target angle considered surgical 'success' was stated and a success rate or percentage achieving success postoperatively was reported. A successful angle was often 0-10PD horizontal deviation (Alkharashi & Hunter, 2017), with some specifying 0-5PD (Wang & Nelson, 2011), 0-8PD (Beauchamp et al., 2003), or 0-15PD (Gigante et al., 2018). Vertical angles considered successful were 0-2PD (Beauchamp et al., 2003), 0-4PD (Biglan et al., 1994), 0-5PD, 0-6PD (Alkharashi & Hunter, 2017) although vertical deviations as large as 12PD HT and 20PD HoT have been considered successful (Adams et al., 2016).

Additional factors could be included in the definition of success. For example, a large prospective multicentre study compared outcomes between different centres (specialist or general) and success was graded based on the preoperative surgical aim. Postoperatively success was graded as within 0-5PD (grade 1 success), 6-10PD (grade 2) or greater that 10PD (grade 3) compared to the surgical goal (Lipton & Willshaw, 1995). The original angle of deviation may be included, for example Cifuentes et al. (2018) reported success criteria of residual deviation up to 10PD and consecutive deviation up to 4PD, with no induced lateral incomitance after surgery for large angle horizontal strabismus. A difference in the ET and XT angle, depending on the strabismus type or aim of the procedure may also be specified. For example, in a large retrospective study reporting re-recessions for recurrent ET, success was considered to be 0-10PD residual ET or 0-8PD consecutive XT (Felius et al., 2001). Outcome measures relating to the specific surgical procedure may also be included. For example, the amount of abduction limitation was an outcome of bilateral lateral rectus (LR) recessions for recurrent XT (Elkamshoushy & Langue, 2019) and incidence of consecutive XT and reoperation rate were outcomes in a long term follow up of surgery for childhood onset ET. Postoperative drift (Alkharashi & Hunter, 2017; Eino & Kraft, 1997), whether reoperation was required (Aletaha et al., 2016; Alkharashi & Hunter, 2017) and complications (Faridi et al., 2007) have also been reported as outcome measures, with some including need for reoperation as failure (Dotan et al., 2014).

3.2.2 Diplopia and BSV

Surgical procedures for planned visual benefit typically included the aim of surgery as an outcome, for example the percentage achieving BSV or improvement in BSV postoperatively (Cifuentes et al., 2018). Surgery for strabismus and psychosocial symptoms would not typically include visual symptoms as outcomes, unless postoperative BSV (Ball et al., 1993) or diplopia occurred. Gusek-Schneider and Boss (2010) included diplopia (yes / no), PCT, VA, BSV and patient satisfaction (yes / no) when reporting postoperative outcomes in secondary sensory strabismus (n=26). The challenge of different outcome measures for different patients was recognised in a retrospective study that grouped patients by strabismus onset, before or after visual maturation (n=255) (Hertle, 1998). Success criteria were divided into sensory and motor success. Sensory success included

restoration of BSV or functional field of BSV. Motor success included orthotropia or heterophoria in primary position and at near. In the absence of BSV and diplopia, motor success included alignment, with a less than 12PD horizontal and less than 5PD vertical deviation considered successful.

3.2.3 Defining success

Increasingly a range of factors have been included in a definition of success to reflect the view that eye alignment is not the only important outcome measure. Hatt et al. (2010a) reported success, partial success and failure outcomes, although their cohort included patients both with diplopia and BSV, and without. Success included no diplopia or visual confusion in primary position or when reading, less than 10PD heterotropia in primary position at both near and distance, no prism or occlusion, and no symptoms relating to strabismus or strabismus surgery. Partial success included the same criteria, but with a less than 20PD deviation and mild or intermittent symptoms (relating to the strabismus or surgery). Failure included diplopia or visual confusion in primary position and when reading, 20PD heterotropia or larger, using prism or occlusion, and moderate or severe symptoms (relating to strabismus or surgery). Their criteria were later refined to include success as having no or rare diplopia, partial success as less than 15PD with diplopia sometimes, with and without a prism, and failure as greater than 15PD heterotropia and diplopia often or always at distance or reading (Hatt et al., 2012a, 2016; Liebermann et al., 2013, 2014). In a large prospective cohort study (n=210), patients with all types of strabismus were recruited to a study investigating QoL and mood, before and after strabismus surgery. Deviation size less than 12PD ET, XT or HT and less than 20PD HoT, no or rare diplopia in primary position and reading, and no prism or occlusion needed were used as the criteria determining success, partial success and failure. Success required all three criteria, partial success required one of three criteria and failure required none of the criteria were met (Adams et al., 2016; McBain et al., 2016b).

3.2.4 Patient perception of the postoperative outcome

Success from the patient's perspective may be different to the clinician's perspective. In recognition of this, some studies included objective and subjective outcomes postoperatively (Frangouli & Adams, 2013) or asked patients to report their eye alignment, binocular function and appearance subjectively (happy / unhappy) (Hertle, 1998). In a retrospective study (n=83) 78% underwent surgery for psychosocial reasons (without diplopia) and both objective and subjective success criteria were used to report the outcomes. 83% of all patients had a successful outcome, both objectively (deviation less than 10PD) and a subjectively (very satisfied) (Sandercoe et al., 2014).

Questionnaires

Increasingly, questionnaires for patients to self-report visual function, QoL, HRQoL and PROMs, both generic and those developed specifically for strabismus, have been used pre and postoperatively (Hatt et al., 2016). Using telephone interviews to complete questionnaires postoperatively (n=128) patients reported satisfactory eye position (98%) and improved self-esteem (85%), abilities to meet new people (65%), interpersonal relationships (27%) and abilities to try new activities (16%). Younger patients reported greater improvements postoperatively and a larger preoperative deviation was associated with greater improvements in self-esteem and self-image postoperatively (Nelson et al., 2008). Interviews have been used to complete questionnaires rather than explore patient perceptions of postoperative outcome (Menon et al., 2002; Ribeiro G de et al., 2014). Menon et al. (2002) reported 97.5% of their cohort (n=40) had improved appearance, relationships with others, self-esteem and self-confidence postoperatively. Postoperatively 37.5% changed future plans and 95% reported trying new activities or things that had previously been avoided.

Ghiasi et al. (2013) used a similar questionnaire to Nelson et al. (2008) to prospectively evaluate changes 3-months after strabismus surgery. All aspects of the questionnaire were reported as improved postoperatively. A high percentage of patients reported improved self-esteem (89%), improved relationships (82%), being able to meet new people (79%), and being better at their job or work (76%) postoperatively. A smaller percentage of patients also reported having improved chances of employment (53%) and being able to try new activities (36%) postoperatively. Gender and direction of strabismus did not affect the results.

Burke et al. (1997) asked patients (n=31) seeking surgery for alignment only to complete questions about psychosocial issues, rating themselves on a 5-point-scale preoperatively and 3-months postoperatively. Patients reported significantly improved psychosocial functioning postoperatively. However, they also reported less than ‘ideal world’ results and that others would rate them less highly than they rated themselves postoperatively. Age did not affect the results, but females and ETs reported greater improvements in psychosocial functioning compared to males and XTs. Greater improvements in HRQoL in females postoperatively has also been reported using the AS-20 (Akbari et al., 2015; Alam et al., 2014; Glasman et al., 2013).

Xu et al. (2012) used their own questionnaire to investigate social and psychological effects of strabismus and surgical correction. None of the cohort (n=56) had diplopia preoperatively and 36% had surgery for alignment only (psychosocial reasons). The most common postoperative outcomes (and the percentage of respondents reporting that outcome) were change in appearance (96%), change in self-esteem or self-confidence (96%), change in relationships with friends (91%), trying activities previously avoided (82%) and changing plans for the future (68%). However, it is unclear which outcomes were gained by those having surgery for alignment only.

Visual function

The VFQ-25 questionnaire is used to measure self-reported visual function and the AS-20 questionnaire is reported to measure HRQoL (section 2.3.2). Visual functioning questionnaires have measured improved visual function after strabismus surgery (VF-14) (Kishimoto & Ohtsuki, 2012). The VFQ-25 was compared to the AS-20 in a prospective study (n=106). In those without diplopia (n=26), the AS-20 was better able to discriminate between surgical success (total or partial) and failure than the VFQ-25, however VFQ-25 scores did improve. In those without diplopia, successful outcomes had significantly higher VFQ-25 scores (composite score, all vision-specific subscales, driving subscale and colour vision subscale) (Hatt et al., 2010a). Akbari et al. (2015) reported good correlation between the AS-20 and VFQ-25 (Persian versions), but did not analyse their results based on surgical success. Jackson et al. (2006) used visual analogue scales to report coping, lifestyle, worry, noticeable strabismus, and strabismus severity on a 0-10 scale, as well as the DAS-24, HADS and the WHOQoLBref. Strabismus surgery (n=46) resulted in significant improvements in QoL, psychological and physical functioning, which were greater in those without diplopia.

AS-20

The AS-20 (Hatt et al., 2009b) has become the most commonly used HRQoL questionnaire in strabismus (Adams et al., 2016; Alam et al., 2014; Glasman et al., 2013; Hatt et al., 2010a; Hatt et al., 2012a, 2016; Hatt et al., 2018; Ji et al., 2020; Koc et al., 2013; Liebermann et al., 2014; McBain et al., 2016b; Sim et al., 2018). Despite not being specific to strabismus with psychosocial symptoms, surgery in these patients has improved both psychosocial and functional aspects of the AS-20 (Alam et al., 2014; Hatt et al., 2010a; Hatt et al., 2012a; Koc et al., 2013). Liebermann et al. (2014) reported all AS-20 functional elements improved postoperatively in patients without diplopia (n=20), with the greatest improvements in stress, worry, needing to take breaks, enjoying hobbies, depth perception and eye strain items. Significant improvements in self-reported visual function after strabismus surgery for psychosocial reasons are difficult to explain, as no visual change was measured using standard clinical vision tests. However, it is possible that a change in binocular field of vision may have occurred as this was not tested (Kushner, 1994; Wortham & Greenwald, 1989).

The AS-20 and A&SQ were used in a prospective study of adult strabismus surgery outcomes (n=61) (Koc et al., 2013). Both questionnaires measured significant improvements in HRQoL 3-months postoperatively. Those with BSV postoperatively had significantly greater improvements in HRQoL scores on the functional subscales than those without BSV, but only when amblyopes were removed from the analysis. The change in overall scores and psychosocial scores (using

both questionnaires) were not significantly different between those with and without BSV, highlighting that visual benefit postoperatively was not required for improvement in HRQoL.

Alam et al. (2014) used the AS-20 in a cohort of older children and adults undergoing first strabismus surgery (n=30). None had diplopia, but it is unclear whether any had BSV. Significant improvements in AS-20 HRQoL were measured 6-weeks and 3-months postoperatively, with a greater improvement in females. Glasman et al. (2013) reported larger improvements in HRQoL (AS-20) in females, those with larger changes of the deviation and those with smaller strabismus postoperatively. Their prospective study of 17-76-year olds (n=86) found surgery led to improvements in all aspects of the AS-20, however BSV and diplopia were not reported. Their cohort may therefore have included some surgery for visual benefit.

Adams et al. (2016) used the AS-20, as well a large battery of QoL and psychosocial measures in a prospective study of patients aged 17-88 years (n=210). A range of aetiologies of strabismus were included and it is unclear how many had surgery for psychosocial reasons, however 44% had no diplopia. Postoperatively there was a reduction in the number of patients reporting poor AS-20 HRQoL, from 85% to 68% at 3 and 6-months postoperatively. Other measures of social anxiety and avoidance, clinical anxiety, and depression also improved significantly. In a study reporting the same cohort (n=210) McBain et al. (2016b) used the AS-20 as the primary outcome measure. Strabismus surgery resulted in significantly improved HRQoL 3-months postoperatively, with no further improvements at 6-months. Improvements in HRQoL were not associated with clinical judgements of success, highlighting that clinical definitions of success may not adequately capture the postoperative result from the patient's perspective. Postoperatively there were improvements in a wide range of psychosocial domains, as well as all aspects of the AS-20 (McBain et al., 2016b). Using a questionnaire to evaluate self-consciousness, Estes et al. (2020) reported improved public (but not private) self-consciousness and improved social anxiety 6-months postoperatively. It is unclear how many of their cohort (n=95) had surgery for psychosocial reasons, as some had diplopia (66%) and depth perception (62%) preoperatively. Using a range of psychological measures Ozates et al. (2019) demonstrated that constant XT (and no BSV) was significantly worse than intermittent XT and BSV. Surgery resulted in significant improvements in all psychological measures for the constant XT group, to the extent that there was no difference between constant and intermittent XT postoperatively.

Patients without diplopia reported significantly lower AS-20 psychosocial subscale scores preoperatively compared to those with diplopia. Interestingly, AS-20 function subscale scores were not significantly different. Postoperatively psychosocial and function subscale scores improved in all patients. Although the improvement in psychosocial subscale score was higher in those without diplopia, they continued to report lower postoperative psychosocial subscale scores than those

with diplopia initially. The only factor predictive of a greater improvement in AS-20 HRQoL was socioeconomic status. Those from a more deprived area had a higher rate of success postoperatively (Sim et al., 2018)

Hatt et al. (2010a) reported strabismus patients without diplopia gained significant improvements in AS-20 HRQoL, particularly if they had a successful result. Even those with 'failure' postoperatively reported AS-20 improvements, leading Hatt et al. (2016) to suggest success should include HRQoL improvements (beyond test-retest variability), in addition to improved alignment and diplopia. Having a distressed personality type, worse diplopia or depressive symptoms postoperatively and coexisting facial abnormalities were associated with postoperatively failure, using the AS-20 as the outcome measure (Hatt et al., 2018). These results highlight that mental health as well as clinical factors influence the outcomes from strabismus surgery, a view shared by others (Adams et al., 2016; McBain et al., 2016b). Hatt et al. (2012a) retrospectively reported outcomes in adults between 5-22 months postoperatively (n=73), described as 1-year results. Those who continued to meet success criteria (less than 10PD alignment and no or rare symptoms) maintained improved AS-20 results at 1-year compared to 6-weeks postoperatively. From 6-weeks to 1-year, those without diplopia showed stable function subscale results and further improved psychosocial subscale results. Ji et al. (2020) used the AS-20 (Chinese version) to investigate successful outcomes 1-year postoperatively, using similar success criteria to Hatt et al. (2016). Patients with BSV and diplopia were included. Despite successful strabismus surgery, 24% of their cohort (n=91) still reported they had strabismus. Those who perceived a deviation postoperatively reported lower AS-20 scores and were more likely to have a larger vertical deviation (Ji et al., 2020).

Whilst motor outcomes (strabismus size) may be more likely to define surgery as successful, the method of AS-20 analysis has been shown to affect the results (Leske et al., 2010). Change in either AS-20 subscale, greater than 95% limits of agreement, was considered difficult to achieve (Hatt et al., 2012b) but relying on motor outcomes only may fail to capture improved symptoms or HRQoL (Hatt et al., 2016). The AS-20 was considered to have excellent test-retest variability and a low chance of a ceiling effect. A change in overall score of 14, psychosocial subscale score of 17.7, and function subscale of 19.5 were considered evidence of real change. Whilst different results were provided for those with and without diplopia (Leske et al., 2010), it was unclear whether the "without diplopia" subgroup was strabismus with psychosocial symptoms only, as it may have included strabismus with BSV. A later evaluation of the AS-20 using Rasch analysis suggested refining the questions, the response options and the subscales to increase responsiveness to change in QoL. This resulted in removal of two questions from the previous function subscale (Leske et al., 2012).

Surgery has been reported to improve and normalise symptoms of anxiety and depression, HRQoL, daily functioning and psychological adjustment postoperatively (Jackson et al., 2006), however others report improved, but not normalised HRQoL (Xu et al., 2016). Patients who perceived they had no strabismus postoperatively achieved greater HRQoL improvement (Xu et al., 2016). Kim et al. (2016) used a self-identity questionnaire to evaluate young adult males at a military service examination. Having strabismus negatively affected self-identity compared to those with no strabismus and those who had previously undergone strabismus surgery in childhood. There was no difference in self-identity between those who had previous strabismus surgery and those without strabismus. With the recognition that QoL is an important outcome from strabismus surgery, focus has shifted to consider whether psychosocial interventions preoperatively could improve QoL and psychosocial outcomes. No trials have been yet been undertaken and this is an area for future research (MacKenzie et al., 2016).

3.2.5 When to measure postoperative outcome

Clinical care of patients following strabismus surgery is likely to vary among different clinicians, hospitals, healthcare systems and countries. Some may be discharged at a specific time point if they are asymptomatic and happy with the surgical result, yet others may be kept under longer review. Strabismus surgery outcomes have been reported at 1-week (Berland et al., 1998), 2-weeks (Dawson et al., 2013), 1-month (Kim et al., 2008), 6-weeks (Alam et al., 2014; Fatima et al., 2009), 3-months (Adams et al., 2016; Alam et al., 2014), 6-months (Adams et al., 2016; Lipton & Willshaw, 1995), 1-year (Currie et al., 2003; Dadeya et al., 2002; Jung & Kim, 2018) and later than 1-year (Currie et al., 2003; Felius et al., 2001). In some studies, the time at which outcome is being reported is unclear (Nelson et al., 2008). Reporting 1-year postoperative outcomes has the advantage of providing longer term data, yet many patients have been discharged and less data available for analysis (Liebermann et al., 2013). Longer term postoperative outcomes may therefore be biased and include a greater proportion of poorer outcomes that have not been discharged.

The last available follow up (Al-Wadaani, 2017; Beauchamp et al., 2003; Berland et al., 1998; Faridi et al., 2007) was commonly used to report postoperative outcomes, but even this was variable. Kim et al. (2008) reported postoperative outcomes following reoperation for sensory strabismus 1-month postoperatively and at the final postoperative visit, which ranged from 1-48 months. In contrast, the last available follow up visit ranged from 6-weeks to 13-years in a study of later surgery for childhood onset ET (Kutschke & Scott, 2004). Specific longer-term studies reporting outcomes after more than 1-year are less common but offer a unique view of postoperative stability and change over time. For example, 2-9 year follow up (Bayramlar & Gunduz, 2006), 3-9 year follow up (Keskinbora et al., 2011) and 10-year follow up (Gigante et al.,

2018) have been reported. A unique Swedish prospective study invited adults who had surgery for childhood ET for review and reported 32-44 years follow up (Ganesh et al., 2011).

On balance, evaluation of strabismus surgery outcomes at, or later than, 3-months represented a useful and achievable time point, unless measuring longer term outcomes were the specific aim. For most patients this was thought to allow sufficient time for healing (Escardó-Paton & Harrad, 2009), for eye alignment to stabilise and for the patient to adapt to their eye position. Measuring QoL outcomes at 6-months postoperatively was not significantly different to 3-months (McBain et al., 2016b).

3.2.6 Delphi study and core outcome sets

A recent Delphi study attempted to identify areas of consensus and disagreement amongst Ophthalmologists when defining success following strabismus surgery (Serafino et al., 2019). A range of different strabismus types and aetiologies were included, however some of the questions included in the Delphi study were pertinent to adults with strabismus and psychosocial symptoms. Of relevance to the current study was the lack of consensus reached on the time point at which postoperative outcomes should be evaluated, the deviation size considered successful postoperatively and how the deviation should be measured (SPCT, APCT or both). Consensus was reached in support of some strabismus conditions having unique outcome criteria (for example, 6th cranial nerve palsy) and for BSV outcomes to be included in the definition of success for some strabismus types. Al Jabri et al. (2019) have also highlighted the difficulty in comparing studies reporting strabismus outcomes due to a lack of ‘core outcome measures’ used. The COMET Initiative ("COMET Initiative: Core Outcome Measures in Effectiveness Trials,") aims to encourage core outcome set development and use in clinical trials. A core outcome set is the minimum set of measurements that should be taken and reported in a clinical trial of a specific condition. Core outcome sets are therefore useful as they allow comparison of study results and outcomes across different studies. Al Jabri et al. (2019) identified the outcome measurements most commonly used and reported in amblyopia, strabismus and ocular motility disorder studies, as well as highlighting that consensus was required to develop core outcome sets for trials and research into these conditions. Of note in the strabismus studies were the most commonly reported core outcome measurements of a near and distance measurement of the deviation, binocularity, HRQoL and adverse events, with some studies additionally reporting visual acuity and control of the deviation.

3.3 Additional outcomes from strabismus surgery

Patients undergoing strabismus surgery for psychosocial reasons may achieve more than just psychosocial benefit, as shown by QoL or HRQoL improvements. Observational studies reporting additional postoperative changes are discussed in detail below.

3.3.1 Visual field

Patients have gained an enlarged peripheral visual field following surgery to reduce ET (Kushner, 1994; Murray et al., 2007; Wortham & Greenwald, 1989). Wortham and Greenwald (1989) reported ten patients with ET who postoperatively gained peripheral visual field, gaining a mean 16-degrees horizontally (range 5-30 degrees). Visual field size was measured using the Goldmann perimeter, I4e target. The gain in peripheral visual field was ipsilateral to the strabismic eye and occurred even in the presence of amblyopia (n=3). Three patients gained some stereopsis postoperatively (range 80” of arc to Titmus fly). This suggested in patients without BSV the suppressed eye contributed to the peripheral field of vision. It also suggested that aligning the strabismic changed the amount, or extent, it contributed to peripheral visual field. Anecdotally four patients reported visual improvement, however patients were not asked to subjectively report visual change postoperatively. No follow up data were presented and comparisons with other patients were lacking. Murray et al. (2007) reported on older children and adults (n=17) with untreated infantile ET gained an expanded field of binocular vision (mean 32 degrees) postoperatively. However, in contrast to Wortham and Greenwald (1989) sensory fusion was always achieved in addition to binocular field expansion (Murray et al., 2007). Kushner (1994) reported that 34 of 35 patients (age 16-62 years) gained an expanded field of binocular vision postoperatively. The patient that did not gain field of binocular vision (n=1) had unilateral poor vision and retinal abnormalities secondary to uveitis, which may have affected the postoperative outcome. Of those who gained field of binocular vision (n=34), 29 had sensory fusion and 5 had suppression postoperatively (Bagolini glasses). Most patients that gained field of vision postoperatively were aware they had improved peripheral vision.

3.3.2 Unexpected binocular vision

Despite surgery for planned psychosocial benefit, unexpected BSV may occur postoperatively. For example, patients with longstanding large angle strabismus (n=8) have achieved good stereopsis, mean 45” of arc (Titmus) (Ball et al., 1993). Eight patients (out of 20) achieved 60-400” of arc (Frisby Near Stereotest (FNS)) or 40-80” of arc (Frisby Davis distance stereotest (FD2)) 1-year postoperatively (Liebermann et al., 2014). Detailed reports of pre and postoperative investigations of BSV in patients with strabismus are lacking. Retrospective studies aiming to identify factors that predict BSV postoperatively can lack complete outcome data (Umazume et al., 1997), leading to difficulty providing data on the proportion of patients who may achieve unexpected BSV or factors

that may predict BSV postoperatively. These factors highlight the importance of assessing potential BSV preoperatively (Ball et al., 1993) and BSV outcomes postoperatively, even when it is assumed no BSV is possible (Murray et al., 2007).

3.3.3 Binocular summation

Strabismus surgery has been reported to improve binocular summation, with a greater effect measured using lower contrast (1.25%) acuity charts. This improvement can mean binocular summation is measured postoperatively, despite binocular inhibition preoperatively. Successful surgical alignment and later onset strabismus have both been associated with greater improvements in binocular summation postoperatively (Pineles et al., 2015). Yet, other studies have shown highly variable changes in binocular summation following strabismus surgery (Chang et al., 2017). Interpreting postoperative binocular summation data only, rather than change as a result of surgery, and interpretation of binocular summation data in isolation, rather than as part of an investigation of pre and postoperative BSV may also be misleading. Further evidence is required to establish whether binocular summation improves following all strabismus surgery, or whether BSV and stereopsis (pre and postoperatively) affect the binocular summation outcome (Kattan et al., 2016).

3.3.4 Task performance

Patients have reported improved ability to perform daily activities (Nelson et al., 2008) and being able to work better (Ghiasi et al., 2013) when completing questionnaires postoperatively. Improved AS-20 function subscale results have been measured postoperatively even though patients have undergone surgery specifically for psychosocial symptoms or had no measurable visual change postoperatively (Alam et al., 2014; Hatt et al., 2010a; Hatt et al., 2012a; Koc et al., 2013; Liebermann et al., 2014). Few studies have measured task performance before and after strabismus surgery. Lee et al. (2013) used a spatial localisation pointing task presented on a touch screen to measure pointing accuracy in patients pre and post XT surgery. Pointing accuracy was reduced 1-day postoperatively, but accuracy improved to preoperative levels at 1-month postoperatively (Lee et al., 2013).

3.3.5 Eye movements

Using a photoelectric eye tracker, Bucci et al. (2009) measured the accuracy and mean velocity of saccades, convergence and divergence, and combined saccades and vergence eye movements, pre and postoperatively. Nine subjects (children and adults) with strabismus were included, six with no BSV pre and postoperatively, although diplopia was not mentioned. Preoperatively, compared to normative data, accuracy was reduced for vergences and combined saccades and vergence; and mean velocity was reduced for saccades and convergence. Postoperatively, accuracy

improved for saccades (at near), vergences and combined saccades and vergence; and mean velocity improved for convergence and combined saccades and divergence.

3.4 Discussion

Most of the evidence describing the outcomes of strabismus surgery in patients without visual symptoms reported improved postoperative ocular alignment and / or improved HRQoL. Yet, even in patients seeking surgery for psychosocial reasons, QoL and HRQoL measures are not used consistently. Different questionnaires have been reported to measure QoL and HRQoL in adults with strabismus. None of the questionnaires were exclusively for strabismus with psychosocial symptoms, however the AS-20 was developed specifically for adults with strabismus and was the most commonly used HRQoL questionnaire and PROM.

There are variable reports of the outcomes of surgery, the time at which outcomes are measured and a lack of consensus on how we should define and measure success after strabismus surgery. Describing the time at which the postoperative outcome was reported is recommended and on balance, the outcome from strabismus surgery appeared to be measured satisfactorily at, or around, 3 months postoperatively. However, there is acknowledgement that the postoperative outcome at 3 months may differ from the longer-term outcome, which may improve further. Additional surgical outcomes, including an expanded field of vison, unexpected BSV, improved binocular summation, improved task performance and improved eye movements have been suggested, but have not been fully investigated. A core outcome set for strabismus has been suggested and there is potential to add to the available evidence by investigating which outcome measures are most relevant to those with strabismus and psychosocial symptoms.

The difficulty in establishing accurate criteria for postoperative success and failure has been highlighted. Similar criteria for ‘success, partial success and failure’ have been used by several studies attempting to categorise and compare surgical outcomes. However, there is the potential to refine and improve these categories, as patients categorised as a failure postoperatively can still report significant improvements postoperatively.

The main gap was the lack of evidence specifically reporting the outcomes of strabismus surgery in adults with psychosocial symptoms. Large heterogeneous cohorts of strabismus patients were often reported, typically with a range of symptoms and differing surgical aims. There was a growing need for robust evidence in this specific subgroup of patients with strabismus and psychosocial symptoms. Some commissioners of NHS services have reported concern that not enough benefit from treatment was proven in adults with strabismus and psychosocial symptoms only. Some areas of England are therefore selectively funding strabismus surgery only for adults with expected

visual benefit postoperatively (Bristol North Somerset and South Gloucestershire Clinical Commissioning Group, 2019).

3.5 Chapter 3 summary

This chapter has reviewed the evidence reporting the outcomes of strabismus surgery undertaken for psychosocial reasons. These patients are expected to gain improved eye alignment and improved QoL or HRQoL postoperatively. Additionally, there was also evidence reporting changes in visual field, task performance, daily activities and eye movements can occur. There is a need to address the gaps in the current evidence, due to the risks to NHS funding for strabismus surgery in adults with psychosocial symptoms. The following chapter will describe the rationale for this study, the research questions, and the study aims and objectives.

Chapter 4. Research aims and objectives

This chapter will briefly describe the rationale, aims and objectives for this study.

4.1 Study rationale

As described in chapter two, having strabismus can cause visual and psychosocial symptoms. Psychosocial symptoms can cause difficulties in a person's daily life, poorer HRQoL and poorer mental health. In addition, vision and task performance difficulties may occur, yet these are rarely measured. Chapter two highlighted the withdrawal of NHS funding for adult strabismus surgery for psychosocial reasons in some areas in England, due to concern that not enough patient benefit from surgery was proven. Robust evidence in this specific subgroup of patients is therefore needed, particularly evidence of postoperative outcomes and patient perspectives of the outcomes. In chapter three, the systematic literature review demonstrated surgery in adults with strabismus and psychosocial symptoms can improve eye alignment and QoL. However, there are gaps in the evidence around other possible outcomes, including whether an expanded visual field, improved task performance and improved eye movements can occur postoperatively. The rationale for this research is to address these gaps in the evidence by investigating the outcomes of strabismus surgery in the specific cohort that are at risk of losing funding for surgery, adults with strabismus and psychosocial symptoms only.

A mixed methods study design was selected to combine qualitative and quantitative research methods. The qualitative phase one explored the outcomes of surgery from the patient's perspective. The findings informed the measures selected for the quantitative phase two. Phase two measured different aspects of vision, task performance, physical symptoms and HRQoL before and after surgery. The feasibility of different aspects of the study were evaluated to inform the design of a future larger trial. The timeline of the study is shown in Figure 4.1

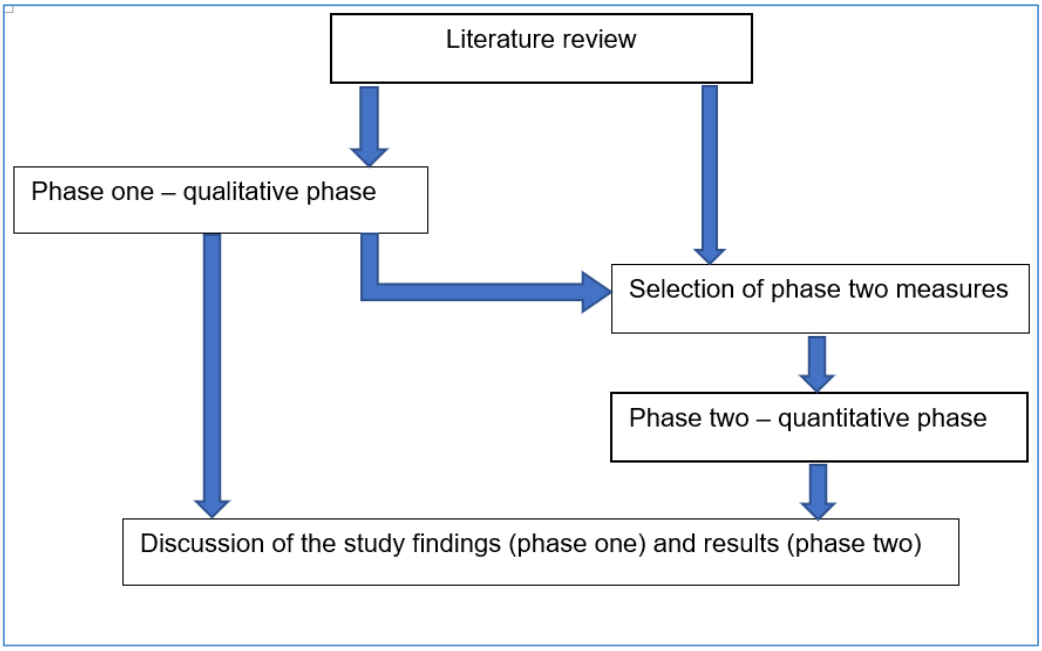


Figure 4.1 Flowchart showing the general timeline of the study

4.2 Research aims

The aims of this feasibility study were to use a mixed methods design to investigate the outcomes from strabismus surgery in a cohort of adults having surgery for psychosocial reasons, and to measure whether benefit could be gained from strabismus surgery in addition to the known psychosocial benefits. The aim is for the results to be disseminated to clinicians and academics to add to the existing evidence of surgical outcomes in this specific patient group. Additionally, the aim was for the study results to be used to inform future updates of current clinical guidelines and recommendations (Das et al., 2017) and future strabismus surgery funding decisions.

The primary research question was:

Overall mixed methods study

1. What are the effects of strabismus surgery undertaken for psychosocial reasons?

The secondary research questions were:

Qualitative phase one

2. What do adult patients who have undergone strabismus surgery for psychosocial reasons feel has changed, or not changed, postoperatively?

Quantitative phase two

3. Is it feasible to measure changes in visual function and task performance after strabismus surgery for psychosocial reasons?
4. Are the tests of visual function and task performance acceptable for patients?

5. Can the acceptable tests of visual function and task performance be refined for a future larger study?
6. Is it feasible to recruit and retain patients to a study that requires measurements of visual function and task performance before and after surgery?

4.3 Research objectives

Qualitative phase one

1. To conduct semi-structured interviews to explore patient perceptions of their outcomes from strabismus surgery for psychosocial reasons
2. To elicit information from the thematic analysis of the semi-structured interviews to refine the design of, and select the measures to be included in, the quantitative phase two

Quantitative phase two

3. To recruit patients who are due to undergo strabismus surgery for psychosocial reasons (surgery group) and volunteers with strabismus, but not planning to have surgery (control group)
4. To conduct phase two, a quantitative repeated measures study, to try and measure the changes patients reported during the qualitative phase one

Overall mixed methods study

5. To analyse and integrate the results from both phases of the study to determine the outcomes of strabismus surgery undertaken for psychosocial reasons
6. To disseminate the results to clinicians and academics by presenting the results in publications and at conferences, so that the results can be used to inform future updates of current clinical recommendations and inform NHS funding decisions.
7. To consider and answer the feasibility and acceptability questions raised by the study (described in section 4.2)

4.4 Chapter 4 summary

This chapter has described the rationale for this study, which will provide new data on the effects of adult strabismus surgery undertaken for psychosocial reasons. The aims and objectives for the overall mixed methods study and the qualitative and quantitative phases have been presented. In chapter 5 the qualitative phase will be introduced, and the qualitative methods described.

Chapter 5. Qualitative Phase: Introduction and Methods

This chapter outlines the rationale for selecting a mixed methods approach for the overall study. The study design for the qualitative phase one is described, including the sampling strategy, recruitment of participants, data collection and analysis. The quantitative phase two will be described later in chapters 7 and 8.

5.1 Mixed methods methodology

A research study may be purely qualitative, purely quantitative, or mixed, in methodology. A research study that includes more than one type of data collection method may be selected to try and broaden the type of data generated (Mason, 2002). A mixed methods research methodology relies on integrating qualitative and quantitative parts within a study, rather than simply including them both as component parts (O'Cathain & Thomas, 2006). A mixed methods approach may be selected to better understand a research problem by combining qualitative and quantitative study data. For example, identifying variables or themes in a qualitative study that can be measured in a quantitative study, or using quantitative data to identify individuals who may expand on the results in a later qualitative study (Hanson et al., 2005). The selection of a mixed methods approach should be supported by reasoned arguments and the rationale behind the methodological choices made (Cameron, 2011).

Combining qualitative and quantitative methods in a mixed methods study can be achieved in different ways (Hanson et al., 2005).The qualitative and quantitative parts can occur in simultaneously in parallel, or may occur sequentially (O'Cathain & Thomas, 2006). For example, the results from one part may inform the design or the selection of methods in the other; or the results may be analysed together (O'Cathain & Thomas, 2006). Whilst different approaches to mixed methods research exist, the common feature is integration of the qualitative and quantitative elements. The purpose of mixing the methods, the sequence of mixing, whether the qualitative and quantitative parts will be equal or unequal, how and when the methods will be integrated should all be considered (O'Cathain & Thomas, 2006).

5.1.1 Theoretical unpinning

In qualitative research, acknowledging that the way we view knowledge and the world around us is important. Our view of knowledge (epistemology) informs our theoretical perspective, or the philosophical stance taken when deciding how knowledge can be demonstrated. Our theoretical perspective also informs our methodological approach and the methods we select (Crotty, 1998). Stating the philosophical underpinning of the research study and the chosen methodological

approach are important parts of acknowledging the researcher's stance from which they have planned their study, as well as viewed and interpreted the results (Hanson et al., 2005).

Silverman (2013) described naturalism as a model that focused on how phenomena are experienced. Observations of factual characteristics or social settings may be analysed to try and describe what an experience or phenomena are really like. Research from a naturalist perspective would tend to involve questioning 'what is going on?' when phenomena are being observed, often in their natural setting. Alternatively, constructivism was described as a model that focuses on how phenomena are constructed in people's everyday activities or lives (Silverman, 2013). Observations of experiences would instead be viewed from the perspective of how they are socially constructed in a particular context. Research from a constructivist perspective would tend to involve questioning 'how is that being perceived?' when phenomena are being observed. The philosophical background to this mixed methods study was from a constructivist perspective, to try and understand how and why patients perceived their lives had changed, or not changed, following strabismus surgery.

Overall mixed research study

A mixed methods methodology was selected with the aim of exploring how patients perceived their lives had changed (or not) postoperatively (phase one), then measuring whether strabismus surgery could lead to additional benefits, in addition to improved eye alignment and improved HRQoL (phase two). This sequential mixed methods design, with the qualitative and quantitative data having equal priority, aimed to generate more insights than would be possible from a qualitative or quantitative study alone.

Analysis of the phase one qualitative findings contributed to the development of theories and hypotheses to be tested, and the selection of the phase two measures (described in chapter 7). Additionally, the phase one qualitative findings and the phase two quantitative results were brought together with equal priority and discussed together in chapter 10. The integration of the results aimed to answer the same research question, but in greater depth (Hanson et al., 2005; O'Cathain & Thomas, 2006).

5.1.2 Qualitative data collection

Qualitative interviews may be unstructured, structured or semi-structured (Mason, 2002). Unstructured interviews would be directed at a topic of interest; yet there would be no pre-set questions to ask or a list of topics to include during the interview. An unstructured interview aims to allow the interviewee to talk openly and freely about the topic and for the interview to flow based on the interviewee's individual experiences. A structured interview would include a pre-determined list of specific questions that are asked to the interviewee during the interview. The questions would not typically be changed or adapted for different interviewees, to ensure all interview participants

received standardised interviews. Semi-structured interviews are a technique in between unstructured and structured interviews. They typically utilise a topic guide, including different topic areas during the interview; yet the questions would be worded openly to encourage the interviewee to talk freely about the different topics (Silverman, 2013). The topic guide may be adapted after each interview to ensure the subsequent interviews include topics arising from previous interviews.

Interviews may be conducted with a participant face-to-face, or using technology to aid communication such as telephone, video call or video conferencing. It is considered good research practice for the researcher to make field notes after qualitative data collection to assist with later data analysis (Mason, 2002). Audio and video recordings of an interview are transcribed before analysis.

Within ophthalmic research, qualitative research studies have increased. Irving et al. (2018) highlighted gaps in patient understanding of sight risk and eye care in Canada. Senthil et al. (2019) found several aspects of QoL were affected in retinal vein occlusion and used the results to develop a QoL measure. Lu et al. (2019) identified that many patients required support to access glaucoma care, despite valuing their vision and trusting the clinician. In retinitis pigmentosa a range of qualitative evidence has emerged, enabling a meta-synthesis of the different sources of qualitative evidence (Garip & Kamal, 2019).

For this study interviews were selected as the qualitative data collection method after consideration of expected recruitment and the timescale of the overall study. A pragmatic decision was made to recruit participants attending eye clinic appointments. The ability for the topic guide to be expanded iteratively during semi-structured interviews was preferred over a structured interview, to allow for new information from participants to emerge from the interviews and be included in subsequent interviews.

5.1.3 Qualitative sampling strategy

Participants should be selected to ensure they represent the population being researched, whilst minimising potential sources of bias (Bryman, 2016). In qualitative research, purposive sampling attempts to overcome a source of bias. Participants are selected based on characteristics that are relevant to the research question, ensuring there is maximum variation within the sample (Mason, 2002). Phase one used purposive sampling of adults with strabismus from the eye clinic at Sheffield Teaching Hospital NHS Foundation Trust (STH NHS FT). This sampling strategy was considered the most practical way to recruit participants who had undergone strabismus surgery for psychosocial reasons within the time frame of the study. Interviews postoperatively, rather than pre and postoperatively, were conducted due to the timescales of the overall study.

5.1.4 Analysis of qualitative interviews

Different techniques are used to approach qualitative data analysis from the theoretical background of constructivism including grounded theory (Charmaz, 2014), narrative analysis, discourse analysis, conversational analysis and interpretive phenomenological analysis (Silverman, 2013). This study used a thematic approach, underpinned by grounded theory to analyse the qualitative data. The interviews were approached without a predetermined hypothesis and theories emerged from the interview data (Charmaz, 2014). Simultaneous data collection and analysis occurred, using the 'constant comparative method' to continually compare new data to developing theories, to widen categories and further develop theories (Charmaz, 2014). This ensured data analysis was based on theory that was grounded in the data (Silverman, 2013). Data collection continued until no new information emerged from the data and 'data saturation' of the categories was reached (Charmaz, 2014).

5.2 Study setting

The setting for the mixed methods study was the Ophthalmology clinic at STH NHS FT, a tertiary referral centre at a large UK teaching hospital. Strabismus patients were typically referred from the Sheffield area or wider Yorkshire and Humber region. They would routinely see an Orthoptist and one of three Ophthalmologists specialising in adult strabismus. Approximately 150 adult strabismus surgeries are performed at STH NHS FT annually.

5.3 Phase one: qualitative methods

5.3.1 Ethical approval

Prior to commencement of phase one favourable reviews from the Research Ethics Committee (REC) (Appendix C) and the Health Research Authority (HRA) (Appendix C) were gained, in addition to local approvals.

5.3.2 Sampling

Purposive sampling of adults who had previously undergone strabismus surgery for psychosocial reasons was used. The characteristics used to ensure maximum variation were gender (male / female) and age group (younger 18-35 years old / older 36 years and above). Gender had been identified as a factor affecting perception of strabismus with females perceiving strabismus more negatively than males (McBain et al., 2014) but also perceiving greater postoperative improvements in QoL compared to males (Burke et al., 1997). Age has not been shown to be a factor in previous strabismus research, however the PPI group felt strongly that recruiting younger

and older patients was important. A pragmatic joint decision (researcher and PPI group) was made for the age cut off (35 years). The younger age group were 18-35 years old. The older age group were 36 years old and above. Ethnicity was not included as a characteristic in the sampling strategy, however no participants were excluded or included based on ethnicity. It is acknowledged that unintentional bias may have been introduced at the recruitment stage as GA was often the clinician assessing potential participants in the Orthoptic clinic to see whether they met the inclusion criteria for the study. This was managed by firmly adhering to the inclusion criteria and inviting all those who met the inclusion criteria for the study, regardless of their outcome.

Data saturation refers to the point at which no additional new information or insights are emerging from the data (Charmaz, 2014). Recruitment of sixteen participants was planned with the acknowledgement that data saturation may occur prior to this. Regular discussion took place with the study supervisors during the qualitative phase to review the possibility of data saturation as well as the overall study timescales.

5.3.3 Recruitment

Patients were invited to participate in a semi-structured interview when they attended a clinical appointment, provided they met the inclusion criteria (see below). The study was explained to them verbally, they were given written information (Appendix D) and the opportunity to discuss the study with the researcher. Participants were offered the choice of an interview following their clinical appointment or at a later date, and the choice of interview location (within the Orthoptic department or University of Sheffield). Written consent (Appendix E) was taken from each participant before their interview and following a brief discussion to ensure they understood the purpose of the semi-structured interview.

Inclusion criteria

- Adults (18 years old or greater, no maximum age limit)
- Strabismus surgery planned and undertaken for psychosocial reasons
- 3-12 months postoperatively

Exclusion criteria

- Unable to give informed consent
- Unable to complete a semi-structured interview in English

Minor amendment to approvals

Due to clinical delays the inclusion criteria of the postoperative period was extended from 3-12 months to 3-24 months postoperatively. This was submitted as a non-substantial/minor amendment to the Health Research Authority for approval (Appendix C) with the support of STH NHS FT Research and Development Department. Alternative options to improve recruitment were

considered and discussed with supervisors, including recruitment from other sites. Due to similar clinical delays at other sites within the region, widening recruitment to other sites was not pursued.

5.3.4 Topic guide

The topic guide (Appendix F) was developed following a review of the literature (chapters 2 and 3). Questions were purposely open-ended to allow for positive changes, negative changes or no change postoperatively. Open questions were asked about a range of different aspects of vision, task performance, daily activities and patient perceptions about their daily life. Positive and negative changes were explored to avoid biasing the interviews towards positive or negative postoperative outcomes. Participant experiences of postoperative outcomes were the specific focus of the interviews, however broader participant descriptions of experiences pre and postoperatively were not discouraged. Opportunities were given for participants to add information they felt was important or had not been covered during the interview. The topic guide was used to help guide participants through the planned topics and ask questions in alternative ways if required. Changes to the topic guide (Appendix F, red text) were made to add prompts, to encourage participants to expand on their responses, and new topics that emerged during interviews (for example driving).

5.3.5 The interviews

GA conducted all the interviews. Using an independent interviewer was considered, but this was deemed not possible for this study. Interviews lasted 15-40 minutes and were audio recorded using a digital recording device (Tascam Linear PCM Recorder DR-05 version 2). Interviews tended to become longer over time as the topic guide was expanded and the interviewer (GA) became more experienced at conducting the interviews and participants were encouraged to give examples to illustrate the points they were making. Field notes were recorded immediately after each interview to capture additional details, possible links between themes and contrasting information. All interviews were transcribed verbatim by GA using Express Scribe Transcription Software (version 6.10 ©NCH Software) as soon as possible after data collection. Each audio recording was listened to again, later, to check transcript accuracy and to correct any mistakes and omissions. Transcripts were anonymised and stored in NVivo (QSR NVivo 10).

5.3.6 Qualitative data analysis

The analysis of the qualitative data was based on thematic analysis following the principles of grounded theory (Charmaz, 2014). One transcript with rich data was read and initially coded by me (GA) and one supervisor (SB) independently. Our coding frames were compared and discussed. A final coding frame was agreed and applied to all the transcripts.

Audio recordings were listened to multiple times. Transcripts were read and re-read to enable familiarisation with the data. Following each interview, constant comparison methods were used to ensure there was simultaneous data collection and analysis, both within and between interviews. Constant comparison was supported by note taking and memo writing.

The codes were used to organise and display the information in NVivo (QSR NVivo 10). The 'one sheet of paper' (OSOP) technique (Ziebland & McPherson, 2006) was used to look at higher level themes that emerged from the data, and to help search for connections and patterns within and across codes and themes. The OSOP technique was used to include all of the themes, examples from the transcripts, codes and categories, including similar information and differing information, or 'deviant cases' (Silverman, 2013). A summary of the initial codes and categories, the development of the themes and the final themes are displayed in Table G-1 (Appendix G).

During transcription and the application of the coding framework to the later interview transcripts (transcripts for participants 8-10), it became apparent that data saturation may have been reached, as no new codes or categories were emerging from the data. Following discussion with supervisors (SB, DB and HD), data collection was stopped after 13 participants had been recruited. Further data collection to achieve a sample size of sixteen, considering the other milestones and timescales of the study, was therefore considered not necessary.

5.3.7 Reflection on qualitative methods

Whilst all researchers should approach a research topic or problem with an open mind and no preconceptions, it is recognised that my background as an Orthoptist included implicit theories and assumptions. All efforts were made to remain open to patient descriptions of positive, negative or neutral experiences postoperatively. However, it is acknowledged that my assumptions and background knowledge may have introduced unintended bias during the interviews or data analysis. Unintended bias may also have been introduced due to the variable postoperative time period at which the interviews were conducted (described in section 6.1) and participant expectations of the purpose and aims of the study. All attempts were made to keep the study information standardised (Appendix D) and not biased towards positive or negative feelings. It is acknowledged that my role as a clinician and interviewer may have influenced participant behaviour during the interviews. Additionally, participant recall of postoperative experiences may have influenced their memories and descriptions of preoperative experiences.

5.4 Chapter 5 summary

This chapter described the rationale for selecting a mixed methods methodology and the theoretical underpinning of the qualitative phase one. The qualitative phase one setting, methods,

recruitment and analysis were described. Chapter 6 will describe the qualitative phase one findings.

Chapter 6. Qualitative Findings

This chapter will present the qualitative interview findings. The interviews aimed to explore adult perceptions of change, or no change, following strabismus surgery undertaken for psychosocial reasons. Postoperative change included positive change (improvement) or negative change (worsening). Throughout the interviews open questions were asked and precautions were taken to avoid demand characteristics and introducing bias.

6.1 Participant characteristics

Thirteen participants were interviewed across the chosen characteristics, including younger females (18-35 years old) (n=4), older females (36 years old or older) (n=3), younger males (18-35 years old) (n=2) and older males (36 years old or older) (n=4). All participants (n=13) selected an interview on the same day as their clinical orthoptic appointment (between November 2017 and February 2018). All had undergone strabismus surgery for psychosocial reasons at STH NHS FT and were on average 12.2 months postoperatively (range 4.5-20 months). All had no diplopia and no potential BSV preoperatively and no diplopia or BSV post-operatively. All participants demonstrated constant suppression on sensory fusion testing and had at least perception of light vision in their poorest seeing eye. A summary of the participants and their clinical characteristics are shown in Table H-1 (Appendix H).

6.2 Qualitative themes

During the interviews, participants talked about their experiences of having strabismus. This included a range of experiences preoperatively or postoperatively, or a combination of pre and postoperative experiences. Four main themes emerged from the analysis of the interviews: vision, task performance, physical symptoms and confidence. These four themes are discussed in detail below. Direct quotes were used from a range of participants to illustrate the findings from each of the themes. Quotes were selected to support the findings presented, but also to ensure information from all participants was included.

6.3 Theme: Vision

Some of the discussion around vision related to problems encountered preoperatively, but almost all participants described ways in which their vision had improved postoperatively. This included how their vision had become clearer or they could focus better; improvements in their peripheral vision; changes in how they used their strabismic eye; control of their strabismus and the strabismus being less of a distraction. These issues are discussed in the following sections:

6.3.1 Vision – Preoperatively

Visual difficulties pre-operatively included vision being described as blurred, not clear, not straight, difficulty focussing and not being able to focus straight.

> *"a lot of blurred vision, it was mainly blur more than anything"*
> *Participant 012, Male, 18-35 years.*

Participants described finding it difficult in busy environments, for example shopping centres. Busy environments with lots of people and movement in the visual scene were perceived to be visually uncomfortable and distracting, leading to difficulties with processing visual information and focussing. Additionally, busy environments raised anxiety and social difficulties due to large numbers of people being present.

> *"I kind of avoided situations like that before, so if it was too busy it was like... maybe it was confusing, so I kind of avoided that type of stuff if I could"*
> *Participant 008, Male, 36+ years.*

Impaired eye movements were described as a consequence of having strabismus, with the perception that the strabismic eye was stuck in the corner and did not move with the other eye. Not being able to move the strabismic eye properly was perceived as a reason for not being able to see properly. Difficulties with peripheral vision were described, this was typically an awareness of having less vision, not being able to see fully to the side and needing to use head movements to compensate for reduced peripheral vision and see people or objects to the side.

> *"if someone was... stood at the side of me…. before I'd just not be able to tell or see or notice at all"*
> *Participant 002, Female, 18-35 years.*

None of the participants felt they used their eyes together preoperatively. Five participants specifically described feeling that they used their eyes independently. Yet, two participants described having some control over their strabismus preoperatively, in which they were able to make their eye position straighter. This ability however was associated with symptoms of eye strain and discomfort.

Having strabismus was reported to be distracting, leading to frustration with their eyes or with their vision. Strabismus was described as being worse when tired or unwell. There was a perception that the strabismic eye was getting in the way or causing participants to be confused about their symptoms and what the root cause of their problems was. Three participants described closing or covering the strabismic eye intermittently to see or concentrate. Closing the strabismic eye occurred most commonly during a specific activity, for example watching television; or to help with concentration during a task at work. Participants associated visual difficulties with symptoms of

dizziness or headaches, which were most common at work, in busy environments or during specific activities requiring concentration, such as driving.

> *"before the operation my eye would probably drift out and it would cause me to get blurry vision and a headache"*
> *Participant 012, Male, 18-35 years.*

6.3.2 Vision - Postoperatively

In addition to postoperative healing, participants described varying periods of adaptation to their new eye position postoperatively. Some were aware of immediate improvements in their vision, yet others described experiencing headaches or bumping into things during their early postoperative period.

> *"I did have some headaches but… I suppose that's to be expected, but that was also my brain getting used to looking through that eye"*
> *Participant 011, Female, 18-35 years.*

Clearer vision and better focussing

Better vision postoperatively was described in a variety of different ways by nine participants. One participant described some visual aspects that were better and some that were worse. In the majority of cases participants spoke about their overall viewing experiences with both eyes open. Participants felt they had better or improved focussing, describing being able to focus more, easier, quicker, or being able to focus their eyes more together. Vision was described as being more in focus leading to being able to focus better on tasks with both eyes open under natural viewing conditions, which participants described as feeling better and easier. Vision was described as sharper and clearer, with an awareness that this led to being able to notice more of their surroundings. Examples of tasks where clearer vision was noticed as beneficial include looking at screens, work related tasks and near activities such as artwork or cooking.

> *"I can definitely focus more now that I've had my squint surgery, on what I'm doing at my desk, on my computer"*
> *Participant 004, Female, 18-35 years.*

Vision was reported to be better or easier, leading to a perception of eyesight being more natural or staying the same and not going blurry. Vision was described as more central and straighter when viewing with both eyes open. This was reported to lead to a beneficial, and more comfortable, feeling of being able to tell that the strabismic eye is in the right place and being able to make more accurate judgements about the position and alignment of objects. Having better vision postoperatively was described by some participants as having more natural or easier vision, which was linked to being able to focus quicker, better or more on tasks. Whilst talking about their vision, activities and examples of activities of every-day life were typically used to describe vision

improvements. For example, participants described being aware of improved vision included driving, watching television, working on computers, walking around, reading signposts and social activities. Having better vision or being able to focus better was also described as enabling greater levels of concentration, longer concentration and being able to perform better at work.

> *"everything since the operation I've been absolutely pleased with, yeah everything, it's just made it a lot better… you've got a lot more, somehow, clearer view, everything's straighter"*
> *Participant 007, Male, 36+ years.*

Four participants reported improvements in their vision in busy environments postoperatively and five reported no change. Perceived improvements in vision whilst in busy environments included being able to focus quicker or easier, having less blurred vision and being less distracted by the busy environment. Postoperative improvements in these difficulties experienced in busy environments may also be linked to having less anxiety and social difficulties because of the improved eye alignment.

> *"it's harder to focus with just one eye than it is to have two eyes, so now when I'm just focussing on something I can just look and see it and then go, whereas before I would have to spend a couple of minutes to… focus on that object that I'm looking at and then go… when there's a lot of people you can't, you've only got a certain amount of time to see that before somebody's pushing in front of you… I don't feel like I have to focus so much, I can just see it and then go"*
> *Participant 011, Female, 18-35 years.*

Three participants described feeling that they had slightly better vision in their strabismic eye postoperatively, for example how vision was perceived when the fixing eye was covered or closed. This included being able to focus easier with their strabismic eye and vision looking straighter.

> *"before if… my good eye was covered up, I'd literally just see like feint colours and blurs… but now if that eye's covered up I can make out more shapes and what shapes are with it… and what the colours are and stuff"*
> *Participant 002, Female, 18-35 years.*

Peripheral vision

Participants would typically experience their peripheral vision as being able to see things at the side, even though they were looking at something straight ahead. An increased field of peripheral vision, or having more peripheral vision, was described by seven participants. This was perceived as better vision, clearer vision or having more vision than before the surgery. Examples of the impact of a greater field of peripheral vision for participants included them being able to see trip hazards, being able to see people standing at the side of them and not having to turn their head as much to see at the side. The personal consequences of having a greater amount of peripheral

vision were described as being more aware of personal surroundings and being safer, as well as feeling better.

> *"it's like being more aware of your surroundings… it helps you feel a bit more safe… I can see more of my surroundings, so you can be a bit more alert… just a bit better and a bit more safe"*
> *Participant 002, Female, 18-35 years.*

Often it is assumed by clinicians that patients with an XT and an expanded field of peripheral vision, or panoramic vision, like or benefit from an increased field of vision. Whilst this may be the case for some patients, one participant with an XT preoperatively disliked having panoramic vision and described this as being aware that his vision was not 'normal'.

> *"it's not normal is it to be able to see like an owl, round the side of your head"*
> *Participant 008, Male, 36+ years.*

Needing to use less head movement to see to the side was reported by four participants and was described as a direct consequence of having greater peripheral vision postoperatively.

> *"it's a lot easier I think, because I don't have to turn, well obviously you do have to turn your head, but I don't have to turn it as far as I used to… I can just look with my eyes, rather than turning my head fully"*
> *Participant 004, Female, 18-35 years.*

Five participants felt their eye movements had improved, this was most commonly described as being able to look further into different positions of gaze. Having improved eye movements was reported to lead to increased peripheral vision and needing to make fewer or less head movements to be able to see to the side. Improved eye movements were also described as the eyes feeling better during eye movements, which was perceived as better postoperatively. Additionally, having a reduced size of strabismus, or improved eye alignment, was perceived as enabling the eyes to move together more. The feeling of improved eye movements combined with the perception of the two eyes moving into different positions together was associated with the feeling of looking through both eyes at the same time.

> *"my squint surgery has definitely made the looking outwards a lot better… I can definitely see more to the side"*
> *Participant 006, Female, 36+ years.*

Using the strabismic eye

Five participants reported being aware that they did not use their eyes together postoperatively and described this as using their eyes separately and independently. However, two participants perceived that they were using their eyes together postoperatively. This was described as being

able to use both eyes to focus on everything at the same time, leading to improved and quicker focussing.

> *"I was trying to use my right eye a lot harder than my left eye, but now I can use them both together, so it's a lot easier to focus on what I'm doing and work, my emails that I'm reading, so that's good"*
> *Participant 004, Female, 18-35 years.*

Four participants reported they felt that they were using their strabismic eye more postoperatively. This was described as it being quicker or easier to take up fixation with the strabismic eye or having greater peripheral vision and needing to make less head movement to see to the side. This was perceived as happening automatically, leading to better and clearer vision. Although participants described using their strabismic eye more or feeling like their eyes were moving together in a coordinated way, they did not associate these perceptions with binocular vision or feeling like they were using their eyes together as a pair.

> *"I don't use both my eyes together… you do feel like... it's easier to use it a bit more now…. I would say more to the side, I mean it is definitely a lazy eye, it's still definitely a lazy eye, it doesn't like being used very often, so I use my other one, but… when I do use it… it is alright, it's a lot better"*
> *Participant 006, Female, 36+ years.*

Two participants felt they had gained control over their strabismus postoperatively, which was described as a positive outcome from surgery.

> *(could you control the squint?) "no, not before… (but now) a little bit I think, yeah… but yeah it seems a lot better now"*
> *Participant 006, Female, 36+ years.*

Two participants described improved vision in their strabismic eye postoperatively. Four participants reported overall improved vision due to feeling less confused, less distracted by their strabismus, or less like their strabismus was getting in the way. These perceptions of improved vision postoperatively were described in association with having less blurred vision and less eye strain. Having less confusion was perceived, by one participant, to be occurring because they felt they were able to look through both eyes postoperatively.

> *"yeah, 100% yeah, it just feels different, it just feels better, yeah, not as confused"*
> *Participant 009, Male, 36+ years.*

Six participants reported improved eye closure, this was described as needing to close or cover the strabismic eye for less of the time when looking at something. This was perceived to have

improved particularly during periods of concentration on a task, for example when watching television.

> *"I don't think I've done it as much as I used to, I still do it now and then, if I'm getting a headache"*
> *Participant 012, Male, 18-35 years.*

Descriptions of unchanged vision

Interestingly, two participants described their vision as being unchanged postoperatively, but then went on to give examples of improved visual ability during their interviews. It is possible that any changes in vision were subtle, not perceptible or of little practical benefit to the participants. However, it could also represent individual differences of understanding of the term 'vision' or perception of what is meant by 'vision' in the context of the question asked. For example, one participant seemed to associate 'vision' with being able to read letters on a VA testing chart in a clinical environment.

> *"there's no change in my vision, but it's helped me to see a bit better, because it's not going out and... like confusing, making it a bit blurry… it's made that better"*
> *Participant 001, Female, 18-35 years.*

6.3.3 Vision - Summary

Participants described a range of visual difficulties preoperatively, including difficulty focussing, their vision not being clear or straight, difficulty seeing in busy environments, strabismus being a distraction, and their eyes not moving properly. Postoperatively participants described a range of changes to their vision, including having clearer vision and better focussing, having more peripheral vision, needing to make less head movements to see at the side, and feeling better able to use their strabismic eye.

6.4 Theme: Task Performance

Participant discussions of task performance and their activities of daily life included their descriptions of how they felt their strabismus affected their task performance preoperatively. Some participants felt their task performance improved postoperatively and they described examples including driving, using computers and screen devices, work-based activities, reading and other near activities. These improvements are described in the following sections:

6.4.1 Task Performance - Preoperatively

Preoperatively difficulties with task performance and undertaking activities of daily life were described. Having strabismus was reported to cause difficulties with driving and reduced

confidence to drive, due to the strabismic eye drifting and the awareness that the strabismic eye wasn't looking straight ahead.

> *"I wouldn't have felt so confident, I wouldn't have wanted to drive looking over my nose if you know what I mean, no I wouldn't have done that"*
> *Participant 007, Male, 36+ years*

Work ability and performance at work were described as difficult due to strabismus. Operating tills and machines were reported to be challenging and having strabismus was associated with making mistakes during tasks at work. Strabismus causing difficulties at work was described as frustrating, causing a loss of concentration and a further reduction in work ability. Work performance was also reported to be negatively affected by having strabismus when work involved interacting with people.

Working on computers and looking at screens (computer screens, mobile phones and tablets) were all reported to be difficult, causing problems with focussing on computer-based tasks and needing to adjust screen brightness and zoom to be able to see images. Reading and performing tasks at near, such as artwork, were described as difficult due to strabismus. Avoiding reading, having problems seeing to read and closing one eye to perform near tasks were all described preoperatively.

> *"if I read for quite a while, you could feel it pulling… you could feel something was dragging it to the side, that's what it felt like anyway"*
> *Participant 009, Male, 36+ years*

Difficulties with balance, clumsiness and bumping into things were all described due to strabismus. Difficulties with balance were reported to occur due to strabismus causing problems concentrating and focussing.

> *"I used to suffer with balance loss, because both my eyes... didn't focus together, I was a bit like falling over or a bit clumsy"*
> *Participant 004, Female, 18-35 years.*

6.4.2 Task Performance - Postoperatively

Driving

Postoperatively driving was reported as improved by seven participants and this was associated with improved vision. Vision was described as clearer, more in focus, and better when looking around. Looking around during driving was described as using the strabismic eye more postoperatively to achieve greater peripheral vision. Having greater peripheral vision postoperatively was reported to make driving easier and lead to increased confidence when driving, due to needing to make less head movements to look to the side.

"*mainly when you're driving and things like that, it's better when you're looking round and things… you do realise the difference… when you're driving and… looking round corners or looking to the left and using that eye more, that's where you sort of notice it a bit better"*
Participant 006, Female, 36+ years

Work

Work tasks and overall work ability were described as improved postoperatively. Examples of specific work tasks such as operating tills and machines were reported as improved, leading to making less mistakes and being better at their job. Improved ability at work was also described as working for longer, being able to work harder, needing to take less rest breaks at work and having improved concentration whilst at work.

"I was only working a day or so in the hairdressers and now I'm doing like 3 days now. I don't want to do any more than that, but I'm working 3 long days now. I probably wouldn't have done that before"
Participant 003, Female, 36+ years.

Some participants described their work ability as better postoperatively because their vision was improved, being able to see better and due to improved focussing. Whereas other participants described improvements in their work ability as occurring because their eye was in a straighter position and it was drifting less. They felt more confident interacting with other people in a work capacity and more confident to take on more work.

"I was always making mistakes… but now, I don't, it's very rare I have any mistakes at all…. (in my job) you have to work fast"
Participant 010, Male, 18-35 years.

Improvements in ability or performance at work also involved descriptions of how using computers and screen-based devices had improved postoperatively. Being able to focus on computer-based tasks, looking at brighter screens, and looking at screens closer or without enlarging images were all described as ways in which using screen devices were improved postoperatively. Work based abilities using computers were described as improved, for example being able to work for longer on the computer and doing more and harder work.

"obviously my left eye was working harder than my right eye, so I couldn't really look at, I couldn't focus on the computer, because it was too bright, I had to turn it down to the lowest brightness, but now I can have it on... full brightness and look at… what I'm doing, I can do what I need to do, because I had to complete a lot of emails and stuff like that, I can focus on my emails a lot better than I used to"
Participant 004, Female, 18-35 years.

Near tasks

Postoperatively near tasks were reported as improved. Participants gave examples of near tasks and practical daily activities where they had noticed differences in their ability postoperatively, including artwork and drawing, woodwork and carpentry, putting keys into door locks, wiring in dark conditions, home DIY tasks and cooking. Participants described being able to focus more or better on near tasks, being able to perform near tasks for longer without taking a break and without needing to close one eye, being more accurate and being able to perform more tasks without getting frustrated. Improved practical daily activities were sometimes described as associated with using their strabismic eye, using both eyes for the task and having improved judgement of position.

> *"now my judgement is a lot better…. if I was trying to thread a needle… I would miss it totally, but now I'm looking through my bad eye I can, I'm getting the judgement right on the hole and getting it through, and not have to like use just one eye to do it, I'm using both"*
> *Participant 011, Female, 18-35 years.*

Postoperatively reading was reported to be improved. Being able to read for longer, reading for longer without taking breaks, being able to read more, and straining the eyes less to read, particularly when reading small print, were all given as examples of 'better' reading after surgery.

> *"I'll be able to like to sit and read for longer without having to take a break and take my eyes off the page"*
> *Participant 002, Female, 18-35 years.*

Balance and depth perception

Improvements in balance, less dizziness, bumping into things less and being less clumsy were described as occurring postoperatively. Improvements in balance were associated with being able to focus with the strabismic eye, feeling like the eyes were working together and having more confidence in their balance postoperatively. Improved dizziness was reported to occur with the strabismic eye drifting less postoperatively.

> *"I used to suffer with balance loss, because both my eyes... didn't focus together, I was a bit like falling over or a bit clumsy, but now they work together I don't think… it's that bad, because my eyes are focussing together on something"*
> *Participant 004, Female, 18-35 years.*

Other tasks were described by participants related to depth perception and the feeling of being able to see depth or judge the position of objects in space. Whilst participants had no measurable BSV, some gave examples of postoperatively feeling they were better able to see depth or judge the position of things, including picking up everyday objects, threading needles or beads, judging steps, judging the heights of obstacles and seeing three-dimensional (3D) computer games.

"(before the surgery) you have to do things slowly, a lot slower, whereas when your eyes are aligned you can just judge it straight away… like now I'm judging it the same really, I'm not tripping over, I know where it is and… I'm not thinking is it a little bit higher than what I'm seeing, or is it lower… I used to do that quite a lot…. when your eye's doing that (misaligned) you do question things because you know that your eye's not right" Participant 011, Female, 18-35 years.

6.4.3 Task Performance - Summary

Participants described a range of task performance difficulties preoperatively. Specific tasks such as driving, using screens such as computers, reading, near work and balance were all described as difficult with strabismus preoperatively. Work ability and performance was also reported to be difficult and frustrating due to strabismus. Postoperatively participants reported they felt better able to perform tasks such as driving, work related tasks, tasks at near including reading. They also reported improved balance and depth perception.

6.5 Theme: Physical Symptoms

Some of the discussion about living with strabismus preoperatively related to the physical symptoms' participants believed were caused by their strabismus. Physical symptoms included pain, discomfort and headaches. Participants described the consequences of these physical symptoms including the need to take pain relief medication and taking time off work. Participants then described ways in which they felt their physical symptoms improved or worsened postoperatively in the following sections.

6.5.1 Physical Symptoms - Preoperatively

Participants described their strabismus as hurting and causing discomfort, eye strain, a feeling of the eye turning, a feeling of tightness, dizziness, watering eyes, eye ache and pain, and headache. Physical symptoms were reported as worse when tired, at work or when unwell. Participants associated their physical symptoms with feeling the presence of the strabismus, the strabismic eye pulling or straining, the strabismic eye feeling tired, exerting effort to try and control their eye position, feeling confused and the feeling of one eye working much harder than the other. The pain caused by strabismus was associated with not being able to focus properly, causing headache, bad temper, stress and depression. The consequences of discomfort and pain caused by strabismus was for participants to take pain relief medication, close one eye, rub their eyes, take breaks from tasks and from work to rest their eyes, and to take time off work.

"I had to take… days off because of it, because of it feeling uncomfortable... I didn't really get a lot of headaches, I did sometimes, but it was just more that the pain, the pulling pain, it was excruciating some days"

Participant 009, Male, 36+ years.

6.5.2 Physical Symptoms - Postoperatively

After the initial postoperative healing period participants described mostly improved physical symptoms, which they attributed to having their eye in a more central position and the eye no longer turning or pulling.

Seven participants described postoperatively experiencing less discomfort, the strabismic eye feeling less tired, the eyes hurting less, less eye pain and less eye strain. In some cases, surgery led to a resolution of these physical symptoms with participants describing the eyes feeling comfortable and no longer experiencing any eye pain or eye strain postoperatively. Experiencing less pulling of the eyes postoperatively was attributed to the eye muscle being tightened and scar tissue being removed. Surgery was described as relieving or lessening the tightness or the tightening feeling caused by strabismus. The strabismic eye feeling less tired was associated with being able to see things better. Needing to close one eye less, or no longer closing one eye, was reported, as well as the eyes feeling more relaxed and more natural. Experiencing less eye pain and the eyes feeling more relaxed was described as leading to being able to concentrate more. The consequence of the eyes feeling more natural was reported as being able to look at things properly and read for longer.

> *"it's a lot better, no more headaches, no more squinting, because of my eye… I was holding my eye to try and get the pain away… now there's no more pain any more or anything"*
> *Participant 011, Female, 18-35 years.*

Headaches were mostly reported as improved postoperatively (n=6) by either reducing in frequency or by stopping and no longer experiencing headaches. If a headache did occur postoperatively, it was reported to occur when tired.

> *"if I went out anywhere or did anything… even riding a bike and looking where you're going, I used to, it was a strange sensation, but I would end up with a headache, yeah I don't get as many now. I do get them sometimes, but not as many, mostly when I'm tired"*
> *Participant 003, Female, 36+ years.*

Participants reported taking less breaks (n=7) due to improved physical symptoms, for example the eyes hurting less and less eye pain. Some even reported no longer needing to take any additional breaks whilst at work, other than regular work breaks and no longer needing to take time off work (on sick leave) due to eye pain. Some participants described no longer being able to feel the presence of the strabismus or their eye turning postoperatively. Less eye watering and improved

headaches were also reported and this varied from experiencing fewer headaches to a complete resolution of headaches.

> *"I take them (breaks), but... I don't need them as often as I used to… yeah, it's a lot better now"*
> *Participant 006, Female, 36+ years.*

Taking additional rest breaks whilst at work was also described as a result of being aware that other people are noticing the strabismus, rather than specifically due to physical symptoms of pain, discomfort or headache. In this situation the participant reported that postoperatively they took fewer additional rest breaks because fewer people at work noticed their strabismus.

Experiencing fewer physical symptoms was often reported during specific activities such as driving, near tasks, work related tasks and reading. Participants associated experiencing fewer physical symptoms with improved task performance, giving examples of being able to read for longer, work for longer, work without needing to take additional rest breaks. They also reported being aware that physical symptoms were lessened during times when symptoms were previously worsened, such as when tired, at work and when unwell.

> *"if I read for quite a while, you could feel it pulling… you could feel something was dragging it to the side, that's what it felt like…. it still pulls a little bit, but nowhere near as bad as it did, and I can read a bit longer now"*
> *Participant 009, Male, 36+ years*

Two participants described some of their physical symptoms worsening, as well as some improving. One participant described increased eye sensitivity and eye watering, which they felt were minor side effects of having surgery.

> *"my eyes water a lot since my surgery… they're a lot more sensitive now than they were before my surgery"*
> *Participant 004, Female, 18-35 years*

Another participant described experiencing worsening headaches, which they reported were a significant side effect postoperatively.

> *"the headaches have got worse though, so that's a negative… the headaches are definitely worse, but I don't know why that is"*
> *Participant 012, Male, 18-35 years*

6.5.3 Physical Symptoms - Summary

Participants described a range of physical symptoms attributed to strabismus preoperatively, including pain, discomfort, eye strain and headaches. These physical symptoms were described as

the reason for taking breaks from work or taking time off work. Postoperatively most participants felt their physical symptoms improved or resolved, for example, participants described less eye strain, less pain and fewer headaches. Having improved physical symptoms was reported to lead to having less time off work and taking fewer additional rest breaks.

6.6 Theme: Confidence and Emotions

In this theme, participants described the impact of having strabismus on their emotions, emotional well-being and confidence preoperatively. Confidence was described in the context of self-confidence, self-perception, confidence when interacting with other people and confidence in social situations. Participants also described ways in which they felt their confidence had changed postoperatively. Improved self-confidence was described in addition to different ways in which participants felt more confident in their eyes, their vision and their abilities. These issues are described in the following sections.

6.6.1 Confidence and emotions - Preoperatively

Strabismus was described as having an extremely negative impact on participant's self-confidence affecting many different aspects of participants lives, including education, work life, social life, personal life and their relationships with other people. Having a noticeable strabismus was described as causing low self-confidence and was reported to lead to feelings of anxiety and depression, avoiding people and social situations. Strabismus was described as holding participants back and causing them to miss out on or decline opportunities.

> *"I was trying to avoid everything"*
> *Participant 011*, *Female, 18-35 years.*

Having strabismus was described as distracting, unpleasant and not nice when other people noticed the strabismus or commented on it, being bullied because of strabismus was also reported. Participants reported suffering from strabismus and how they would spend a lot of their time thinking about their strabismus, how it affected them and their lives. Different techniques were used to try and hide their strabismus from others or avoid social interactions.

> *"other people could tell because they'd commented on it... which was very upsetting and very hurtful"*
> *Participant 004, Female, 18-35 years.*

6.6.2 Confidence and emotions - Postoperatively

Having straighter eyes postoperatively was described as making a big impact on all of the participants' lives, making them feel better and leading them to no longer worrying about their eyes and their eye position. Participants (n=12) described themselves as happier, feeling better in

themselves, having improved confidence and no longer suffering with stress, anxiety or worry postoperatively. Feeling better in themselves and having more self-confidence was linked to overall improvements in participants' lives, feeling more able to do things and socialise more.

> *"it's miles better, basically everything about it has just improved my life, I feel a lot more comfortable about things, more relaxed, rather than really tense all the time"*
> *Participant 009, Male, 36+ years.*

Postoperatively participants (n=13) reported being able to interact with and communicate with people much more. They described being able to go out in public, look at people, make eye contact and have face-to-face communication. In a social setting participants reported being able to meet new people and feeling more relaxed and confident when meeting new people. They reported that with straighter eyes they were able to speak to people and socialise more and they were no longer avoiding social situations or communication with people. Feeling able to face people, feeling less embarrassed and more comfortable in social situations as well as feeling able to fit in socially were all described as consequences of having straighter eyes.

> *"I found it a lot easier to talk to people in person and communicate with people"*
> *Participant 002, Female, 18-35 years.*

In a work setting participants reported an improved ability to interact with and communicate with people led to them being more confident at work and being given more 'face-to-face' work opportunities. Participants reported being perceived as more friendly and less rude at work because they were able to communicate with others better. Having increased confidence at work and being able to communicate more with others led to participants describing themselves as being better at their jobs postoperatively.

Having straighter eyes led to participants feeling a lot more confident and less stressed. Participants described no longer receiving comments about their strabismus or being treated differently. The ability to look at another person and communicate with them were described as important factors in postoperatively having the self-confidence to put themselves forward for opportunities such as applying for a new job, going for a job interview, applying for a university course and starting a new career.

> *"I'm more willing to go for an interview, whereas before… I'd stay in the same job… because I'm getting older I want a better career and that's holding me back (referring to strabismus), because I just knew… if they'd see my eye I'm thinking they're not going to want to employ somebody like that… so, (postoperatively) I've recently applied for uni"*
> *Participant 011, Female, 18-35 years.*

Having greater confidence in their ability to perform a task or activity was also reported by two participants postoperatively, in addition to greater self-confidence and social confidence. Participants reported that more confidence in their abilities gave them greater confidence to try new activities. Being more confident in their ability to drive and their abilities at work were both reported as examples of feeling more confident in their abilities.

> *"well I'm certainly working as much as I was... probably more after the operation, you know, I'm more confident at taking more jobs on"*
> *Participant 007, Male, 36+ years.*

Postoperatively three participants also described having greater confidence in their eyes and their vision, which was typically described when performing a task. Having greater confidence in their eyes and vision was reported as a significant factor in feeling able to try new activities. Examples of having greater confidence in their eyes and vision postoperatively included having the confidence that their eyes will see a car when driving; that their eyes see clearer postoperatively; and that they will move to see an object without causing pain or headache. Having confidence in what they see was also linked to no longer needing to double check everything and giving them additional confidence that they could drive as part of a job.

> *"I wouldn't have applied for that kind of job, anything to do with driving, but now I would, I'm confident to think that my eyes are good enough to do that kind of thing"*
> *Participant 008, Male, 36+ years.*

6.6.3 Confidence and emotions - Summary

All participants perceived that strabismus had an extremely negative impact on their confidence and emotions preoperatively, as well as a number of ways in which they perceived their confidence and emotions improved postoperatively. All reported the positive impact of having improved eye alignment on their interactions with other people, in both social and professional settings. Postoperatively improved self-confidence was common. Improved confidence in vision, the eyes and in abilities were also reported. Participants reported improved self-confidence and emotions postoperatively had a large positive impact on their lives.

6.7 Chapter Summary

This chapter has described the semi-structured interview and thematic analysis findings. Postoperatively, adults who have undergone strabismus surgery for psychosocial reasons, can report a range of improvements in their vision, task performance, physical symptoms and confidence and emotions. Worsening of symptoms or the development of new 'negative' symptoms postoperatively was much less common than participants reporting positive outcomes after

surgery. Some of the changes reported in this chapter may be measurable using standard clinical measures and existing QoL, HRQoL or PROM questionnaires. However, it is likely that some of the changes described may not measurable with standard clinical tests. These considerations were considered when selecting the measures for the quantitative phase two. The process of selecting the phase two measures will be described in chapter 7.

Chapter 7 – Selection of the Quantitative Measures

This chapter describes how the findings from the qualitative phase (chapter 6) and the literature evidence (chapters 2 and 3) were used to inform the design of the quantitative phase two. The process of selecting the quantitative measures of visual function, task performance, physical symptoms, and confidence and emotions to be performed pre and postoperatively is described. The further refinement of the quantitative measures was informed by a review of the literature and discussion with supervisors and the advisory group. The literature evidence supporting the refinement of the measures is presented in this chapter. In the following chapter (chapter 8) the quantitative methods will be described.

7.1 Five stage method to select the quantitative measures

The findings from the semi-structured interviews and thematic analysis (chapter 6) were considered and combined with literature evidence of strabismus and strabismus surgery outcomes (chapters 2 and 3). The aim was to include quantitative measures that could measure the changes reported by participants during the qualitative phase one, whilst being mindful of testing burden for participants recruited to the quantitative phase two. An overview of the five-stage method used to select and refine the measures for the quantitative stage is shown in Figure 7-1. Regular discussions with academic supervisors, clinical supervisors and the study advisory group took place during the five stages to inform the process and aid in the selection, elimination and refinement of the possible measures.

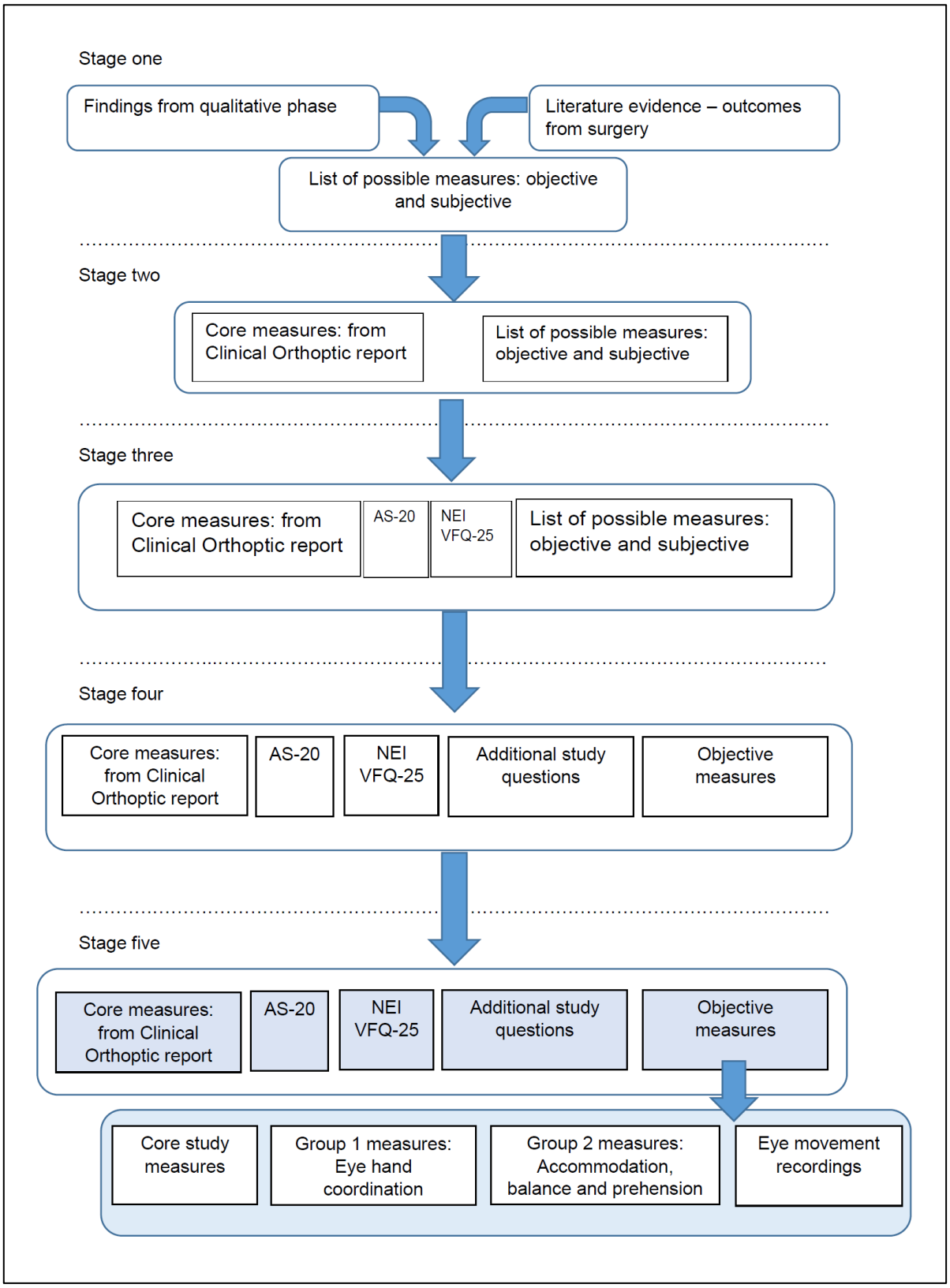


Figure 7-1. A flow chart illustrating the five-stage selection process of the measures for the quantitative phase two.

Stage one

During stage one, the list of measures suggested by the findings from qualitative interviews (chapter 6) was combined with the measures and outcomes reported in the literature (chapters 2 and 3). This combined list of potential outcome measures is presented in Table I-1 (Appendix I.1). This resulted in a combined list of measures under consideration, including objective measures or clinical tests, subjective self-reports, or instances where both objective and subjective measures were suitable. The time to complete all of these measures would have been too long to be practical during the study. To refine this list further the combined list of measures (objective, subjective or both) was prioritised by ranking each item in order of the frequency it had been described or mentioned during the semi-structured interviews (described in Table I-2, Appendix I.2).

Stage two

During stage two, standard 'core' clinical orthoptic measures were separated from the list of possible study measures. VA, cover test (CT), BSV, ocular movements (OM) and PCT were considered suitable for extraction from the clinical orthoptic report, rather than being repeated during the study visit (Table I-3, Appendix I.3). The list of possible study measures was refined further by removing measures considered more suited to a subjective self-report than an objective measurement (Table I-4, Appendix I.4). Further discussion with supervisors and the advisory group refined the selection of the objective measures further, by removing tests considered time consuming, overlapping with other measures or more suited to a subjective self-report (Table I-5, Appendix I.4).

Stage three

During stage three, possible PROMs were reviewed (Appendix I.5). The AS-20 (Hatt et al., 2009b) and VFQ-25 (Mangione et al., 2001; Mangione et al., 1998) were selected for inclusion in the study. Both were suitable for pre and postoperative use in adults with strabismus, without specifically measuring the effect of amblyopia. The AS-20 was reported to be more sensitive to change in HRQoL compared to the VFQ-25 in adults with strabismus (Hatt et al., 2009a). The VFQ-25 was selected to measure the impact of visual disability and symptoms on health and daily functioning, although the VFQ-25 has been reported as a vision related QoL measure (Fieß et al., 2020), as the results have an implied effect on QoL.

Stage four

During stage four, the coverage of the AS-20 and the VFQ-25 questionnaires were reviewed and compared to the qualitative interview findings (Table I-6, Appendix I.6). Additional study questions were developed to enable wider coverage of the themes and topics arising from the semi-

structured interviews whilst avoiding duplication of the topics already covered by the AS-20 and the VFQ-25 (Table I-7, Appendix I.6). A long list of possible questions was refined into thirty-three additional study questions with input from supervisors and the advisory group (Table I-8, Appendix I.6). Questions covered the themes of vision (Q1-Q11), task performance (Q12-Q17), physical symptoms (Q18-23) and confidence and emotions (Q24-Q33) (Appendix I.7). Responses to the questions followed a similar format to the AS-20, with five possible responses 'none of the time', 'a little of the time', 'some of the time', 'most of the time' and 'all of the time'. Questions were worded in such a way to avoid double negatives within the question and responses, and responses being ordered and predictable. For example, 'none of the time' did not mean better or worse for every question. The development of an 'item bank' of possible patient reported outcome measurement questions has been reported by Pesudovs (2010) however, questions specifically relating to strabismus and psychosocial symptoms have not been included in the item bank. Therefore, for this study, questions were developed to capture the qualitative interview findings whilst avoiding overlap with the AS-20 and VFQ-25 (Table I-7, Appendix 1.6).

Stage five

During stage five, the list of possible objective study measures was refined and divided into core study measures, which included measures of vision and task performance. The remaining objective study measures were divided into groups (one and two) and an optional additional eye movement recording (EMR). Group one measures included eye hand coordination. Group two measures included accommodation, prehension and balance. The measures included at each stage of the quantitative phase are shown in Figure 7-2.

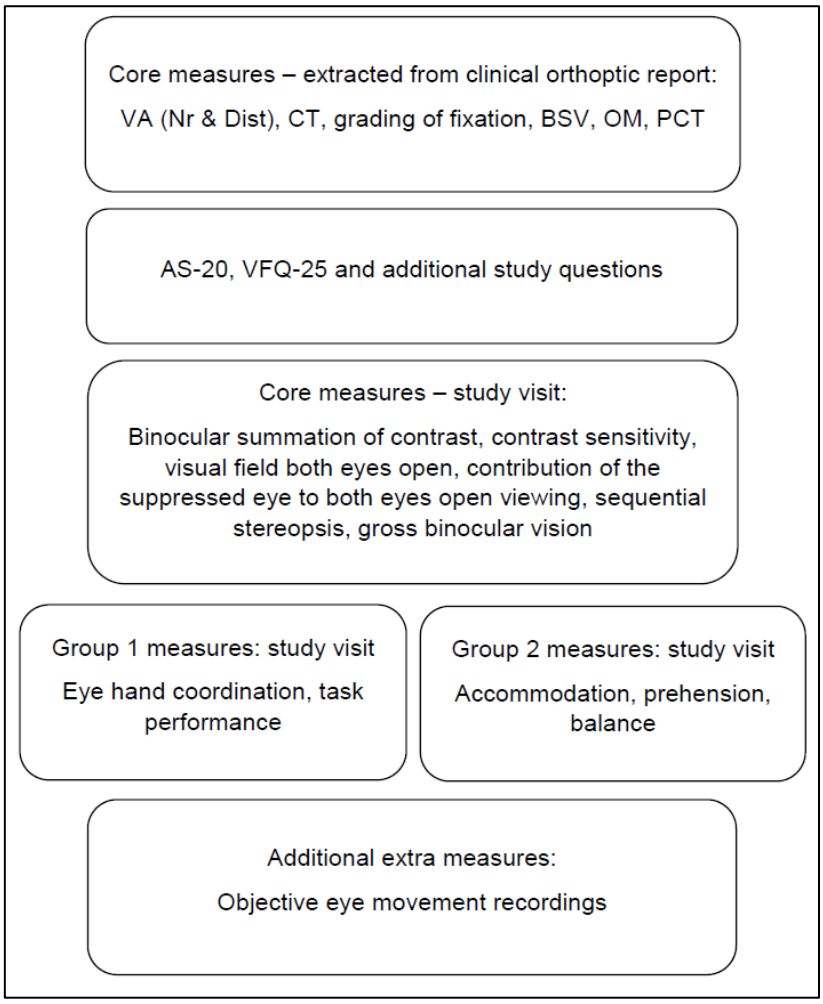


Figure 7-2. The overall structure of quantitative phase two, including the measures used at each stage.

7.2 Evidence informing the selection of the quantitative measures

The following sections will describe the literature evidence that informed the selection and refinement of the quantitative measures. The quantitative measures are presented in the following sections, core measures extracted from the clinical orthoptic report (section 7.3). PROMs (section 7.4). Core measures to be completed during the study visit (section 7.5). Further measures to be completed during the study visit in either group one (section 7.6) or group two (section 7.7). Optional EMR to be completed during an extra study visit (section 7.8).

7.3 Core measures – extracted from the clinical orthoptic report

Extracting information from the clinical orthoptic report was selected to reduce patient testing burden and avoid duplication of tests as part of a clinical and experimental study visit. The measures typically performed as part of an orthoptic and ophthalmological assessment of an adult

patient with strabismus and psychosocial symptoms included VA, CT, investigation of BSV, OM and PCT.

7.3.1 Visual acuity (VA)

LogMAR VA charts have increased VA measurement accuracy with greater standardisation of optotype size, spacing and progression down the chart. Scoring each optotype allows part lines to be scored more accurately, as each letter correctly seen is accounted for. The gold standard ETDRS VA chart was originally developed for research (Early Treatment Diabetic Retinopathy Study Research Group, 1991) but is increasingly used clinically. Near and distance ETDRS chart versions are available. Typically, high contrast VA charts are used in clinical practice, although low contrast VA charts are also available. Measuring low contrast VA has been suggested to be a better measure of visual function than contrast sensitivity in certain clinical conditions, for example multiple sclerosis (Balcer et al., 2017). High contrast VA charts however, have better test-retest consistency than low contrast charts (Carkeet & Bailey, 2017).

When performing a VA test, participants are typically asked to read out as many letters as they can see on the chart. For a VA test with five optotypes per line, the termination rule of stopping when four or more mistakes are made on a line has been recommended to ensure the most accurate testing and recording methods (Carkeet & Bailey, 2017). Each letter correctly seen on the final line of the test is then scored, with each letter on an ETDRS chart equalling 0.02 logMAR and each complete line read equalling 0.10 logMAR. Distance VA is typically measured and recorded monocularly, with participants wearing their optimum refractive correction. At times, VA with both eyes open may be recorded, for example at near when the result is considered more important as an overall measure of visual performance at that distance. Near VA is also considered relevant to the performance of tasks at near (Sheedy et al., 1991). The high contrast logMAR ETDRS VA chart was selected for phase two.

7.3.2 Cover test (CT) and fixation grading

The CT is the gold standard method to identify manifest and latent strabismus. Typically performed at near and distance, the CT includes ‘cover-uncover’ and ‘alternate-cover’ components. During the CT the fixation behaviour of the deviating eye is observed. Fixation preference relates in part to VA, or level of amblyopia, in the deviating eye (Cotter et al., 2009; Şener et al., 2002; Wright et al., 1986). Fixation behaviour typically includes a description how well the deviating eye is able to take up and hold fixation during the CT. Wright et al (1986) used ‘alternates’, ‘holds well’, ‘holds briefly’, ‘holds momentarily’ or ‘will not hold’ to categorise fixation preference. Others have used five similar descriptions to grade fixation, such as free alternation (no fixation preference) (4), holds fixation well (holding fixation through a blink) (3), holds fixation briefly (holds fixation for 2-3 seconds) (2), holds with difficulty (can take up fixation but unable to hold) (1), and no fixation (unable to hold

fixation) (0) (Şener et al., 2002). The duration that fixation was maintained can be additionally be used in the grading. Cotter et al. (2009) used grade A 'spontaneous alternation', grade B 'holds well' where fixation was for at least 3 seconds, occurred during smooth pursuit and was held through a blink, grade C 'holds momentarily' where fixation was held for 1-3 seconds, or grade D 'does not hold' where fixation was held for less than one second.

In phase two fixation preference was observed at near, with refractive correction, if worn. A descriptive scale was developed, as none of the published scales were deemed satisfactory. The following criteria were used:

1. Unable to take up fixation with deviating eye (unable to take up fixation)
2. Able to take up fixation with deviating eye, but unable to hold fixation (momentary fixation)
3. Able to take up fixation with deviating eye and able to hold fixation for up to 5 seconds (brief fixation)
4. Able to take up fixation with deviating eye and able to hold fixation for longer than 5 seconds, but unable to alternate fixation (holds fixation)
5. Can alternate fixation if specifically asked or encouraged to do so (voluntary alternation)
6. Freely alternates without specific effort to alternate fixation (free alternation)
7. Heterophoria

7.3.3 Binocular single vision (BSV)

BSV tests are performed during an orthoptic assessment, typically with the patient wearing their optimum refractive correction. Sensory fusion, motor fusion then stereopsis would be investigated at near and distance. In this study, preoperatively, Bagolini glasses were used to demonstrate participants had suppression at near and distance. Potential BSV was investigated in free space with Bagolini glasses and loose prisms or using the Synoptophore. Evidence stratifying the postoperative outcome of strabismus surgery by the quality of the BSV predicted preoperatively or measured postoperatively is lacking. Therefore, any evidence of potential BSV was considered as the surgery having a functional aim, even in cases where potential sensory fusion, or gross BSV, was demonstrated only. The postoperative diplopia test was performed in free space with prisms at near and distance (from 20PD undercorrected to 20PD overcorrected). Only those with suppression, no potential BSV, no diplopia and no risk of postoperative diplopia met the inclusion criteria. Postoperatively, suppression was not assumed, as diplopia (Wang et al., 2019) or unexpected BSV can occur (Ball et al., 1993; Liebermann et al., 2014). Bagolini Glasses were used to establish the presence of sensory fusion, diplopia or suppression postoperatively. If sensory fusion was present, at either near or distance, the prism fusion range (PFR) was tested in combination with Bagolini glasses and stereopsis was tested using the FNS, Wirt and the FD2.

7.3.4 Ocular movements (OM)

A standard orthoptic examination of OM uses a pen torch for fixation, whilst the examiner moves the light slowly and smoothly from primary position into each of the eight positions of gaze. In each position, ductions and versions are assessed by observation and graded using a 9-point scale. 0 denotes no limitation, restriction, underaction or overaction. Overactions are graded using +1 to +4. Underactions, restrictions or limitations are graded using -1 to -4, with -1 equating to 75% of movement (Ansons & Davis, 2014; British and Irish Orthoptic Society, 2015). In this study, the 0 to +/-4 grading scale was used to grade OM in each position of gaze.

7.3.5 Prism cover test (PCT)

The PCT measures the size of a deviation at near and distance. Typically, the alternate prism cover test (APCT) is used to measure the deviation by using maximum dissociation and recording the maximum amount of prism prior to reversal of the deviation (Ansons & Davis, 2014). However, the simultaneous prism cover test (SPCT) can be used to measure only the manifest component and is preferred by some (Serafino et al., 2019). Sometimes both the SPCT and APCT are recorded if an additional latent component is observed during the CT (Deacon & Gibson, 2001). The Synoptophore is an alternative clinical method for measurement of the deviation. Whilst the Synoptophore is suitable for measurements of the distance deviation in primary, secondary and tertiary positions of gaze, it has proximal cues that can affect the deviation measured (Ansons & Davis, 2014).

Variability in PCT measurements can be induced by holding prisms incorrectly, such as not in the position that they have been calibrated for, and by altering other factors such as examiner position (Johnson et al., 2004). Standardisation of PCT technique is therefore important. Variability in APCT measurements of horizontal deviations less than 10PD, when measured by different clinicians, may be within inter-examiner variability (de Jongh et al., 2014). Others have suggested greater inter-examiner variability occurs for larger deviations. For eso deviations larger than 20PD, a difference of 12PD was considered real change, but for smaller eso deviations (10-20PD) a difference of 6PD was considered real change (Holmes et al., 2008). de Jongh et al. (2014) did not find the size of the horizontal deviation to affect the inter-examiner variability. Instead they found that variability between examiners differed between different examiner pairings, highlighting that inter-examiner results between two examiners performing the PCT cannot necessarily be generalised to all examiners.

In addition to primary position deviations, the PCT can be used to measure deviations in secondary and tertiary positions of gaze (Ansons & Davis, 2014). Care must be taken to standardise testing conditions to minimise error induced by prism placement or position of the head. Refractive

correction is typically worn for PCT measurements to ensure the patient is able to see the required fixation target with each eye and make the appropriate amount of accommodation at each fixation distance. Measurements of the deviation in secondary and tertiary positions of gaze can be more problematic when wearing glasses due to a prismatic effect and frame size. For this reason, measurements of the deviation in secondary and tertiary positions of gaze may be performed without glasses, providing the patient is able to fixate on the required fixation target adequately. Clinical judgement is therefore required to select the most appropriate conditions for PCT measurements to ensure accuracy and repeatability.

The APCT was used in phase two to measure the maximum size of the deviation. Measurements were performed by one examiner where possible using a standardised technique. Primary position measurements were performed with refractive correction. APCT measurements in secondary positions of gaze were performed without refractive correction if possible, but with refractive correction if the participant was unable to accurately fixate on the distance target without refractive correction.

7.4 Patient reported outcome measures (PROMs)

The negative effects of strabismus on HRQoL (Chang et al., 2015; Ritchie et al., 2013; Satterfield et al., 1993) have been discussed in chapter 2. The most common reason for undergoing strabismus surgery when there is no predicted visual benefit preoperatively, is for psychosocial reasons to improve HRQoL (Alpak et al., 2014; Jackson et al., 2006; Nelson et al., 2008). The outcomes from surgery for psychosocial reasons have been discussed in chapter 3. The AS-20, VFQ-25 (Appendix I.5) and the additional study questions (Appendix I.7) were selected as PROMs for phase two.

The original AS-20 questionnaire and scoring method was used to measure HRQoL (Hatt et al., 2009b) (Appendix J). The VFQ-25 instructions for administering and scoring the questionnaire was used (Appendix K). The additional study questions were administered using a standardised set of instructions and scoring method (Appendix L).

7.5 Core measures – study visit

The study design was selected to ensure that all study participants performed a select ‘core’ set of vision and task performance measures.

7.5.1 Binocular summation of contrast

Binocular summation is considered a measure of binocular function, but not a measure of BSV. Binocular summation is the improved performance of a visual task binocularly, or both eyes open,

compared to monocularly, using the better eye. Binocular inhibition, is the improved performance of a visual task monocularly, using the better eye, compared to binocularly. To measure binocular summation a visual threshold task is performed with both eyes open (binocularly) and compared to the performance of the task monocularly, with the better eye (with the worse eye covered). Traditionally binocular summation has been investigated in a laboratory setting, using psychophysical experimental techniques to accurately measure binocular and monocular thresholds for different visual detection tasks. Computer displays of contrast threshold tasks can be used (Pineles et al., 2014; Thompson et al., 2011) sometimes with the addition of shutter goggles to manipulate the contrast of the stimuli seen by each eye (Baker et al., 2007; Vedamurthy et al., 2007). Electrodiagnostic testing of visual evoked potentials has also been used to measure detection thresholds of different contrast stimuli (Pineles et al., 2013; Plainis et al., 2011). Clinical tests have been used to measure binocular summation. High contrast VA (Plainis et al., 2011), contrast sensitivity (Pelli Robson) charts (Pineles et al., 2014; Pineles et al., 2013) and low contrast VA charts have been used (Pineles et al., 2014).

Binocular summation of contrast can be measured using ETDRS high and low contrast (2.5% and 1.25%) VA charts. Threshold VA has been measured binocularly (both eyes open) and monocularly with the better seeing eye, in a dimly lit room, although exact room luminance or lighting conditions were not specified (Chang et al., 2017; Pineles et al., 2015; Pineles et al., 2014; Pineles et al., 2013; Tandon et al., 2014). The results are typically recorded as the number of letters correctly seen using the same VA testing and scoring reported by the Early Treatment Diabetic Retinopathy Study Research Group (1991). The minimum number of letters is scored as 0 and the maximum number of letters can be scored as 70 (Kattan et al., 2016; Pineles et al., 2013; Tandon et al., 2014) or 100 (Chang et al., 2017; Pineles et al., 2014). In earlier work a binocular summation score was calculated by dividing the binocular score by the better eye score (binocular score / better eye score) (Pineles et al., 2013). More recently the binocular summation score was calculate as the difference between the binocular letter score and the better eye letter score (binocular letter score – letter score better eye) (Chang et al., 2017; Kattan et al., 2016; Pineles et al., 2014; Tandon et al., 2014).

Binocular summation scores are typically presented and analysed using the calculated score. Scores can be further categorised into binocular summation, binocular inhibition or indeterminate VA scores, with a difference of 5 letters or more suggested for categorisation. For example, if binocular VA was 5 letters (or more) better than monocular (better eye) VA, binocular summation was present. If binocular VA was 5 letters (or more) worse than monocular (better eye) VA, binocular inhibition was present. In there was a difference of less than 5 letters in binocular and better eye VA, indeterminate VA was present (Chang et al., 2017; Kattan et al., 2016; Pineles et al., 2014; Tandon et al., 2014). A combination of values can be presented in the analysis with

calculated binocular summation scores, the percentage of participants falling into each category (binocular summation, indeterminate VA or binocular inhibition), and the percentage of participants that demonstrate a change in binocular summation score of 5 letters of more following strabismus surgery all being considered relevant (Pineles et al., 2015).

The low contrast VA charts used for measurements of binocular summation are typically the Sloan letter charts at 2.5% and 1.25% contrast levels. These can be presented on a solid surface or presented in an illuminated cabinet. There are limited chart versions making it impossible to randomise letters or change chart types between monocular testing of each eye and binocular testing. Electronic charts presented on liquid crystal display (LCD) screens have addressed this issue by allowing presentations of multiple different chart types and randomisation of letters between presentations to prevent a learning effect. However, evidence of the standardisation and the comparability between the contrast levels of optotypes presented in the different chart formats is lacking.

In phase two, the Thomson Test Chart Xpert 3Di was used to present the high (100%) and low contrast (10% and 5%) VA charts at 4m to allow randomisation of the letters before every chart presentation. British standard letters were selected, rather than Sloan letters, as this allowed 5 letters to be presented at 4m for all VA levels. The wider spacing between the Sloan letters allowed only 4 letters to be presented at 1.0 logMAR at 4m. The British standard letters included D E F H N P R U V Z, which had 6 letters in common with the Sloan letters (D K O H N S R C V Z), yet it avoided the O and C which may be considered interchangeable in some studies as they can easily be confused (Bailey & Lovie–Kitchin, 2013; Elliott et al., 1991). Binocular summation, or inhibition, scores were calculated from the letter scores for each level of contrast (binocular letter score – better eye monocular letter score). The maximum number of letters was 100. Additional analysis included the percentage of patients in the categories of binocular summation, indeterminate, and binocular inhibition.

7.5.2 Contrast sensitivity

Contrast sensitivity testing relies on participant identification of optotypes of reducing contrast until their threshold for correctly identifying optotypes at low contrast is detected. Laboratory based measures of contrast may use gratings and measurement of the contrast sensitivity function curve at multiple different spatial frequencies (Pelli et al., 1988). Letter based contrast sensitivity charts have been developed specifically for measuring contrast in clinical conditions, with the aim of detecting reduced contrast sensitivity. The Pelli Robson chart was developed for use at 3m but may be used at 1m in patients with low vision. Three letters at each level of contrast appear on the chart. Two of the three letters should be correctly identified for the triplet to be scored as seen (Pelli et al., 1988). The Pelli Robson contrast sensitivity test was printed and has two versions,

limiting the use of the test in a repeated measures study. Scoring each letter correctly seen, taking a mean result of both versions of the chart and considering the O and C as interchangeably correct have all been suggested as methods of increasing test reliability and sensitivity (Elliott et al., 1991).

The Mars letter contrast sensitivity test was also printed, with three chart types and 0.5m test distance (Arditi, 2004). It was developed to improve testing accuracy (Elliott et al., 1991) and test-retest reliability (Haymes et al., 2006) compared to the Pelli Robson. Whilst the results between the Mars and the Pelli Robson tests have been described as comparable (Arditi, 2004), the two tests may have different normative values and further normative data may be required for the Mars test (Haymes et al., 2006). Computerised test charts, such as the Thomson Test Chart Xpert 3Di, presented on LCD screens allow randomisation of letters and prevention of learning effects on repeated testing. The printed Pelli Robson and Mars contrast sensitivity tests were compared to Thomson chart presentation of the Pelli Robson (Thayaparan et al., 2007). The Mars and Pelli Robson printed charts were the most repeatable and had the greatest agreement. The Thomson chart results were slightly less repeatable with the properties of the LCD screen suggested as the reason for the result. Whilst the LCD screen was left on for 'several minutes' before testing to ensure the luminance had reached peak level (mean 304 cd/m^2) (Thayaparan et al., 2007), it is possible that the screen set up affected the results more than the computerised presentation of the chart. A small learning effect of repeated testing with the three Mars charts and two Pelli Robson charts has been shown by Haymes et al (2006).

For all letter based optotype testing, clear instructions for standardising the chart presentation (British Standard, 2003), the instructions given to the participants (Elliott et al., 1991; Haymes et al., 2006), the termination of the test (Carkeet & Bailey, 2017) and the scoring of the results (Early Treatment Diabetic Retinopathy Study Research Group, 1991; Elliott et al., 1991) are important (Bailey & Lovie–Kitchin, 2013) to improve testing and measurement standards. In this study, the Thomson Test Chart Xpert 2000 was selected to present the contrast sensitivity test in letter triplets to allow randomisation of the letters between each test presentation. At the threshold level reached, contrast sensitivity was scored as seen if two of the three optotypes were correctly identified. Whilst scoring each letter correctly seen as 0.05 log units has been reported to be more sensitive and reliable for the paper based Pelli Robson (Elliott et al., 1991), the Thomson Test Chart Xpert 2000 presentation of contrast sensitivity stimuli did not show stimuli at regular 0.15 intervals of reducing contrast to allow 0.05 per letter scoring.

7.5.3 Visual field both eyes open

A visual field test is typically performed uniocularly, however a visual field performed both eyes open can measure the visual field of the patient experienced during typical viewing conditions. Kinetic perimetry has been used to measure the binocular visual field, or a visual field plotted both

eyes open. The Goldmann perimeter has been used to document pre and postoperative peripheral visual fields using a I4e target in adults with esotropia (Kushner, 1994; Wortham & Greenwald, 1989). Wortham and Greenwald (1989) tested 1-2 months postoperatively (without glasses), whereas Kushner (1994) tested 5-8 weeks (mean 6 weeks) postoperatively with glasses.

Alternative methods of testing visual fields under binocular viewing conditions have been described. Joosse et al. (1997) and Joosse et al. (1999) reported a static perimetry method using two perimeters (at 90 degrees) simultaneously viewed through mirrors to test visual fields under binocular and monocular conditions. This method was used to investigate suppression scotomas in microtropia and small angle ET (Joosse et al., 1997) and intermittent and constant XT (Joosse et al., 1999). Kaban et al. (1997) reported measuring Goldmann visual fields uniocularly and combining them into a 'field of binocular vision' plot in healthy volunteers. The Humphrey visual field analyser has been used to perform static perimetry tasks with blue stimuli on a yellow background using yellow filters and different occluders to investigate suppression in adults with strabismic amblyopia (Barrett et al., 2012, 2013). Perimetry tasks were performed both eyes open (with minimal dissociation) and with the non-amblyopic eye (fellow eye) occluded. Suppression was said to be measured when performance was reduced during both eyes open viewing compared to monocular viewing with the amblyopic eye (Barrett et al., 2012, 2013). Whilst the Humphrey visual field analyser was used in the experimental investigation of the contribution of the apparently suppressed eye, it was not used to measure the size of the visual field of each participant.

In this study it was important to measure the size of the peripheral visual field. A kinetic visual field was plotted using a Goldmann perimeter with a I4e target using a similar technique to Kushner (1994) and Wortham and Greenwald (1989). Due to glasses frames potentially obstructing the view of the test and the variability amongst glasses wearers, glasses were not worn for visual field testing.

7.5.4 Gross binocular single vision (BSV) or stereopsis

Stereopsis is the ability to see depth from disparity and one of the cues used to make judgements about the depth of objects in daily life (Levi et al., 2015). Stereopsis is considered to be visually advantageous and requires BSV. It has been shown to lead to functional improvements in daily life, measured by motor skills tasks such as a peg board, threading and pouring (O'Connor et al., 2010a) as well as highly skilled motor task performance such as surgical procedures (Al-Saud et al., 2017; Bloch et al., 2015). Whilst clinical tests are used to investigate stereopsis or measure stereoacuity, some participants without clinically measurable stereopsis have been shown to perceive depth from 3D screen presentations, such as television or computer game devices. The

reasons for this perceived depth effect are unclear, but may include monocular clues, motion clues or features specific to the screen (Tidbury et al., 2014).

The investigation of BSV performed as part of the clinical orthoptic visit was discussed (section 7.3.3). It remained possible that some ability to perceive depth, or gross stereopsis, may be present, even though clinical evidence of BSV was not measurable. Clinical tests such as Bagolini glasses are considered minimally dissociative, yet they do not include any depth information. The Lang two-pencil test can be used as a qualitative stereopsis test detecting gross stereopsis (Ansons & Davis, 2014), although the results have been estimated to be equivalent to 3000-5000" of arc, depending on arm length and the interpupillary distance of the observer (LaRoche & Von Noorden, 1982). Unsurprisingly the Lang two-pencil test has been shown to be less accurate at predicting positive and negative BSV results, compared to Bagolini glasses and the TNO stereotest, despite attempts to modify the test to make it more challenging by occluding the tips of the pencils during the task (Mojon, 2009).

Improved performance with binocular viewing, compared to monocular viewing with the fixing eye only, may be considered as evidence that the deviating eye is contributing information to improve performance. Following this principle, a task involving the judgement of depth, or the position of objects in space, could be compared under binocular viewing and monocular viewing, with the strabismic eye covered. Whilst the disparity of the objects could be a measure of the stereopsis, other ways of measuring the task may be more important in a comparison of binocular and monocular performance. Saladin (2005) described the qualitative factors of speed, reliability-robustness and strength of percept as important aspects to consider, in addition to the minimum disparity detectable (seconds of arc), when thinking about stereopsis. The time taken to detect differences in depth (speed), how constant or changeable their ability to detect depth is (reliability-robustness) and how strong their sensation of perceiving depth is (strength of percept) are alternative considerations not classically measured or captured by clinical stereopsis tests (Saladin, 2005). This view was shared by O'Connor and Tidbury (2018) who reported clinical investigations of stereopsis are unlikely to capture all aspects of depth perception, particularly relating to dynamic real world scenarios.

In this study a measure of gross BSV or stereopsis was needed. A new coarse stereotest (CST) was under development by Professors Davis and Frisby using a similar principle to the FD2 (Ansons & Davis, 2014), and the dynamic three rods test (Matsuo et al., 2014). The CST presented several dark rods of varying height and width within a white viewing box, with one rod adjusted to be closer to the participant. When the front of the box was opened and revealed, the participant reported which rod was closer to them and their response was timed. On repeated testing, this allowed the test to quantify the number of times the participant selected the correct rod, the

disparity correctly identified, and the time taken to make the decision. These measures were chosen in attempt to capture both the accuracy and the speed of the stereopsis 'decision' by the participant. The CST was performed both binocularly and monocularly (with the strabismic eye covered) to compare the two viewing conditions.

7.5.5 Sequential stereopsis

In the absence of BSV, judgements of depth and the relative position of objects in space remain possible. People without BSV typically use monocular clues and cues to help them make depth judgements, including object size, position, colour, shading, and motion, as well as making head movements to facilitate using these cues (Levi et al., 2015). In addition to monocular cues, eye movements also contribute information about depth, even when all monocular cues are removed, and the head is still. Enright (1991) described this phenomenon as 'sequential stereopsis' when he investigated stereopsis using judgements of distance disparities under different viewing conditions. Using eye movements to look between two stimuli and make sequential comparisons of their position in space allowed improved judgements of depth. Importantly, sequential stereopsis was present even when one stimulus was visible to one eye only, as the other stimulus fell within the blind spot of the other eye. Simultaneous perception of the two stimuli was therefore not required to make judgements about their relative distance disparity (Enright, 1991).

Enright (1996) described an experimental set up to measure sequential stereopsis. Two patterned stimuli were presented behind viewing holes, in a task requiring the participant to judge the distance of the two stimuli relative to each other. The participant was required to move one of the stimuli, to make them equidistant from their viewing position. If the stimuli were sufficiently 'fine', or high pass filtered, they were visible only during foveal fixation, not during eccentric viewing. This experimental set up meant that simultaneous viewing of both fine stimuli was not possible, instead sequential stereopsis was required, where the observer was required to look from one stimulus to the other repeatedly, to be able to make judgements about their distance relative to each other. Simultaneous viewing of both stimuli was possible only when the stimuli were 'chunky' or unfiltered, as they could be seen by the fovea and peripheral retina. Once the participant had adjusted the stimuli and judged them to be equidistant from their viewing position, the researcher recorded the distance between the two stimuli. Multiple trials were completed and mean and standard deviation (SD) results were calculated (Enright, 1996). Random dot stimuli were superior to textured sandpaper (Enright, 1996; Frisby et al., 1997). Frisby et al (1997) used a similar experimental set up but recorded the mean absolute percentage error under monocular and binocular viewing conditions. Fixed and standardised head and target positions were required to ensure the results were robust (Taroyan et al., 2000).

In this study, sequential stereopsis was measured using a set up similar to Enright (1996) and Frisby et al (1997). Two different stimuli were used, an unfiltered random fractal texture with multiple grey levels and a high pass filtered version of the texture to give spatial frequencies only above 16 cycles/degree (Frisby et al., 1997; Moores et al., 1998). The unfiltered texture (comparison stimulus), was visible during foveal and eccentric viewing and the high pass filtered texture was visible during foveal viewing only. Both eyes open and monocular (with the strabismic eye covered) performance of the sequential stereopsis task was measured. Monocular viewing of the sequential stereopsis task was considered a control condition (Enright, 1991; Frisby et al., 1997). As some depth information can be gained from moving one eye (monocular kinetic information) (Jones & Lee, 1981), both eyes open performance was needed to be better than monocular performance, for an advantage from both eyes open viewing to be considered significant.

7.5.6 Contribution of the suppressed eye to both eyes open viewing

As previously described improved performance with both eyes open viewing, compared to monocular viewing with the fixing eye only, can be considered evidence that the deviating eye was contributing information to improve performance of a visual or physical task. Using this approach, the clinical tests already described are considered measures of the contribution of the apparently suppressed eye, including binocular summation of contrast (section 7.5.1), visual field (section 7.5.3), gross stereopsis measured using the CST (section 7.5.4) and sequential stereopsis (section 7.5.5).

Alternative methods of investigating the contribution of the suppressed eye have used filters or shutters to manipulate the viewing conditions and allow the presentation of stimuli to each eye separately or both eyes simultaneously. A Humphrey visual field analyser was used by Barrett et al (2012, 2013) to perform a static visual field test using blue stimuli presented on a yellow background. The test was performed both eyes open and additionally using a yellow filter over one eye to create dichoptic viewing. The blue stimulus was only seen by the eye not looking through the filter, but the eye looking through the filter could still see the background. The test was performed with the yellow filter over each eye, measuring the visual field of the dominant eye and of the non-dominant eye of strabismic amblyopes (Barrett et al., 2012, 2013). Griffiths et al (2011) described two experiments using saccadic eye movement tasks, recorded with an infra-red eye tracker in subjects with manifest strabismus and suppression. Stimuli were viewed through liquid crystal shutter goggles, which allowed the presentation of stimuli to both eyes, or to the fixing eye or the strabismic eye only. One experiment used a distractor stimulus and the other experiment presented disconjugate stimuli, however both eyes were open during the tasks. Despite subjects with strabismus being unaware of seeing the stimuli presented to their suppressed eye, their saccadic latency, planning and adaptation was affected by these stimuli (Griffiths et al., 2011).

These methods were used because they allowed manipulation of the viewing conditions without changing the motor position of the eyes during the task. In the presence of a manifest strabismus, occlusion of the fixing eye changes the position of the strabismic eye as it takes up fixation. In this study however, the performance of the strabismic eye only was not typically under investigation. Occlusion was only used to cover the strabismic eye, meaning a change in eye position was not expected as the fixing eye would maintain fixation throughout. Due to the selection of other tests to be performed both eyes open and during monocular viewing with the fixing eye (binocular summation of contrast, visual field, CST, sequential stereopsis and balance) no additional test of the contribution of the apparently suppressed eye was selected.

7.5.7 Task performance

One task performance measure was selected for the core study visit measures, the grooved pegboard pin insertion task, performed with both eyes open, under bright and dim lighting conditions. The discussion of the different task performance tests is included below in section 7.6.

7.6 Group one measures – study visit

7.6.1 Eye hand coordination and task performance measures

The terms eye hand coordination and task performance are similar. For the purposes of this study, the description ‘task performance’ was used. Tasks or daily activities have been reported to be improved following strabismus surgery, but the evidence does not specify or explore which tasks or daily activities may be improved (Beauchamp et al., 2005a; Hatt et al., 2012a; Nelson et al., 2008). A range of different tasks can be used to measure task performance and the measurement can include time (for example, completion time) accuracy (for example, by scoring errors made during the task), or both.

To investigate task performance and binocular vision participants with BSV have performed a variety of tasks with binocular and monocular viewing, typically by occluding one eye (Jones & Lee, 1981; Read et al., 2013; Sheedy et al., 1986). Task performance has been used in addition to standard clinical measures to compare the effect of an intervention, for example the correction of presbyopia (Sheedy et al., 1991), correction of refractive error (Sheedy et al., 1992), or a neurosurgical procedure (Raw et al., 2017). Task performance and visual measures have been used in before and after studies, to measure the effect of watching 3D television (Read et al., 2016; Read et al., 2015). Task performance studies have also compared participants with different levels of stereoacuity (Lenoir et al., 1999; Mazyn et al., 2007; Murdoch et al., 1991) or other clinical factors like sensory fusion, motor fusion and VA (O'Connor et al., 2010b). Grouping participants

based on levels of stereoacuity has been combined with binocular and monocular task performance testing in an attempt to further differentiate the benefit of stereopsis (Fronius & Sireteanu, 1994; Grant et al., 2007; Mazyn et al., 2004; O'Connor et al., 2010a; Schiller et al., 2012; Suttle et al., 2011). Participants with normal BSV wearing lenses to simulate vision reduction in one eye (Piano & O'Connor, 2013) and extended periods of occlusion (for up to 5 days) were used to investigate how much task performance was reduced by acquired monocular viewing (Sheedy et al., 1986).

Not all tasks yielded the same results in task performance studies. Schiller et al. (2012) reported all tasks (needle threading, rod insertion and a touch screen task) were able to differentiate between different levels of stereopsis, but their rod insertion task had greater variability in the results. Sheedy et al. (1986) found tasks requiring stereopsis were better able to measure the significant advantage of binocular visual performance compared to monocular performance, such as their inserting pointers into straws task and bead threading task. Others have suggested tasks requiring finer motor skills were better at measuring binocular advantage. For example, bead threading was shown to differentiate between different levels of stereoacuity (O'Connor et al., 2010a) and between BSV and suppression (O'Connor et al., 2010b). Bead threading and a peg insertion task (Purdue pegboard) were found to be better task performance measures than water pouring (Piano & O'Connor, 2013). Other tasks considered poorer at measuring binocular advantage included counting letters within text and throwing (Sheedy et al., 1986).

Jones and Lee (1981) suggested lighting level during the task was an important factor, with binocular advantage more evident in dimmer lighting conditions. Standardising testing conditions and practice prior to task performance was an important consideration as some tasks had a large practice effect, such as the bimanual cup stacking task (Foerster et al., 2011). Prior experience may also affect performance, for example previous experience of making depth judgements was suggested as a possible reason for participants having better than average (or expected) task performance despite lacking stereopsis (Murdoch et al., 1991). The ability to make head movements, or not, may affect task performance results. Jones and Lee (1981) proposed their allowance of head movements during task performance was a likely reason for differences in their results compared to others. They hypothesised that during monocular task performance depth information could still be gained by moving one eye or moving the head, as this would result in 'monocular kinetic' information.

The task performance measures selected for this study were the grooved pegboard (Lafayette Instrument Company, 2014-2015) performed both eyes open in bright light and dim light. This was selected as a core measure to be completed during the study visit (section 7.5.7) similar to the studies of Almuklass et al. (2017) and Sheedy et al. (1986). The additional factor of lighting was

selected as the findings of Jones and Lee (1981) suggested that any advantage of using the two eyes, rather than one, may be more evident under dimmer lighting conditions.

As part of the group one study visit (section 7.6) additional task performance measures were chosen. The Purdue pegboard (Lafayette Instrument Company, 2015) was a standardised and validated test that has been used in other studies (peg insertion task) (O'Connor et al., 2010a, 2010b). In this study the peg insertion task, performed one handed using the preferred hand, and the more complex two-handed assembly task were selected. Bead threading, using larger and smaller beads was selected, similar to the tasks used by (Jones & Lee, 1981; O'Connor et al., 2010a, 2010b; Piano & O'Connor, 2013; Sheedy et al., 1986). Different bead sizes had been suggested to create different levels of difficulty within the task (O'Connor et al., 2010a, 2010b; Piano & O'Connor, 2013).

A two-dimensional (2D) touchscreen task was used to quantify pointing and spot touching (Lee et al., 2013). Pointing to virtual objects in 3D space was shown to improve with binocular viewing. Disparity was thought to improve the control of the moving hand as well as the estimation of object position (Hu & Knill, 2011). Pointing has been shown to be abnormal in strabismus (Fronius & Sireteanu, 1994) with differing results in alternating and constant manifest strabismus (Mann et al., 1979). The Lang two-pencil test qualitatively assesses the ability to localise and touch an object in 3D space (Ansons & Davis, 2014; LaRoche & Von Noorden, 1982; Mojon, 2009), but it is difficult to measure. Tasks such as buzz wires have been used to quantify task performance in 3D space (Read et al., 2013). Touchscreen tasks have the drawback of being performed on a 2D surface, rather than in 3D, yet a touchscreen task was used successfully by Schiller et al. (2012) and (Lee et al., 2013) to measure task performance in strabismus. In this study a pointing task was selected to be used on a large touchscreen to investigate pointing and touching spots using a similar task to Lee et al. (2013).

Other screen based tasks available to measure dynamic task performance include the Clinical Kinematic Assessment Tool (CKAT) where a stylus is used (Culmer et al., 2009) and the iPad based Lee-Ryan eye hand coordination application (L-R app) (Lee et al., 2014). The L-R app used tracing, with time and accuracy scored (Lee et al., 2014) and the CKAT used tracking, aiming and steering (similar to tracing) with all tasks scored individually (Culmer et al., 2009). The CKAT had been used in paediatric (Cunningham et al., 2019; Flatters et al., 2014; Shire et al., 2016; Snapp-Childs et al., 2015) and adult studies of task performance (Raw et al., 2012; Raw et al., 2017). The CKAT was selected for this study as it was more widely reported and measured a different task to touching static spots on a 2D touchscreen.

7.7 Group two measures – study visit

7.7.1 Accommodation

Accommodation has been measured in other studies investigating task performance and visual performance. Sheedy et al. (1992) found no difference in accommodation between glasses and contact lens wearers. Accommodation was a jump task, alternating between charts at 40cm and 3m. The time to perform the task and the number of errors made were recorded. As part of their investigation of the effect of watching stereoscopic 3D television, Read et al. (2016) measured near point of convergence with the RAF rule and the accommodative convergence / accommodation (AC/A) ratio using the gradient method at 6m. Near point of convergence did not change following 3D television watching, but AC/A ratio tended to increase, although there was less increase in those watching more stereoscopic 3D content. The selection of the accommodative task in both studies was specific to their topic of investigation. A jump task was chosen by Sheedy et al. (1992) to investigate the hypothesis that contact lens wear would adversely affect accommodation during a jump task where eye movements would be required (Sheedy et al., 1992) and an AC/A ratio was selected by Read et al. (2016) to investigate the strength in the relationship between accommodation and convergence. For this study accommodation was selected to be measured objectively, both eyes open using a near target and a distance target.

7.7.2 Prehension

Prehension, the action of taking hold of an object, using a grasping movement of the hand, is typically measured using detailed kinematics of the hand during the reach and grasp. Participants can wear infrared emitting diodes (IREDs) detected by motion capture cameras during a prehension task (Connolly & Goodale, 1999; Greenwald et al., 2005; Loftus et al., 2004; Melmoth & Grant, 2006; Melmoth et al., 2007; Mon-Williams & Dijkerman, 1999; Servos et al., 1992; Watt & Bradshaw, 2000). Additionally, IREDs may be placed on the object to measure object movement (Watt & Bradshaw, 2000) or liquid crystal shutter goggles may be used to present to the experiment monocularly and binocularly (Grant & Conway, 2015; Melmoth et al., 2009).

BSV has been shown to improve prehension (Loftus et al., 2004) by making it faster and more accurate (Grant et al., 2007; Melmoth & Grant, 2006). Binocular cues were used faster than monocular cues (Greenwald et al., 2005). Stereopsis (Servos et al., 1992) and vergence (Mon-Williams & Dijkerman, 1999) have been shown to be particularly important in prehension. Vergence was shown to be used in the planning of the reach (Melmoth et al., 2007), visual information was used during the reach (Connolly & Goodale, 1999), binocular information was used in the final part of the reach (Melmoth & Grant, 2006) and the grasp (Melmoth & Grant, 2006;

Watt & Bradshaw, 2000) and disparity was important to the control of the grasp (Melmoth et al., 2007).

Amblyopia (Grant et al., 2007), reduced stereopsis (Grant et al., 2014; Melmoth et al., 2009) and both amblyopia and reduced stereopsis (Grant & Conway, 2015; Suttle et al., 2011) have been associated with worse (slower and less accurate) prehension. Improving VA with amblyopia treatment, did not improve prehension as much as improving stereopsis (Grant et al., 2014) however, prehension studies have tended to focus on amblyopia and level of stereopsis, rather than on strabismus. Amblyopia prehension studies have included different types of strabismus (microtropia, small angle strabismus with abnormal BSV and constant strabismus with suppression) in either a reduced or no stereopsis group. This made it difficult to draw precise conclusions about prehension in different subtypes of strabismus (Grant & Conway, 2015; Grant et al., 2007; Grant et al., 2014; Melmoth et al., 2009; Suttle et al., 2011). All reductions in stereopsis led to poorer prehension, although prehension with coarse stereopsis (Wirt stereotest 100-3000” of arc) was better than with no stereopsis (Melmoth et al., 2009).

Prehension was considered important to include. For this study a large-scale optical motion capture system was not feasible, instead a portable system Boxed Infrared Gross Kinematic Assessment Tool (BIGKAT) was selected. IREDs placed on the hand and the BIGKAT was selected to capture the position of the IREDs in space, and relative to each other, as the participant reached to grasp an object. The kinematics of the hand during reach to grasp movements were measured.

7.7.3 Balance

Balance, movement and stability may be measured using IREDs and motion capture systems, similar to prehension (Buckley et al., 2010b; Chapman et al., 2012; Shaheen et al., 2018). Participants with normal BSV, under binocular and monocular viewing conditions, and participants with reduced stereopsis have been used to investigate the effects of stereopsis on balance and movement. During foot placement onto moving targets Chapman et al. (2012) found monocular information was sufficient for the initial stages of the step, but binocular information was required for the final part of the movement to ensure the foot was placed accurately. Buckley et al. (2010b) reported stereo deficient participants were more cautious crossing higher obstacles, however their reduced stereopsis group ranged from constant strabismus and no stereopsis to microtropia and reduced stereopsis. Interestingly, monocular performance was always worse than binocular, suggesting both eyes contributed to the performance of the task, even in the reduced stereo group. Shaheen et al. (2018) measured walking on stairs and found in people with AMD, those with more careful negotiation of steps in poor lighting had better balance.

Chung et al. (2018) videoed completion of a mobility course incorporating obstacles and hazards to score performance, including collisions and completion time. Hayhoe et al. (2009) videoed a walking task to analyse foot clearance when stepping over obstacles in combination with EMR to measure what participants were looking at. Monocular performance led to slower task performance, a higher step over obstacles, looking at obstacles for longer and overall poorer (more cautious and slower) performance. Floor pressure sensors and worn accelerometers have been used to measure balance and stability during task performance. No effect of 3D television viewing was found on balance and coordination (Read et al., 2016; Read et al., 2015). Timed completion can also be used to measure mobility tasks. West et al. (2002) measured activity completion time, including a 4-metre walk, ascending and descending stairs, and a timed get-up-and-go test. Older participants with poorer vision and contrast sensitivity performed worse on these mobility tasks. Hospital records have also been analysed to look for relationships between strabismus surgery and falls, fractures or musculoskeletal (MSK) injuries. Pineles et al. (2017) found no association between strabismus surgery and a reduced risk of falls, fractures or MSK injuries.

When measuring balance, tests need to be sufficiently sensitive but allow a range of participants to complete the task. Adult participants with chronic pain were often unable to complete a balance beam and the one-foot balance had poor test-retest variability. Walking was used as a proxy balance measure as it adequately captured gross balance problems (Harding et al., 1994). The balance subtest from the Bruininks-Oseretsky Test of Motor Proficiency clinical test battery was used by Zipori et al. (2018) in a paediatric study. Strabismic children with and without amblyopia had reduced balance compared to controls, particularly with challenging balance tasks. No difference was found between ET and X or X(T). A platform measuring vertical force was used by Matsuo et al. (2006) to measure increased body sway and centre of pressure shifting towards the heel in children 3-days post-strabismus surgery. Preoperatively having no stereoacuity was associated with greater body sway compared to having stereoacuity. A force platform was also used by Bucci et al. (2016) to measure posture (centre of pressure surface area, length and mean speed) in children with a range of constant and intermittent strabismus types, before and after (2-6 months) strabismus surgery. Postural control improved postoperatively, although the majority had intermittent deviations and BSV.

The Wii balance board (Nintendo, Kyoto, Japan) has been suggested as an appropriate clinical tool for measuring standing balance (Clark et al., 2010). Whilst not as accurate as balance laboratory pressure plates, it gave an accurate measure of force and centre of pressure, particularly on repeated measurement using the same Wii balance board (Bartlett et al., 2014). Studies using the Wii balance board in clinical populations have found it compared favourably to pressure plate systems in stroke survivors (Llorens et al., 2016), but in patients with multiple sclerosis and minimal balance impairment, it tended to overestimate some measures of balance

(Severini et al., 2017). Imaizumi et al. (2018) used the Wii Balance Board to measure centre of pressure during a balance task with and without visual feedback. They found visual feedback helped participants control their centre of pressure when they specifically instructed to do so. In this study static balance was considered more feasible to measure than a movement task and the Wii balance board was selected as the measurement device.

7.8 Eye movement recording - optional extra study visit

EMR can be performed using a range of different eye trackers. Increasingly video-based eye trackers have been used due to increasing data quality and reducing cost. Eye trackers can measure gaze direction during a task (Kurz et al., 2017), or can measure eye movements made in detail, for example smooth pursuit, OKN, saccades and fixation. The scleral search coil was used to record eye movements in an adult patient with incomitant strabismus. Despite no BSV, Bucci et al. (1999) concluded the patient had some rudimentary binocular cooperation between the two eyes, due to changes in saccadic behaviour measured during monocular and both eyes open viewing.

In strabismic children EMR before and after strabismus surgery have been performed monocularly using a photoelectric eye tracker (infrared limbal Skalar IRIS device) (Bucci et al., 2002) and binocularly using a different device (Oculometer, Bouis) (Bucci et al., 2009; Bucci et al., 2006). Bucci et al. (2006) and Bucci et al. (2002) included a range of strabismus types, with and without BSV and recorded eye movements in the early postoperative period (2-8 weeks). Postoperatively saccadic disconjugacy reduced, but post saccadic drift was unchanged. Bucci et al. (2002) concluded improved eye alignment or central adaptation could have caused the improved saccade conjugacy, however they were unsure whether BSV affected their findings. Saccadic and vergence latency were largely unaffected by surgery, although latency did reduce in some individuals (Bucci et al., 2006). Later, in follow up recordings improved saccade and vergence accuracy and improved speed for some of the vergence eye movements were measured (Bucci et al., 2009).

Niechwiej-Szwedo et al. (2010) used a Chronos video-based eye tracker and measured increased saccadic latency in adults with anisometropic amblyopia. Induced blur did not affect saccades, reaching and eye hand coordination to the same extent as having anisometropic amblyopia (Niechwiej-Szwedo et al., 2012). Adults with amblyopia, strabismus without amblyopia and normal controls had their eye and hand movements recorded. Amblyopia and reduced stereopsis, but not strabismus, were associated with increased latency (saccade and limb movement) and less precise localisation of the target (Niechwiej-Szwedo et al., 2017). Taylor et al. (2017) analysed search time, fixation duration, number of saccades per second and saccade amplitude using the Eyelink II. Worse visual search performance was measured in participants with AMD compared to

controls. The Eyelink II was used to measure fixation and an infrared system (ReadAlyzer) was used to measure eye movements in children during reading both eyes open. Anisometropic amblyopes were found to have slower reading, increased saccades and less stable fixation (Kelly et al., 2017). The Eyelink II (Açık et al., 2009) and Eyelink 1000 (Açık et al., 2010) have been used to measure fixation behaviour when participants with normal vision viewed different scenes. Ju et al. (2018) used the Eyelink 1000+ to measure vision and motor performance in children to evaluate a sports intervention. Saccades and the time to achieve foveal fixation were used as the measures of vision performance. Whilst the task was performed both eyes open, only data from the dominant eye were analysed.

Busy and complex visual environments may be uncomfortable or difficult for some patients with strabismus to experience (Pineles, Lee, et al., 2014) therefore visual tasks incorporating busy and complex visual environments were considered suitable to measure visual performance in strabismus. In this study the Eyelink 1000+ was used for EMR, specifically fixation, smooth pursuit and saccades. Complex visual scenes were incorporated into the stimuli to try and simulate busy environments and real-world viewing tasks. Due to technical difficulties with EMR simultaneously from both eyes in the presence of large angle strabismus, the tasks were performed both eyes open, but the EMR were measured from the fixing eye.

7.9 Control group

From the reviewed literature evidence, it was apparent that a control group was used infrequently in studies of strabismus or ocular conditions. Studies with a control group had the advantage of being able to measure how much performance was affected by the condition under investigation, or measure how much performance changed due the repeated measures study design. Healthy adult volunteers with ‘normal’ vision may be useful as a control group to some extent, but a control group of adults with strabismus, no BSV, who are not undergoing strabismus surgery was considered a more robust and representative control group. An ‘ideal’ control group would be a population of adults with strabismus who had no previous surgery, or treatment for strabismus, yet this was considered difficult to achieve in practice. Therefore, previous treatment for strabismus was allowed in the control group. This study used a strabismus control group to establish the repeatability of the tasks in the quantitative phase and to ensure any conclusions drawn about the effects of surgery were not due to an inadvertent learning effect. It is possible that any change observed in a repeated measures study design may be due a practice or learning effect, therefore the inclusion of a control group was considered an important part of the interpretation of whether change in the surgery group had occurred postoperatively.

7.10 Chapter 7 summary

In chapter 7, the process of selecting and refining the quantitative phase two measures was described. The measures were selected based on the findings from the qualitative phase (chapter 6). Where several different clinical or experimental tests were available, the literature evidence was used to select and refine the measures further. In chapter 8 the methods used in the quantitative phase will be described.

Chapter 8. Quantitative Methods

In the previous chapter 7 the selection and refinement of the measures for the quantitative phase were described. In this chapter 8, the methods used in the quantitative phase two of the study will be described. In the following chapter (chapter 9) the quantitative results will be reported.

8.1 Quantitative study design

The quantitative phase two was a prospective repeated measures study, with a before and after design. Participants were recruited to the surgery group or the control group. The surgery group included adults with constant strabismus, no BSV and no potential BSV, who were planning to undergo strabismus surgery for psychosocial reasons. The control group included adults with constant strabismus, suppression and no BSV, who were not planning to undergo strabismus surgery.

8.2 Ethical approval

Ethical approval from a REC (Appendix C) and approval from the HRA (Appendix C) were granted prior to commencement of the quantitative phase. Following the lockdown in England due to the COVID-19 pandemic, a non-substantial amendment was submitted to and authorised by the study sponsor (STH NHS FT) (Appendix C).

8.3 Quantitative setting, sampling and recruitment

8.3.1 Setting

The setting for the mixed methods study was the Ophthalmology Department, STH NHS FT (described in section 5.2) and the Division of Ophthalmology and Orthoptics, Health Science School, University of Sheffield.

8.3.2 Sampling and recruitment

Recruitment of fifty-five participants to the surgery group and twenty participants to the control group was planned. The surgery group size was selected to combine the smallest size of the feasibility study and larger study for small to medium standardised effect sizes (Sim & Lewis, 2012). As the study was a feasibility study, a sample size calculation was not undertaken to achieve statistical power.

Participants were recruited to the surgery group from the Ophthalmology Department, STH NHS FT. Participants were recruited to the control group from volunteers known to the University of

Sheffield as staff, students or 'Patients as Educators'. All participants were adults with constant strabismus, suppression and no evidence of BSV (section 7.3.3). Surgery group participants were approached in the orthoptic clinic and invited to participate in the study if they met the inclusion criteria and had decided to undergo surgery for psychosocial reasons. Control group participants were approached by letter or email and invited to participate in the study if they met the inclusion criteria. Written information (Appendix M and N) and an opportunity to discuss the study were given to each participant prior to them deciding on participation.

Inclusion criteria

- Adults (age 18-years or older, with no maximum age limit)
- Heterotropia of any size with suppression (without diplopia and without BSV)
- Any measurable VA (up to light perception)
- Surgery group - offered and decided to undergo strabismus surgery for psychosocial reasons

Exclusion criteria

- Patient unable to complete required study visits.
    - Surgery group: one pre-operative visit (within one month of surgery) and one post-operative visit (3-12 months after strabismus) surgery.
    - Control group: visit one and visit two (3-12 months after visit one)
- No measurable VA in poorest seeing eye, e.g. no perception of light
- Co-existing health condition known to affect the structure and function of extraocular muscles, e.g. Graves' orbitopathy
- Unable to understand or communicate in English
- Surgery group - surgery planned for a reason other than to improve eye alignment for psychosocial reasons, e.g. to regain BSV or eliminate diplopia symptoms
- Surgery group - no strabismus surgery performed

8.4 Quantitative methods and data collection

After agreeing to participate in the study written informed consent was taken from all participants (Appendix O). Two study visits were conducted at least 3 months apart. Visit one for the surgery group was typically the preoperative assessment and visit two was typically the 3-months postoperative assessment. The control group selected convenient study visit dates.

Surgery group participants had clinical information extracted from their orthoptic report (described in section 7.3). The researcher (GA) performed the majority of these orthoptic assessments in the same clinical space. Clinical history, recent refractive correction, near and distance VA, CT and

fixation grading, investigation of BSV, OM and PCT in primary and secondary positions of gaze were recorded. Control group participants had these tests performed during their study visit by GA.

The same university room was used for all study visits, using standardised instructions and testing methods. All tests were positioned in the same place within the room for each visit. The Thomson Test Chart Xpert 2000 was used to display binocular summation and contrast sensitivity stimuli on a HP ZR24W screen (24inch widescreen, resolution 1920 x 1200 pixels). Prior to data collection and regularly during testing the screen was calibrated using the X-Rite Colormunki display. The screen was turned on at least 30 minutes prior to data collection. Participants were always seated on a height adjustable chair to ensure they were comfortable and at an optimum height for each task. The room was well lit and room lighting was kept constant, unless otherwise stated.

For all tests the participant wore appropriate, up to date, refractive correction, unless otherwise stated. All tasks were performed both eyes open. Some tasks were additionally performed monocularly, to allow comparison between monocular and both eyes open performance, these are described below. Monocular performance was assessed with full occlusion of the strabismic eye, using an adjustable opaque plastic patch, unless otherwise stated. All participants had the following tests performed at visit one and visit two, using the same method, in the same order, unless specified (Figure 7-2).

8.4.1 Visual acuity – near and distance

Near VA

Near VA was tested with a 100% contrast reduced ETDRS chart at 40cm. Each side of the chart was presented randomly during testing. Participants were asked to read aloud the lowest line they could see. Pointing to the beginning of the row of letters, from the side, was allowed. Pointing underneath individual letters was not allowed. The whole of the chart was always visible. Encouragement was given to see the letters, especially around the threshold level, but participants were not encouraged to guess if they were unable to see the letters. If a letter was read incorrectly, a correction of the letter was allowed. If a participant repeatedly guessed a letter incorrectly, this was not scored as correct to prevent continuous guessing. The ETDRS chart contained Sloan letters (D K O H N S R C V Z). Each letter had to be identified correctly to be scored as correctly seen. No letters were accepted as interchangeably correct. Results were scored in logMAR using the termination rule of four mistakes on a row (Carkeet & Bailey, 2017).

Distance VA

Distance VA was tested monocularly. A 100% contrast ETDRS chart, presented in a back illuminated cabinet, was used at 3m or 4m (depending on the room). Right and left eye testing was randomised during the test. An opaque black plastic occluder was used to occlude the non-viewing

eye. Chart versions were changed during testing to prevent participants remembering the letters. The same testing and scoring methods were used as for near VA. For participants with less than (or equal to) 1.080 VA, the crowded logMAR VA test was presented at a reduced test distance and scored up to 1.875 logMAR. If VA was less than 1.875, counting fingers (yes or no, test distance recorded), hand movements (yes or no) and perception of light (yes or no) were used.

8.4.2 Cover test and fixation grading

The CT was performed at near (33cm) & distance (4m or 6m) using a target visible to the poorer seeing eye. The CT results were described using standard clinical terminology (British and Irish Orthoptic Society, 2015). Fixation was observed at near during the CT and in free space. One fixation score was given, using a 1-7 grading scale (described in section 7.3.2). Originally a 1-6 scale was planned, however postoperatively heterophoria (grade 7) was required and was added to the scale (non-substantial amendment, Appendix C).

8.4.3 Binocular Single Vision

All participants had the presence of sensory fusion or suppression tested using Bagolini glasses at near (33cm) and distance (4m or 6m). Preoperatively the surgery group had potential BSV investigated using loose prisms in free space in combination with Bagolini glasses. Additionally, potential BSV was investigated using the Synoptophore if the Bagolini glasses result was unclear, or if the history suggested potential BSV was more likely. Participants were only included in the surgery group if no BSV and no potential BSV was demonstrated.

Postoperatively if there was evidence of sensory fusion, further BSV tests were performed. Motor fusion was measured using the PFR with Bagolini glasses (to confirm the end point of the test if suppression occurred). Stereoacuity was measured using the Wirt stereotest and FD2. If stereoacuity was better than the Wirt fly (3552” of arc) and the first animal (800” of arc), testing was also performed using the FNS.

8.4.4 Ocular movements

OM were tested without glasses using the standard clinical technique. A pen torch was held at approximately 50cm from the participant and moved from primary position, into each of the secondary and tertiary positions of gaze. Ductions and versions were tested in each position using a black plastic occluder. Results were documented diagrammatically. Grading of 0.5 intervals was allowed. Small limitations less than -0.5 could be graded as -0.25.

8.4.5 Prism Cover Test

Measurements of the primary and secondary position deviations were performed using the APCT with loose prisms. A fixation target was used appropriate to the level of VA of the worst seeing eye for each participant. Primary position measurements were performed at near (33cm) and distance (4m or 6m) with glasses, if required. Secondary position (right gaze, left gaze, upgaze and downgaze) measurements were performed at distance without glasses. If glasses were removed for secondary position measurements, the distance primary position measurement was repeated without glasses, to allow the primary position measurement to be directly compared to the secondary position measurements. If a participant was unable to fixate with each eye for an accurate APCT, a prism reflection test (Krimsky's test) (Ansons & Davis, 2014) was performed at near to measure the deviation.

PCT results (PD) were recorded separately for the horizontal and vertical elements of the deviation. Eso and hyper deviations were reported as positive values. Exo and hypo deviations were reported as negative values. Absolute values of the horizontal and vertical deviations were used in the analysis of the deviation, to avoid horizontal and vertical deviations of opposite directions returning a mean of zero. The change in deviation was calculated by subtracting the visit one deviation from the visit two deviation. The magnitude of the change in the deviation was considered important, but also whether the deviation had reduced or increased in size. Positive values were assigned to improved (reduced) horizontal deviations and negative values were assigned to worsened (increased) horizontal deviations.

8.4.6 Binocular summation of contrast

Binocular summation of contrast was measured using British standard letters presented on the screen (screen set up described in section 8.4) using the Thomson Test Chart Xpert 2000. Testing was performed at 4m, or at 1m if the participant was able to see less than 20 letters at 4m. Each letter had to be identified correctly to be scored as correctly seen. No letters were accepted as interchangeably correct. Letters were presented at 100%, 10%, then 5% contrast. At each level of contrast participants were tested monocularly (using an opaque black plastic occluder) and then both eyes open. For monocular testing the better seeing eye was tested first. If the participant had equal vision, the right eye was tested first. The letters were randomised prior to each presentation to prevent memorisation of the letters. The same testing method was used as for VA (section 8.4.1).

The VA threshold was recorded in logMAR (same method as near and distance VA) and the number of letters correctly seen for the right eye, left eye and both eyes open at 100%, 10% and 5% contrast (maximum 100 letters – minimum 0 letters). The number of letters correctly seen was scored using the ETDRS protocol (Early Treatment Diabetic Retinopathy Study Research Group, 1991). A binocular summation score (letter scoring method) was calculated for each contrast level (100%, 10% and 5% contrast). The binocular summation categorisation method was additionally reported.

Binocular Summation Letter Scoring Method

Binocular Summation score (number of letters) = both eyes open score (number of letters) - monocular score of the better eye (number of letters)

Binocular Summation Categorisation Method

Binocular Summation = a binocular summation score ≥ +5

Binocular Indeterminate = a binocular summation score -4 to +4

Binocular Inhibition = binocular summation score ≤ -5

8.4.7 Contrast sensitivity

Contrast sensitivity was tested both eyes open to threshold at 1m. Sloan optotypes were displayed in triplets at 1.5 logMAR on the screen (described in section 8.4) using the Thomson Test Chart Xpert 2000. The optotypes were randomised for each presentation. Each participant was shown the test at 100% contrast (0.0 log contrast sensitivity) and asked to read out the letters as they were gradually reduced in contrast. All letters had to be read correctly to be scored as correctly seen, no letters were accepted as interchangeably correct. The optotypes were reduced in contrast until the participant was unable to read any of the three optotypes accurately (Elliott et al., 1991). At the threshold level of contrast (log contrast sensitivity score) participants were given 20 seconds to try to see the optotypes (Elliott et al., 1991; Haymes et al., 2006).

8.4.8 Sequential stereopsis

Sequential stereopsis was tested using a custom-made rig placed on a table, using a similar technique described in section 7.5.5 (Enright, 1991, 1996; Frisby et al., 1996; Frisby et al., 1997; Moores et al., 1998; Taroyan et al., 2000). Descriptions of the sequential stereopsis rig and set up are shown in Appendix P. The participant placed their chin and forehead on a rest and kept their head still during testing. Head position and movement was constantly monitored. Participants were shown the task as a demonstration whilst standardised instructions were given to move the

stimulus on the right until it appeared equidistant to the stimulus on the left, from their position on the chin rest.

Participants looked through each of the viewing ports, continually making eye movements to look through each port in turn. Only the stimuli were visible behind the viewing ports. The participant gave the examiner instructions to move the stimulus on the right. Whilst the examiner was adjusting the position of the stimulus on the right, the digital callipers measuring the difference between the left (static) and right (moveable) stimuli were not visible. When the participant was satisfied the two stimuli were equidistant from their viewing position, the examiner read off the digital calliper, the distance between the stimuli in mm. After the demonstration trial, one practice trial was allowed before the experimental trials were performed and the results recorded.

Six trials were completed for each of the four viewing conditions: both eyes open unfiltered, monocular unfiltered, both eyes open high pass filtered, and monocular high pass filtered. This sequence of testing was chosen as it was assumed to progress from easiest to hardest to give participants the greatest chance of performing the task well. The starting position of the stimuli were randomised during testing. At visit two the same method of testing was used as at visit one, however the stimuli in front of the left eye and the right eye were both rotated 180 degrees, so that the stimuli positions or views through the viewing ports could not be remembered.

The results were recorded as the distance between the stimuli in mm when the participant considered them to be equidistance from their viewing position. The absolute percentage error of each presentation was calculated using the following formula.

Absolute percentage (%) error = (absolute value of the recorded error (mm) / distance of left static stimulus from the eye (mm)) x 100

A mean absolute % error value was calculated from the six absolute % error values for each viewing condition. A higher mean absolute % error equated to worse performance of the test.

8.4.9 Gross binocular single vision (BSV) or stereopsis

Gross BSV was assessed using the CST, which presented stimuli at 3094” of arc using a similar principle to the FD2 (set up shown in Appendix Q). The participant sat with their forehead positioned on a rest to prevent head movement. The CST was presented at eye level with the midpoint of the height aligned with the lateral canthus and the midpoint of the width aligned with the nasal bridge. The participant viewed the CST at 1m (eye to the position of furthest rod).

Prior to presentation of the CST five rods were selected (from the total of 10) and placed on the magnetic base. Unselected rods were hidden from view. The rod positions were changed and moved with the viewing door closed. Four rods were positioned at the back of the magnetic base and one was moved to the front of the base, 25 cm apart. The horizontal distance between the rods was also adjusted for each presentation. The selection of the rods, their position during each presentation and their counterbalanced test order were decided prior to testing to ensure the CST was consistent for all participants (Appendix Q).

Participants were given standardised instructions explaining the test and demonstrating the five rod positions. One practice trial (both eyes open) was given. Ten experimental trials were shown for each viewing condition. For each presentation participants reported which of the five rods was closest to them, as quickly as possible. The rod reported (correctly or incorrectly) and the time taken to respond (seconds) were recorded. If the incorrect rod was reported by the participant, both the rod (A-J) and the position (1-5) were recorded. The CST was performed under both eyes open and monocular viewing conditions. The viewing condition used first was alternated between participants.

The recorded results included accuracy (number of presentations reported correctly out of a maximum 10 presentations), response time (total time to respond, correctly or incorrectly, to all the presentations (seconds)) and the Rate Correct Score (RCS) (Vandierendonck, 2017).

Rate Correct Score (RCS) (number of correct responses per second) = number of correct responses / by the total reaction time for all 10 presentations (seconds)

8.4.10 Visual field

The visual field was tested using the Goldmann Perimeter 940 following calibration and the recommended testing procedures for the Goldmann (Haag-Streit International). The participant was set up with their chin and forehead placed on the rest. The fixing eye maintained central fixation and fixation was monitored during testing. Standardised instructions and a demonstration of the stimulus were given to participants. Six points were presented as practice trials (horizontal meridian, vertical meridian and two further diagonal points). Testing began when the participant reliably responded to the practice trial points whilst maintaining central fixation.

Peripheral vision was tested using the 4Ie target. The stimulus light was moved slowly and steadily at approximately 2 degrees / second from the periphery (unseen) in towards the centre (seen). Participants pressed the buzzer to report they were able to detect the stimulus in their peripheral

vision. This precise point was recorded manually with an ‘x’ on the paper chart inserted into the back of the Goldmann. Every 30 degrees was tested as a minimum and more intermediate points were tested if required, for example if there was a larger than expected difference between two points. A minimum of 4 points, one in each quadrant, were rechecked for accuracy. The visual field both eyes open was tested at visit one and visit two. Monocular testing of the fixing eye was performed at visit one only.

The visual field was drawn by connecting all of the peripheral points on the paper chart. The total area of the visual field and the horizontal meridian (Kushner, 1994; Wortham & Greenwald, 1989) were reported. The horizontal meridian was reported in degrees (total number of degrees across the 0-180 line in which the participant had seen the stimulus light). The total area of the visual field was measured by dividing the whole area of the visual field into twelve 30-degree triangles and calculating the total area of all twelve triangles (Buckley et al., 2010a). Triangle lengths were measured in mm (from the centre of fixation to the edge of the peripheral visual field). The following formula was used to calculate the area of each 30-degree triangle in excel. The total area of the visual field was reported in degrees^2.

Area of triangle (degrees^2) =
½ * (length of triangle a / 1.2) * (length of triangle b / 1.2) * SIN (30/57.2958)

1.2 degrees = 1 mm
SIN 30/57.2958 used in excel to convert the calculation from radians to degrees

8.4.11 Grooved pegboard

The grooved pegboard (model 32025) (Lafayette Instrument Company, 2014-2015) was performed both eyes open only. The participant was seated with the pegboard placed in front of them, centrally, on a flat table. The bottom edge of the pegboard was 10cm from the edge of the table and the dish (containing the pegs) was above the insertion holes. Twenty-five pegs were placed in the dish before the start of the task. Five pegs were inserted into the pegboard as a demonstration. A practice trial using 5 pegs (top row) was given before the task began. Participants were given standardised instructions to insert the pegs into the pegboard holes, one at a time, as quickly as possible without skipping any holes. The specific pattern of peg insertion was for the top row to be completed before moving down to the next row. The first hole to be completed on each row had to be opposite to the hand they used for the task. Holes had to be completed sequentially, moving along the row, without leaving gaps. Only one peg could be picked up at a time. Participants used their preferred hand for the task. If a peg was dropped during the task and was within easy reach of the participant, they could pick it up and continue the task.

Participants completed the pegboard task in bright (5.38 lux) and dim (2.15 lux) lighting. The starting light level was alternated between participants. The room lights were turned off and a task lamp was used. A grey plastic filter was secured over the lamp to ensure the light was the same colour in both lighting conditions. The lamp position and distance from the pegboard was kept constant during testing. One-minute adaptation time to each lighting condition was given before the task began. Testing order at visit one was repeated at visit two for each participant. The results recorded were the time taken (seconds) to complete the pegboard (all 25 pegs) and any pegs dropped.

For those participants completing the group one study visit measures (task performance) the following methods were used.

8.4.12 CKAT

The CKAT was presented on a Dell Latitude 7285 laptop with the tablet screen removed and placed in front on the participant on a flat table. Standardised instructions were given to explain the task and to hold the stylus in the preferred hand, keeping it in contact with the screen during the task. Participants performed the CKAT both eyes open and read the standardised instructions displayed on the screen before each part of the task. The CKAT task battery had been described in detail (Culmer et al., 2009; Cunningham et al., 2019). CKAT methods are summarised below. Additional details are described in Appendix R. Testing order at visit one was repeated at visit two for each participant.

Tracking

A red dot appeared in the middle of the screen. Once the participant touched the red dot with the stylus, the dot turned green and began to move in a sinusoidal horizontal figure of eight shape. Participants were instructed to always keep their stylus in contact with the screen and follow the dot as it moved around the screen. The speed of the dot increased after 3 revolutions of the shape. Three test speeds were tested (slow, medium and fast). The tracking task was firstly presented without any background guide and was then repeated with a background guide (black solid line showing the path of the dot). Tracking results were a measure of accuracy during the tracking task (Root Mean Square Error (RMSE) mm). Tracking results were recorded for the slow, medium and fast target speeds, both with and without the background guide. A smaller accuracy value equated to better performance on the tracking task.

Aiming

The aiming task began with the participant placing their stylus on the ‘start’ position. On the screen a red dot appeared at the end of their stylus. A green dot then appeared on the screen in a

different position and the participant was instructed to move their stylus (and red dot) to the green dot as quickly and accurately as possible, keeping their stylus in contact with the screen throughout. When the stylus made contact with the green dot it disappeared, and another green dot appeared in a different location on the screen. 75 aiming dots were shown. Aiming results were mean path length time (seconds) for the first 50 trials (dots). A shorter aiming time (seconds) equated to better performance of the aiming task.

Steering

The participant placed their stylus on the 'start' position (green circle). On the screen a tracing path appeared, and participants were required to trace the path on the screen as accurately as possible until reaching the 'finish' position (red circle). They were instructed to keep within the square box and keep their stylus on the screen. The square box 'paced' the participant, highlighting the portion of the path they should be completing. The pacing box moved along the path at regular intervals to allow completion of the path in 36 seconds. This prevented rushing or moving too slowly. When the finish position was reached the completed path disappeared and another path appeared. The same tracing path was used for each of the six presentations, three of the tracing paths were upright (shape A) and three were inverted (Shape B). Steering results were a penalised path accuracy score (pPA) for shape A and shape B. A lower steering pPA score equated to better performance of the steering task. A mean pPA for shape A and shape B were calculated. If no significant difference was found between shape A and shape B steering, an average was taken, generating one value for steering pPA.

8.4.13 Purdue pegboard

The Purdue pegboard (model 32020) (Lafayette Instrument Company, 2015) was performed with both eyes open. The pegboard was placed flat on a table with the cups at the top. The bottom edge of the Purdue pegboard was flush with table edge. The cups contained (left to right) pins, washers, collars and pins. Two tasks were performed, the pin insertion task then the assembly task.

Pin insertion task

Using their preferred hand participants picked up the pegs, one at a time, and placed them into the holes (using the holes on the side of their preferred hand). As many pegs as possible were inserted in 30 seconds. A demonstration (5 pegs) then a practice trial (5 pegs) was given prior to the task.

Assembly task

Using both hands participants had to construct the assemblies using a pin, washer, collar and then a washer (using the holes on the side of their preferred hand). As many assemblies as possible

were completed in 60 seconds. A demonstration (2 assemblies) then a practice trial (2 assemblies) was given prior to the task.

Pin insertion task results were recorded as the number of pins. The assembly task score was calculated using the standard scoring method. Higher scores equated to better task performance.

> Assembly task score = (number of correctly completed assemblies x 4) + number of pieces in any part completed assemblies that were correctly placed
>
> 4 = number of parts in completed assembly

8.4.14 Bead threading

Bead threading was performed with both eyes open. Participants threaded beads onto a knitting needle. 20 large beads (yellow plastic, 8.75mm maximum diameter, 6.48mm height, 3.82mm diameter hole in the centre) were threaded onto a large knitting needle (Aero size 9, 3.70mm diameter). 20 small beads (red plastic, 4.48mm maximum diameter, 4.74mm height, 2.85mm diameter hole in the centre) were threaded onto a small knitting needle (Aero, UK size 12, 2.64mm diameter). Each knitting needle was held horizontally, directly in front of the participant, in a retort clamp placed on a flat table. The clamp base was positioned 30cm from the table edge. The needle was positioned 19.5cm from the edge of the table and 15cm above the table, with the tip pointing to the right. Each needle was clamped so that 22cm of the needle was available for the threading task. The beads for the task were placed in a round dish (10.5cm diameter, 2.5cm high) which was secured to the table directly in front of the participant, in the middle of the needle, 5cm rom the edge of the table. A demonstration of the task (5 large beads) and a practice trial (20 large beads) were given. The large bead task was completed before the small bead task. Using their preferred hand participants threaded the beads one at a time onto the needle. Each bead was placed onto the needle and positioned fully to the end of the needle before the next bead was selected. If a bead was dropped the researcher replaced the dropped bead into the dish. Participants were not allowed to touch the needle to avoid them guiding the beads into place. The time taken to thread all 20 beads was recorded (seconds) for the large beads and the small beads. Shorter bead threading time equated to better task performance.

8.4.15 Touch screen spatial localisation (TSL task)

Prior to data collection, the touchscreen used to display the TSL was validated (Appendix S). During the TSL task the screen was placed flat on a table, in landscape orientation, 10cm from the table edge. The screen was positioned directly in front of the participant so they could touch all

areas of the screen. Participants had to use the same preferred finger and hand throughout the task, with long sleeves rolled up and no dangling jewellery, to prevent false touches of the screen. The TSL task was completed with both eyes open. During the TSL task a blue spot (3mm diameter) was presented on the screen and the participant was required to touch the spot as quickly and accurately as possible. Each spot remained on the screen until it was touched. Before each blue spot appeared, a red spot (3mm diameter) was shown in the centre of the screen for 3 seconds to encourage participants to begin by looking at the centre of the screen.

Each trial consisted of 10 blue spots appearing one at a time in a predetermined location (shown in Figure 8-2). Three trials were presented in the same order, but during each trial the 10 spots were presented in a random order. A total of 30 spots were presented in the same location for each participant, but in a different order. Prior to the TSL task, a demonstration trial (10 random spots) and a practice trial (10 random spots) were given.

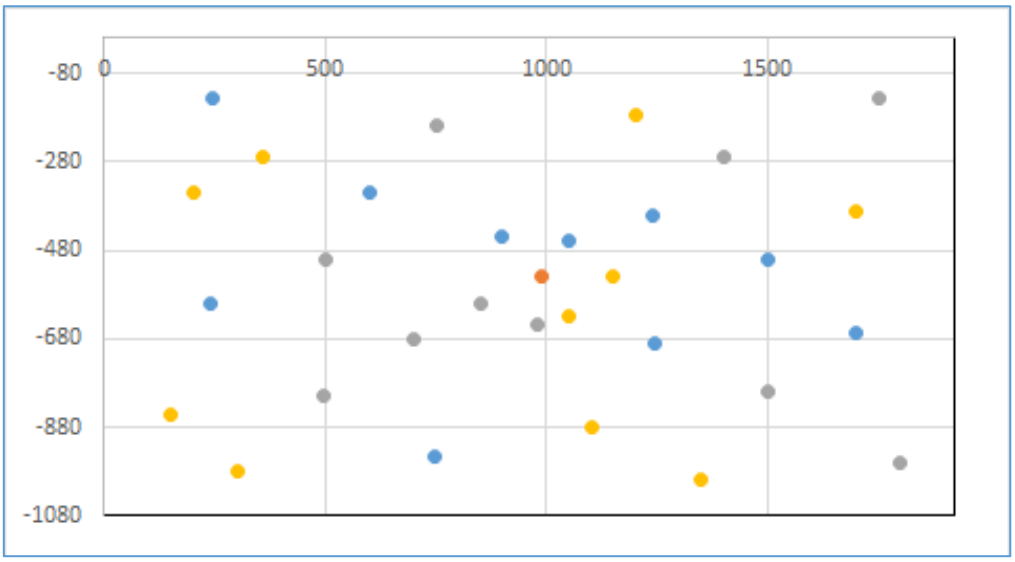


Figure 8-1. The location of the TSL spots presented in the three trials

The centre location of the screen, where the red spot was presented is shown by the orange spot. The spots in trial 1 (1-10) are shown in blue. The spots in trial 2 (11-20) are shown in grey. The spots in trial 3 (21-30) are shown in yellow.

The results recorded for each spot included the response time (time taken to touch the spot (ms)) and the location of the screen touched by the participant (pixels). The accuracy of each spot touch was calculated as the difference between the centre of the spot shown and the centre of the point touched on the screen. Accuracy was calculated in pixels and converted into mm.

Accuracy (pixels) = $\sqrt{}$ ((spot x – screen touch x)^2 + (spot y – screen touch y)^2)

Accuracy (mm) = accuracy (pixels) / 3.2

A value of 3.2 pixels / mm was generated from the screen dimensions. Screen = 600mm x 338mm, 1920 pixels x 1080 pixels. (1920 / 600 = 3.2. 1080 / 338 = 3.2)

8.4.16 PROMs (AS-20, VFQ-25 and additional study questions)

If the AS-20 PROM was completed as part of the clinical visit, the results were used in the study. If the AS-20 PROM had not been completed, participants were asked to complete the AS-20 as part of the study PROMs during the study visit. The study PROMs were presented on paper as one pack containing the additional study questions, AS-20 and VFQ-25. All 20 questions in the AS-20 were completed. The 'self-administered' version of the VFQ-25 was completed (all 25 questions) and the optional additional questions were not used. For each PROM the participant read the 'instructions for patients' before completing the questions without interruption. The researcher was available to answer any questions. Further details for each PROM are described in the following sections.

Additional study questions

The instructions, questions and scoring method used are described (Appendix L). The raw results were recorded on a spreadsheet and then converted into the scores reported for the additional study questions. The total score (33 questions) was used in the analysis (minimum 0 – maximum 3300). Subscale scores were also reported for vision (questions 1-11), task performance (questions 12-17), physical symptoms (questions 18-23) and confidence and emotions (questions 24-33). Subscale scores were converted to a score out of 100 and reported as a mean score each subscale (minimum 0 - maximum 100). Higher scores reflected better vision, task performance, physical symptoms and confidence and emotions.

AS-20

The AS-20 instructions, questions and scoring method used are described in Appendix J. The AS-20 results were recorded on a spreadsheet. A higher AS-20 score reflected better HRQoL. A mean overall question score (all 20 questions, minimum 0 – maximum 100), a mean psychosocial subscale score (questions 1-10, minimum 0 – maximum 100) and a mean functional subscale score (questions 11-20, minimum 0 – maximum 100) were reported. Additionally, the number of participants reporting a change in AS-20 score (from visit one to visit two) greater than the 95% limits of agreement (Leske et al., 2010) were reported.

VFQ-25

The VFQ-25 instructions, questions and scoring method used are described in Appendix K. The raw results were recorded on a spreadsheet and then converted into recoded VFQ-25 scores. From the recoded VFQ-25 scores, a mean composite VFQ-25 score was calculated and used in the analysis (minimum 0 - maximum 100). A higher VFQ-25 composite score reflected a higher, or better, self-reported visual function. Additionally, the number of participants reporting a change in VFQ-25 (from visit one to visit two) greater than the 95% limits of agreement (Leske et al., 2010) were reported.

8.4.17 Group two study visit measures: accommodation, prehension and balance

Prehension was included as one of the additional measures, using the BIGKAT portable device. Technical difficulties with the BIGKAT meant prehension was removed as one of the additional study measures.

Additionally, some participants completed the group two study visit measures (accommodation and balance). However, the methods are not reported here, and the results are not included in chapter 9. Due to COVID-19 and the lockdown experienced in England, research activity and data collection during phase two had to be paused. Insufficient surgery participants completed both visit one and visit two accommodation and balance measures to allow meaningful comparison to the control group.

8.4.18 Eye movement recording

All participants were offered an optional and additional part of their study visit to take part in the EMR. The Eyelink 1000+ was used with participants placing their chin and forehead on a rest. Stimuli were presented on a screen placed 931mm from the participant. EMR were performed both eyes open, but the Eyelink 1000+ recorded monocularly from the fixing eye. The set-up, calibration and validation are described in Appendix T. Glasses were worn for EMR if they were required to see the target clearly. Data was automatically ‘parsed’ into fixation, a saccade or a blink during recording. Prior to analysis of the recorded data, data was visually inspected in the Eyelink data viewer program to ensure the data was characterised properly.

Fixation

Two practice trials (target in primary position against both the picture and jumbled backgrounds) were shown prior to the ten experimental trials. A 1-degree diameter cross was used as the fixation target (Figure 8-3).

Figure 8-2. The fixation target used during EMR

1-degree (visual angle) diameter black cross with a white centre

Participants were encouraged to find the cross and fixate in the centre as accurately as they could, but they were not told the location of the cross on the screen. The fixation cross was randomly presented in primary position, horizontal gaze (±10°) and vertical gaze (±8°), against a picture background or a jumbled background for 30 seconds (Figure 8-4).

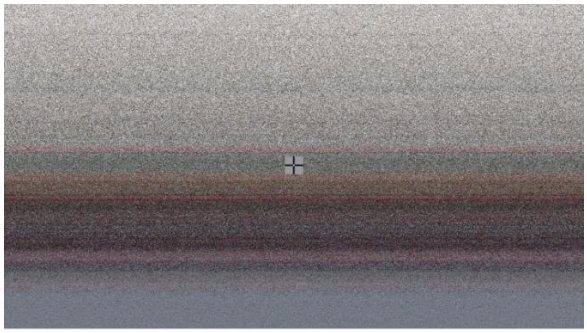


Figure 8-3. The fixation target presented in primary position against picture background 1 and jumbled background 1.

Ten trials were included in total (5 against the picture background and 5 against the jumbled background). Two pictures were used (background 1 and 2) and each background had a jumbled version created by jumbling up the pixels present in the image along horizontal rasters (Figure 8-5). For visit one, half the participants had background 1 and half had background 2. The backgrounds were alternated at visit 2.

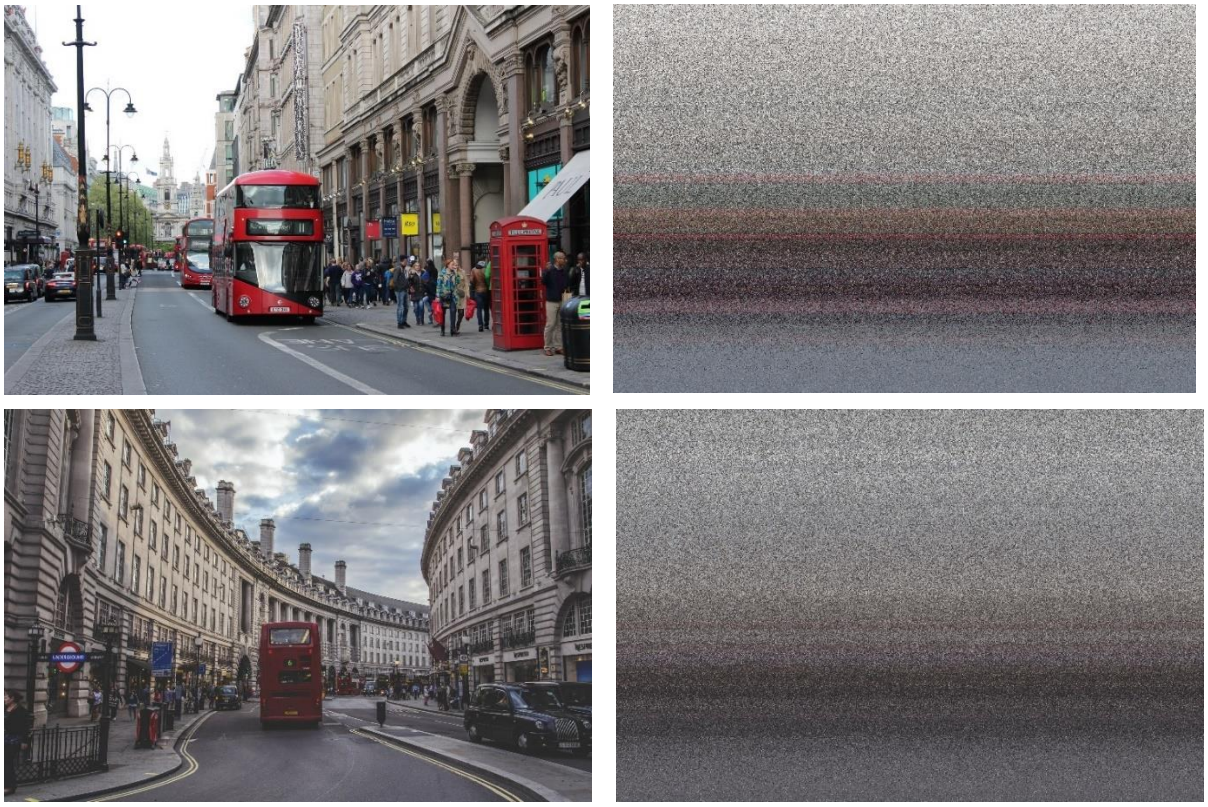


Figure 8-4. Background images used during fixation EMR.

Top left = background 1. Top right = jumbled version of background 1.
Bottom left = background 2. Bottom right = jumbled version of background 2.

Background 1 = Image of a London street scene accessed from: https://pixabay.com/en/london-shopping-england-2268971/ (free for commercial use) (psteph, 2017)
Background 2 = Image of the London street scene https://pxhere.com/en/photo/773176 (free for commercial use) (Unknown photographer, 2017)

Fixation results were analysed by importing circular interest areas (0.5-degree, 1-degree and 2-degrees diameter) around the centre of the target in each of the 5 fixation positions and for each of the backgrounds (Figure 8-6). An interest area report was created to extract the interest area fixation percentage data (%) for each interest area and for any of the fixations falling outside the interest area. This 'interest area fixation %' represented the percentage of all fixations in the trial falling in each of the interest areas or falling outside the 2° interest area, anywhere else on the screen. 0.5° interest area fixation % + 1° interest area fixation % + 2° interest area fixation % + interest area fixation % outside the interest areas = 100% total fixation of the trial.

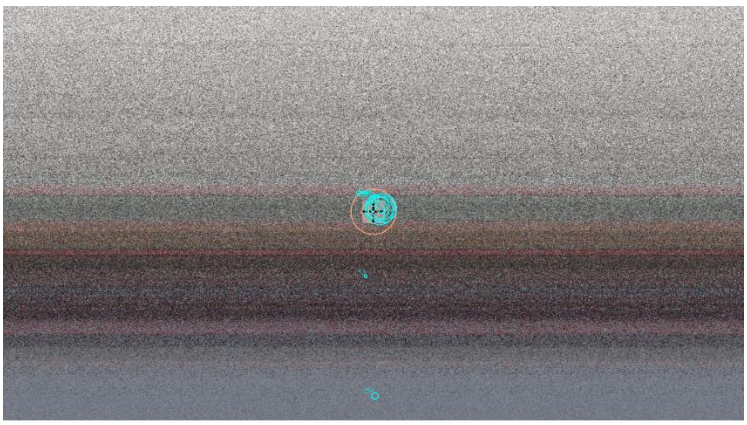


Figure 8-5. Interest areas (orange circles) overlaid on the fixation results for primary position fixation, jumbled background 1.

Interest areas = 0.5°, 1° and 2° diameter circles (orange) with the centre of the circle in the same position as the centre of the fixation cross.
Fixation results appeared as a turquoise circle indicating fixation position on the screen. A larger circle represented a longer duration of fixation.

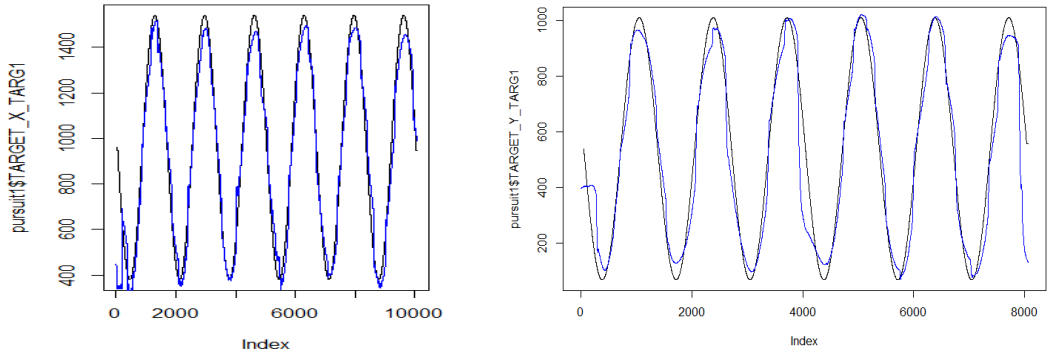
Smooth pursuit

Smooth pursuit was tested horizontally and vertically in separate trials. For each direction a practice trial was given before the experimental trial. A 1-degree diameter cross (Figure 8-3) was used as the target against a grey background. Participants were encouraged to look at the centre of the cross as it moved across the screen. The target was presented as a sinusoidal stimulus moving at mean velocity 24°/second with constant amplitude for six cycles, beginning in the centre. Horizontal smooth pursuit was ±10° amplitude (20° total excursion) with a frequency of 0.6Hz. Trial time = 10,000ms (10 seconds). Vertical smooth pursuit was ±8° amplitude (16° total excursion) with a frequency of 0.75Hz. Trial time = 8,000ms (8 seconds).

Any highly atypical data was inspected, and movement considered non-smooth pursuit was excluded prior to analysis. A sample report was extracted from the Eyelink data viewer program. Pursuit gain, pursuit accuracy and length of fixation for each part of the trial considered a fixation (and not a blink or a saccade) were analysed (Appendix T). Data from the first part of the smooth pursuit movement, with the eye moving from a stationary position to reach the target, was removed prior to calculation of the mean gain, mean accuracy and total fixation.

Smooth pursuit data were velocity gain (eye velocity / target velocity) (value 0-1), accuracy (RMSE in degrees of visual angle) and fixation duration (ms) during the trial. Smooth pursuit velocity gain of 1 equated to the eye moving with equal velocity to the target and a velocity gain of less than one equated to the eye moving with slower velocity than the target. A velocity gain value closer to 1

therefore equated to better smooth pursuit, or a better ability to follow the smooth pursuit target at the appropriate velocity. Smooth pursuit accuracy was a mean measurement of the difference between the eye position and the target position during the trial. A larger accuracy value (RMSE in degrees of visual angle) equated to worse smooth pursuit accuracy. Smooth pursuit fixation duration was a total measurement of the time (ms) the eye was in a 'fixation' during the trial. Longer total fixation during the trial was considered better than shorter total fixation during the trial, as the eye was able to maintain fixation on the target for longer. Example smooth pursuit traces are shown in Figure 8-7.



A. B.

Figure 8-6. Horizontal and vertical smooth pursuit traces

Horizontal smooth pursuit (A) and vertical smooth pursuit (B). Target movement (black) and eye movement (blue)

Saccades

Saccades were tested horizontally (10° amplitude) and vertically (8° amplitude) in separate trials. For each direction practice trials (6 saccades) were given before the experimental trials (n=20). A 1-degree diameter cross (Figure 8-3) was used as the target against a grey background. Participants were encouraged to look at the centre of the cross as accurately as they could before and after it moved position. In the experimental trials, 20 saccades were randomly presented to the left or right and up or down. The cross always started in the centre of screen and randomly moved to a new position horizontally or vertically.

Individual saccade results were inspected prior to analysis to exclude any inaccurate saccade data prior to analysis. Any saccades occurring during or immediately after a blink were excluded. Any saccades inaccurately characterised, with latency <80ms or with peak velocity <100°/second or >1000°/second were removed. Saccades were analysed in the Eyelink data viewer program using a saccade report. Saccade amplitude, peak velocity and latency data were extracted. For horizontal and vertical saccades, mean gain, mean peak velocity and mean latency were

calculated. Saccade gain (saccade amplitude / target amplitude) was a value with minimum 0 and maximum greater than 1. A saccade gain of 1 equated to the eye making a saccade of the same amplitude as the target amplitude. A saccade gain of less than one equated to the eye making a hypometric saccade. A saccade gain of greater than one equated to a hypermetric saccade. Saccade peak velocity was reported in degrees (of visual angle) / second. A higher peak velocity equated to a faster peak velocity. Saccade latency was measured in time (ms). A lower latency equated to a faster saccade.

8.5 Analysis

Analysis of the quantitative data was performed using IBM SPSS Statistics for Windows, Version 25 (SPSS Inc., 2017). The distribution of the data was analysed by considering the distribution of the data in a histogram, the skewness and kurtosis of the data, the Q-Q plots and the Shapiro-Wilk results. The Shapiro-Wilk test of normality was used due to the small sample size. Data that appeared to be taken from a normally distributed population were displayed graphically using the mean (±2 SD) and analysed using parametric analysis. For parametric analysis, Levene's test of equality of error variances was used to test the null hypothesis that the error variance of the dependent variable was equal, or not significantly different, across the groups. Data that appeared to be taken from a non-normally distributed population were displayed graphically using the median (and IQR) and analysed using non-parametric analysis.

8.5.1 Study hypotheses

The overall hypothesis was that there was an effect of strabismus surgery. The null hypothesis was that there was no effect from strabismus surgery.

Direction of an effect was not included in the hypothesis as it was recognised that whilst surgery was intended to have a positive effect (improved eye alignment and HRQoL), outcomes from surgery are not always positive and may be negative. It was also recognised that the additional measures performed may have improved, but they may also have worsened postoperatively, in addition to not changing (or no effect of surgery).

Visit and task performance measures (all measures except the AS-20 and the additional study questions)

For the vision and task performance measures the hypotheses at each visit were as follows:

Visit one

At visit one the hypothesis was that the null hypothesis was true and there was no difference between the control group and the surgery group. The alternative hypothesis was that there was a difference between the control group and the surgery group at visit one.

Visit one to visit two

The hypothesis was that from visit one to visit two there was a change in the surgery group, but there was no change in the control group. The null hypothesis was that there was no change in the surgery group or the control group. The alternative hypothesis was that there was no change in the surgery group, but there was a change in the control group.

Visit two

At visit two the hypothesis was that there was a difference between the control group and the surgery group. The null hypothesis was that there was no difference between the surgery group and the control group.

HRQoL measures (AS-20 and additional study questions)

For the HRQoL measures, the surgery group were expected to have lower HRQoL scores than the control group, the hypotheses at each visit were therefore:

Visit one

At visit one the hypothesis was that there was a difference between the control group and the surgery group, with the surgery group having lower HRQoL than the control group. The null hypothesis was that there was no difference between the control group and the surgery group.

Visit one to visit two

The hypothesis was that from visit one to visit two there was a change in the surgery group, but there was no change in the control group. The null hypothesis was that there was no change in the surgery group or the control group. The alternative hypothesis was that there was no change in the surgery group, but there was a change in the control group.

Visit two

At visit two the hypothesis was that there was no difference between the control group and the surgery group. The null hypothesis was that there was a difference between the surgery group and the control group.

The two independent groups were the surgery group and the control group. The statistical analysis of the quantitative data compared the results from visit one between groups and the results from visit two between groups. The results from visit one and visit two were compared within groups. The within groups and between groups comparisons were given equal consideration when interpreting change from visit one to visit two. Firstly, for all analyses, the within groups change in the surgery group was compared to the control group to inform the decision on whether the surgery group results differed to those seen in the control group. This allowed interpretation of whether any change in performance in the surgery group differed from the expected practice effect in a similar population. Secondly, for all analyses, the between groups comparisons, at both visit one and visit two, were used to inform the decision on whether the surgery group performed better than, worse than, or no different to the control group at each visit.

Results were displayed graphically using red text to denote between groups analysis at visit one, blue text to denote between groups analysis at visit two and within groups analysis in black text. Statistical significance was shown on graphs using the notation (NS $p>0.05$, * $p\leq0.05$, ** $p\leq0.01$, *** $p\leq0.001$, **** $p\leq0.0001$).

Parametric analysis was predominantly a two-way mixed analysis of variance (ANOVA) to analyse the main effect of group, the main effect of visit and a two-way interaction between group and visit on the dependent variable. The between subjects factor was group and the within subjects factor was visit. The dependent variable was the measurement of performance of the task. Analysis of covariance (ANCOVA) was also performed using the covariate baseline size of distance deviation (PD), to investigate the effect of the size of the deviation at visit one and a possible interaction with group. Additional within subjects factors were included in the ANOVA where appropriate and are described in chapter 9. Planned comparisons were performed between groups. Where post-hoc t-tests were performed to explore within groups results, Bonferroni correction was used to try and control the familywise error rate.

Non-parametric analyses used the Independent Samples Mann-Whitney U Test to compare between the groups at visit one and at visit two. The Wilcoxon signed rank test was used to compare repeated measures within the groups, from visit one to visit two.

The level of significance was accepted as $p\leq0.05$. P values that rounded down to 0.05 were not considered significant (example $p=0.054$). Effect size was calculated. $r=0.1$ was considered a small effect, $r=0.3$ a medium effect and $r=0.5$ a large effect size. A detailed description of the analysis steps was included in chapter 9 the first time each analysis was used. Subsequent analyses then followed the same steps.

8.5.2 Postoperative judgement of success, partial success and failure

Postoperative outcomes were categorised as success, partial success or failure, using the same criteria as Hatt et al. (2012a) except the APCT was used rather than the SPCT for the measurement of the deviation (summarised in Table 8-2). This definition of success, partial success and failure was selected as it was the most commonly used and cited by other authors (section 3.2.3).

Table 8-1. Criteria used to define postoperative success, partial success or failure at visit two		
	(Hatt et al., 2012a)	Phase two
Success	Heterotropia <10PD (primary position, in the Distance and Near) measured by SPCT Diplopia or visual confusion = none or rare, primary position at distance and for reading No prism, Bangerter foil or occlusion	Heterotropia <10PD (primary position, Distance and Near) measured by APCT Heterophoria Diplopia or visual confusion: none or rarely reported by patient, in primary position at distance and for reading No prism, Bangerter foil or occlusion
Partial success	Heterotropia ≤15PD (primary position, in the Distance and Near) measured by SPCT Diplopia or visual confusion = none / rare / sometimes, primary position at distance and for reading (with or without prism) Prism (yes / no) No Bangerter foil or occlusion	Heterotropia ≤15PD (primary position, in the Distance and Near) measured by APCT Diplopia or visual confusion: none / rare / sometimes reported by patient, in primary position at distance and for reading (with or without prism) Prism (yes / no) No Bangerter foil or occlusion
Failure	Heterotropia >15PD (primary position, in the Distance and Near) measured by SPCT Diplopia or visual confusion = always / often, primary position at distance and for reading Prism, Bangerter foil or occlusion (yes / no)	Heterotropia >15PD (primary position, in the Distance and Near) measured by APCT Diplopia or visual confusion: always / often reported by patient, in primary position at distance and for reading Prism, Bangerter foil or occlusion (yes / no)

For a successful or partially successful outcome, all criteria had to be met. If one of the 'failure' criteria were met, then the outcome was classed as a failure. If the criteria for more than one outcome group was met (for example, meeting both the success and partial success outcomes), the patient was allocated to the group with the better outcome (Hatt et al., 2012a).

8.6 Chapter 8 summary

This chapter has described the methods used in the quantitative phase of the study. The two independent groups were the surgery group and the control group. All participants invited to participate in the study were adults with constant strabismus and suppression, with no diplopia and no BSV. The quantitative methods were repeated at two separate visits. For the surgery group this was preoperatively and at least three months postoperatively. For the control group this was at two study visits at least three months apart. The quantitative methods included measures of vision, task performance, physical symptoms and confidence and emotions. In chapter 9 the quantitative results are presented.

Chapter 9. Quantitative Results

In the previous chapter 8 the methods used in the quantitative phase two and the hypotheses for phase two were described. In this chapter 9 the results from the different measures included in the quantitative phase two will be reported. Firstly, descriptive statistics of the quantitative cohort are presented (section 9.1). The analyses of deviation size are presented (section 9.2), followed by analyses of the PROMs (section 9.3), which include questions relating to physical symptoms and confidence and emotions. Vision measures are then presented (section 9.4), followed by task performance measures (section 9.5) and EMR (section 9.6). The key quantitative findings are summarised (section 9.7) and presented as the measures that have improved, worsened or remained unchanged after surgery.

9.1 Descriptive Statistics

Thirty-five participants were recruited to phase two of the study and 27 completed both study visits. The measures completed by each of the participants (n=27) (visit one and visit two) are summarised in Appendix U. Twenty participants were recruited to the surgery group and twelve completed both study visits. Due to covid-19 restrictions eight participants were unable to attend visit two and are not reported in this thesis. All surgery group participants underwent surgery to improve their eye alignment and psychosocial symptoms. Fifteen participants were recruited to the control group, all completed both study visits. No strabismus treatment was received by the control group during the study. The clinical demographics of all twenty-seven participants are shown in Table 9-1.

The surgery group (n=12) all had an XT preoperatively. The control group included participants with both XT (n=5) and ET (n=10). Amblyopia was present in nine of the twelve participants in the surgery group and in seven of the 15 participants in the control group. Previous strabismus surgery was reported in eight of the twelve participants in the surgery group and in 13 of the 15 participants in the control group. Two members of the control group had no previous strabismus treatment. Two of the control group also reported receiving BT injections as part of their previous strabismus treatment. Participant 24 had received BT at age 11 and 12 years and participant 8 had received BT intermittently in adulthood when their residual angle secondary ET increased due to experiencing psychosocial symptoms.

Table 9-1. Clinical demographics of phase two participants							
Group	Participant number	Gender	Visit 1 strabismus	Visit 2 strabismus	Strabismus onset	Previous treatment for strabismus	Amblyopia
Surgery	01	F	Consecutive XT	Consecutive residual ET	ET < 2 years	Sx age 17 for ET	Strabismic amblyopia
	03	M	Consecutive XT	Consecutive residual ET with accommodative element	ET < 6 months	Nil, spontaneous consecutive XT drift	Strabismic amblyopia
	05	F	Consecutive XT, DVD	Consecutive residual ET, DVD	ET < 6 months	Sx age 18/12 for ET	Nil
	07	F	Consecutive XT	Consecutive residual ET	ET < 4 years	Sx for ET < 4 years	Strabismic amblyopia
	09	M	Primary XT & HT	Consecutive ET & HT	XT in childhood	Nil	Nil
	10	F	Consecutive XT	Consecutive residual ET with accommodative element	ET < 4 years	Sx for ET as a child	Strabismic amblyopia
	13	F	Consecutive XT	Residual consecutive XT	ET < 4 years	Sx age 5 & age 21 for ET	Strabismic & anisometropic amblyopia
	15	M	Residual XT	Residual X	XT in childhood	Sx age 11 for XT	Nil
	16	M	Consecutive XT & HT	Residual consecutive XT	ET age 5	Sx age 10 & age 16 for ET	Strabismic amblyopia
	18	F	Consecutive XT	Residual consecutive XT	ET in childhood, consecutive XT noticed after head injury 2019	Nil	Strabismic amblyopia
	20	F	Consecutive XT	Residual consecutive XT	ET age 15 months	Nil, spontaneous consecutive XT drift	Strabismic amblyopia
	21	M	Consecutive XT	Residual consecutive XT	ET < 2 years	Sx age 2 for ET	Strabismic & anisometropic amblyopia
Control	2	F	Residual ET	Residual ET	ET from birth	Sx age 3 for ET	Nil
	4	F	Consecutive XT	Consecutive XT	ET from birth	Sx age 4 for ET	Nil
	6	F	Consecutive XT, DVD	Consecutive XT, DVD	ET < 6 months	Sx age 6/12 for ET Sx age 2, 5 & 16 for XT	Strabismic amblyopia
	8	F	Secondary ET & HT	Secondary ET	ET from birth	Sx age 20 for ET Occasional BT for ET	Stimulus deprivation & strabismic amblyopia
	11	F	Primary ET	Primary ET	ET age 4	Nil	Strabismic amblyopia
	12	F	Consecutive ET	Consecutive ET	XT < 4 years	Sx age 19 for XT	Strabismic amblyopia
	14	M	Residual ET & HT	Residual ET & HT	ET from birth	Sx age 9/12 & 9 years for ET	Strabismic amblyopia, HSK in fixing eye

	17	F	Residual ET, DVD	Residual ET, DVD	ET < 6 months	Sx age 1 for ET	Strabismic amblyopia
	22	F	Residual ET Nr, Consecutive XT Dist	Residual ET Nr, Consecutive XT Dist	ET age 1	Sx age 2 for ET	Nil
	23	F	Residual ET & HoT, DVD	Residual ET & HoT, DVD	ET from birth	Sx age 8/12 for ET	Nil
	24	F	Consecutive XT & HT	Consecutive XT & HT	ET from birth	Sx age 2 for ET Sx age 5 for XT BT age 11 & 12 for XT Sx age 14 for XT	Nil JIA and intermittent uveitis
	25	F	Residual ET	Residual ET	ET < 2 years	Sx age 7 for ET	Nil. Refractive Sx for myopia age 41
	27	F	Infantile ET with accommodative element, DVD & MLN	Infantile ET with accommodative element, DVD & MLN	ET in childhood	Nil	Nil
	29	F	Residual XT	Residual XT	XT from birth	Sx age 4 for XT	Strabismic & anisometropic amblyopia
	31	F	Consecutive XT	Consecutive XT	ET < 2 years	Sx age 2 for ET	Nil

Key to table:

M: male, F: female.

Consecutive: strabismus has changed direction, typically following surgery, example from ET to XT or from XT to ET.

Residual: strabismus remains in the same direction following surgery, but is typically a smaller angle, example from large XT to smaller XT.

Infantile ET: ET of onset less than 6 months with features of MLN and DVD.

Consecutive residual ET: original strabismus was an ET, then became consecutively XT, during study had Sx for XT and became consecutively ET but with a residual ET

Residual consecutive XT: original strabismus was an ET, then became consecutively XT, during study had Sx for XT and becomes a smaller XT

Sx: surgery, JIA: juvenile idiopathic arthritis, HSK: herpes simplex keratitis.

The surgical procedures performed on the surgery group and the surgical outcomes are presented in Appendix V. Postoperative outcomes in the surgery group included suppression (n=8), suppression with the aid of a prism in the distance (n=1) and unexpected BSV (n=3). Unexpected BSV was measured in participants 03, 10 and 15. Of these three, participant 15 achieved the best BSV, with evidence of sensory fusion, motor fusion and stereopsis (110” of arc, FNS) postoperatively. Participants 03 and 05 achieved gross BSV at near only. Participant 03 achieved sensory fusion at near (Bagolini glasses) and an intermittent stereopsis response when looking at the Wirt fly (3552” of arc). Participant 10 achieved sensory fusion at near only (Bagolini glasses), but no stereopsis. Participant 05 was included in the suppression outcome, as they described mostly suppression, with an occasional fleeting cross or diplopia response (Bagolini glasses) at near.

Participant 07 preoperatively had an XT (PCT near 40PD XT 4PD LHT, distance 25PD XT 4PD LHT) and had a postoperative overcorrection of their deviation. One week postoperatively they had an ET and a -3 limitation of abduction (PCT near 8PD ET 6PD LHT, distance 25PD ET 8PD LHT). This limitation and the consecutive ET gradually reduced in size postoperatively. Initially blenderm occlusion was worn. Three months postoperatively a 12PD base out Fresnel prism was fitted to place the diplopic image into the patient’s suppression area. This Fresnel prism was then reduced in size over a further 3 months, as the deviation gradually relaxed out to a smaller ET without diplopia (with suppression). Six months postoperatively (when study visit two was conducted) participant 07 had a small consecutive residual ET with suppression (PCT near 8PD ET, distance 20PD ET 4PD LHT) with a -1.5 limitation of abduction. Despite having suppression, they preferred to wear a 10PD base out Fresnel prism in the distance to more comfortably achieve suppression. This prism was worn for both eyes open testing during visit two.

9.1.1 Timing of visits

The mean (SD) time between visit one and visit two (days) was calculated for both groups, shown in Figure 9-1. These data were normally distributed.

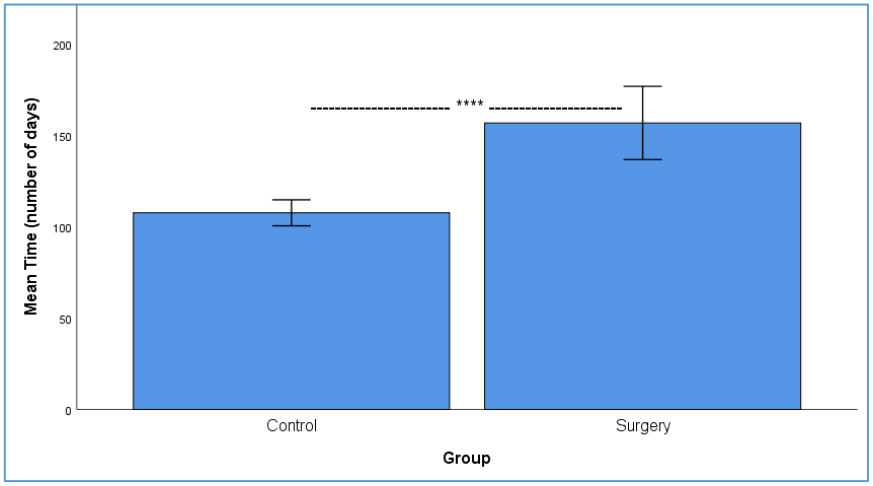


Figure 9-1. Bar chart showing the mean (±2 SD) time (days) between visit one and visit two for both groups

**** p≤0.0001

The surgery group had a significantly longer time between visit one and visit two (M=156.7 days, SD=34.7) compared to the control group (M=107.5 days, SD=13.7), t(13.8)=4.62, p<0.0001, r=0.78 (independent samples t-test).

9.2 Deviation

The surgery group all had XT, with six of the twelve participants additionally having a smaller vertical element to their strabismus. All participants in the surgery group underwent surgery to reduce the size of their XT. Participant 5 had surgery (left superior rectus (SR) recession) to additionally reduce their dissociated vertical deviation (DVD).

The control group had ET (n=9) and XT (n=5). Participant number 27 was included in the control group despite having no measurable deviation at near or distance with refractive correction. They had alternating DVD, which spontaneously elevated throughout testing. Without refractive correction they had a small constant ET. They had no evidence of BSV, had constant suppression and gave a history of having an infantile ET with an accommodative element since childhood. They had received no previous treatment for their ET, other than refractive correction.

9.2.1 Prism cover test (PCT) measurements

The primary position, horizontal and vertical deviation at visit one and visit two (absolute values of the alternate PCT (PD)) are shown in Figure 9-2. The surgery group results are shown on the upper graph and the control group on the lower graph. There appeared to be little difference in the size of the vertical deviation between the two groups and from visit one to visit two, for both groups. The horizontal deviation, at near and distance, appeared larger in the surgery group than

the control group at visit one. From visit one to visit two, the horizontal deviation in the surgery group reduced and minimal change was evident in the control group. At visit two, the horizontal deviation, at near and distance, in the surgery group appeared smaller than the control group.

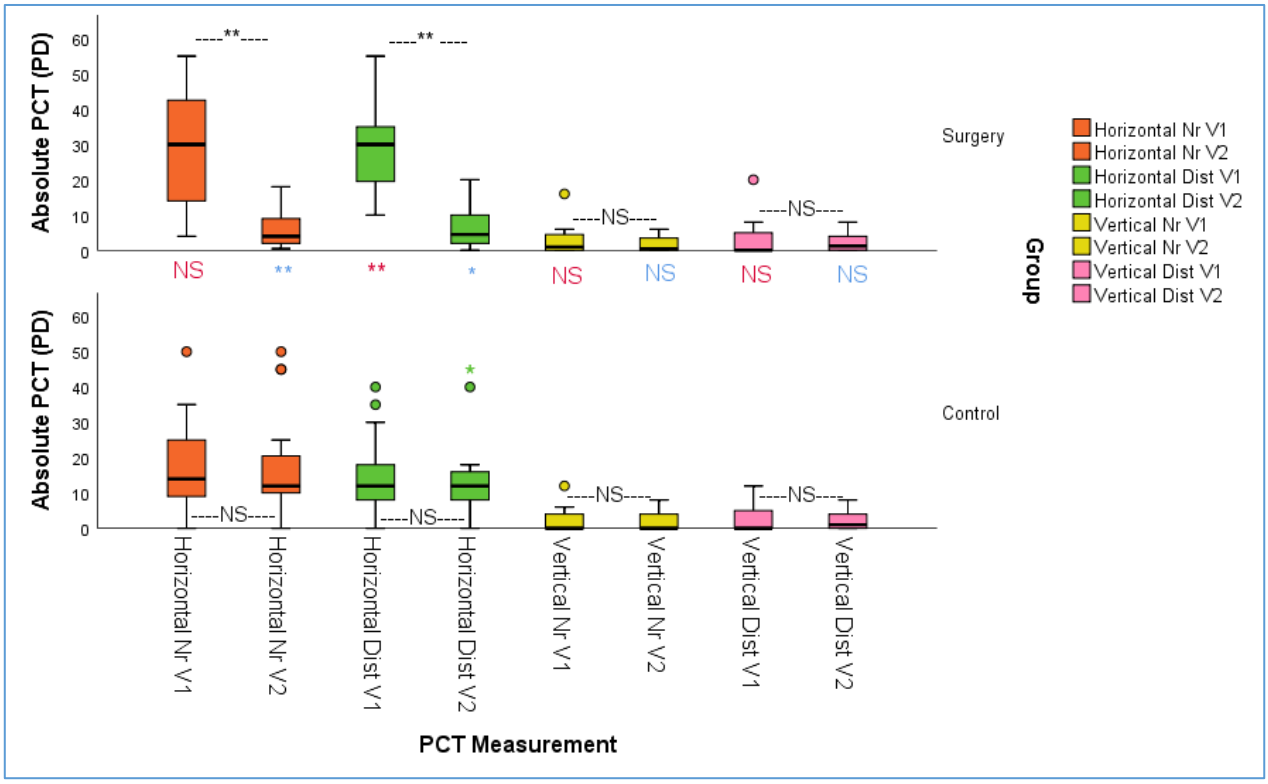


Figure 9-2. Boxplot of the horizontal and vertical primary position PCT results, at near and distance, at visit one and visit two, for both groups.

Absolute values of the PCT (PD) are shown. Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-

Non-parametric analysis of PCT

The PCT data was not normally distributed. The steps in the non-parametric analysis are described here, for later non-parametric analysis the same steps will be followed.

Visit one

The Independent Samples Mann-Whitney U test was used to test the hypothesis that there was no difference between the groups at visit one (section 8.5.1). These between groups results are shown in Figure 9-2 in red text. At visit one PCT measurements in the surgery group did not significantly differ from the control group, for the horizontal deviation at near (surgery Mdn=30, control Mdn=14), vertical deviation at near (surgery Mdn=1, control Mdn=0), and the vertical deviation at distance (surgery Mdn=0, control Mdn=0)). However, PCT measurements, distance

horizontal deviation, in the surgery group (Mdn=30) were significantly larger than the control group (Mdn=12), U=142.5, z=2.57, p<0.01, r=0.495.

Visit one to visit two

The Wilcoxon signed rank test was used to test the hypothesis that from visit one to visit two there was a change in the surgery group and no change in the control group (section 8.5.1). Repeated measures of the PCT at visit one and visit two were compared for both groups. These within groups results are shown in Figure 9-2 in black text. There was no significant change in the PCT in the control group at near or distance (p>0.05). The surgery group had significantly smaller horizontal deviations postoperatively at near (visit one (Mdn=30), visit two (Mdn=4), T=0, p=0.003, r=0.60) and at distance (visit one (Mdn=30), visit two (Mdn=4.5), T=0, p=0.002, r=0.62).

Visit two

The Independent Samples Mann-Whitney U Test was used to test the hypothesis that there was a difference between the groups at visit two (section 8.5.1). These between group results at visit two are shown in Figure 9-2 in blue text. At visit two, PCT measurements in the surgery group were significantly smaller than the control group for the near horizontal PCT (surgery group Mdn=4, control group Mdn=12), U=33, z=2.793, p=0.004, r=0.51; and for the distance horizontal PCT (surgery group Mdn=4.5, control group Mdn=12) U=46, z=2.154, p=0.03, r=0.44. The vertical PCT, near and distance, were not significantly different in the surgery group (near Mdn=0.5, distance Mdn=0) and the control group (near Mdn=0, distance Mdn=0) at visit two, p>0.05.

9.2.2 Change in deviation size

The change in horizontal deviation (PD) was calculated by subtracting the visit one PCT from the visit two PCT for both the near and the distance deviation. Positive values were assigned to improved (reduced) deviations and negative values were assigned to worsened (increased) deviations. The change in primary position horizontal deviation PCT (PD) in both groups are shown in Figure 9-3. The surgery group showed a reduction in PCT at near and distance (positive change), whereas the control group had a slight increase in PCT (negative change).

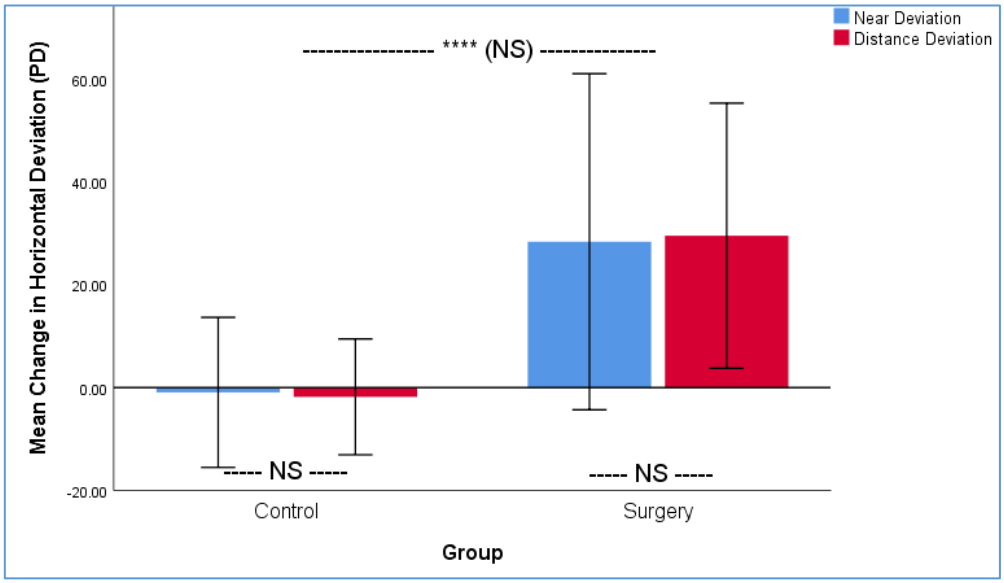


Figure 9-3. Bar chart showing the mean (+/- 2SD) change in horizontal deviation (PD), from visit one to visit two, for both groups

Improved deviations shown as a positive value. Worsened deviations shown as a negative value.

Statistical significance reported as NS = not significantly different ($p>0.05$), **** = significant ($p\leq0.0001$). The two-way mixed ANOVA results are shown and the ANCOVA results are shown in brackets. The difference within the groups is shown in black text underneath the bars. The difference between the groups is shown in black text above the bars.

Parametric analysis of the change in horizontal deviation

The change in deviation size data were normally distributed. The steps in the parametric analysis are described in detail below, for later parametric analysis the same steps will be followed.

Firstly, a two-way mixed ANOVA was performed using the change in horizontal deviation as the dependent variable. The between subjects factor was group (surgery or control group). The within subjects factor was test distance (near and distance). Secondly, analysis of the change in the horizontal deviation was repeated with a covariate (ANCOVA) to analyse whether the covariate was a significant factor. The covariate used was the absolute value of the size of the distance horizontal deviation at visit one (baseline size of distance deviation (PD)).

Change in the horizontal deviation – two-way mixed ANOVA

Levene's test was significant, $F(1,25)=15.51$, $p<0.001$, indicating that the error variance in the two groups was significantly different and the assumption (required for the ANOVA) had not been met. However, the ANOVA is reported here for comparison to the ANCOVA result below. The main effect of group was significant, $F(1,25)=54.57$, $p<0.0001$, $r=0.83$, supporting a significant difference between the change in deviation in the control group ($M=-1.37$) and the surgery group ($M=29.96$)

(shown in Figure 9-3 in black text above the bars). The main effect of test distance was not significant, F(1,25)=0.02, p=0.88, supporting no significant difference between the change in deviation at near and distance (shown in Figure 9-3 in black text underneath the bars). The two-way interaction, group x test distance, was not significant, F(1,25)=1.02, p=0.32, supporting no significant interaction.

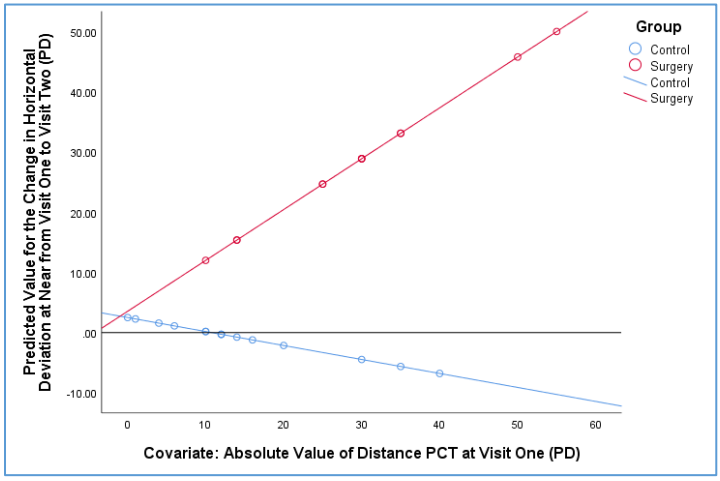
The surgery group had a large improvement in their deviation after surgery, and the control group had a slight worsening of their deviation on repeated measurement. The change in deviation at near was not significantly different to the change in deviation at distance. There was no significant difference in the change in deviation at the different test distances in the different groups.

Change in the horizontal deviation - ANCOVA

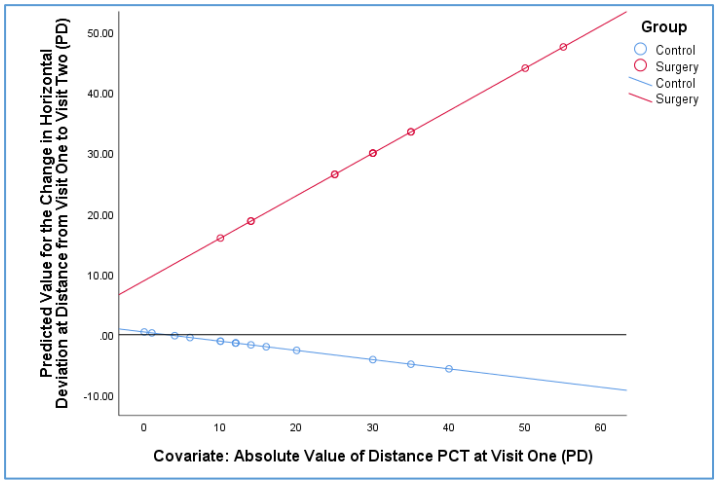
With the covariate (baseline size of distance deviation (PD)) included, Levene's test was no longer significant for the change in the near deviation, F(1,25)=3.60, p=0.07, or the change in the distance deviation, F(1,25)=1.72, p=0.20, indicating the error variance in the two groups was no longer significantly different. This suggested some of the unexplained variance could be explained by the covariate and the error variance had reduced.

The main effect of group was not significant, F(1,23)=0.50, p=0.49, r=0.15, indicating that overall there was no significant difference in the change in deviation between the groups (shown in Figure 9-3 in black text, in brackets). The main effect of test distance was not significant, F(1,23)=0.60, p=0.45, indicating there was no significant difference between the change in the deviation at near and the change in the deviation at distance. The effect of the covariate (size of horizontal deviation, visit one) was significant, F(1,23)=5.10, p<0.05, r=0.43, indicating that those with larger horizontal deviations at visit one had larger changes in the horizontal deviation from visit one to visit two.

There was a significant two-way interaction between group x covariate (size of horizontal deviation at visit one), F(1,23)=14.08, p<0.001, r=0.62, indicating that having surgery and having a larger horizontal deviation at visit one, led to a significantly larger change in the horizontal deviation at visit two. The interaction is shown in Figure 9-4, where the change in the near (A) and distance (B) horizontal deviations can be seen for the size of the baseline deviation in the surgery and control groups.



A



B

Figure 9-4. Scatter plot showing the interaction effect of the covariate on the change in horizontal deviation, from visit one to visit two, at near (A) and distance (B), for both groups

Surgery groups shown in red and control group shown in blue.

There were no significant two-way interactions between test distance x group, $F(1,23)=2.98$, $p=0.10$, or test distance x covariate, $F(1,23)=0.15$, $p=0.70$. The change in deviation at near and distance remained not significantly different in the different groups or in those with different sized deviations at visit one. The three-way interaction test distance x group x covariate was not significant, $F(1,23)=1.85$, $p=0.19$, indicating that having surgery and having a larger deviation at visit one was no longer a significant interaction when test distance was included. Having surgery and having a larger deviation at visit one did not lead to a significant difference in the change in deviation at near and distance.

With the covariate included in the model, there was no main effect of group and no main effect of test distance. There was a main effect of the covariate, those with larger deviations at visit one had larger changes in their deviation from visit one to visit two. There was no interaction between group and test distance. There was an interaction for group x covariate, indicating that those who had

surgery had a larger change in their deviation if they had a larger deviation at baseline. The interactions test distance x group and test distance x group x covariate were not significant, as the change in deviation at near and distance was not significantly different, even when it interacted with group or group x covariate.

9.2.3 Postoperative judgement of success, partial success and failure

Seven (7/12) outcomes were a success, four (4/12) were a partial success and one (1/12) was a failure postoperatively (Table 9-2).

Table 9-2. Postoperative outcomes for each of the surgery group participants using the success, partial success and failure criteria.

Participant number	Phase two postoperative outcome (success / partial success / failure)			Phase two postoperative outcome given (success / partial success / failure)	Reason for partial success or failure
	Deviation size	Diplopia	Prisms / Foil / Occlusion		
1				Success	
3				Partial success	PCT Nr 10ET
5				Partial success	Diplopia sometimes (when fixing with affected eye)
7				Failure	PCT: Dist 20ET 4LHT, Prisms to put back into suppression area, intermittent diplopia without prisms
9				Partial success	PCT Dist 10ET 8RHT
10				Success	
13				Success	
15				Success	
16				Success	
18				Partial success	PCT Nr & Dist 10XT 2LHT
20				Success	
21				Success	

Foil = Bangerter foil. Green = success, orange = partial success, red = failure. Nr = near, Dist = distance.

Summary of the deviation results

At visit one participants in both groups had similar deviation sizes, although the surgery group had significantly larger distance horizontal deviations (Figure 9-2). The vertical deviations in both groups were small and were not significantly different. Surgery led to a significant reduction in the horizontal deviation at near and distance, but not the vertical deviation, as expected (Figure 9-2). At visit two the surgery group had significantly smaller horizontal deviations than the control group (Figure 9-2). The vertical deviations remained small and not significantly different between groups. The analysis of the change in the horizontal deviation, from visit one to visit two, showed that the covariate, size of horizontal deviation at visit one, was significant and also interacted with group (Figure 9-4). Those who had surgery and had a larger horizontal deviation at visit one, had a significantly larger change in the size of their horizontal deviation (from visit one to visit two), and the change in the deviation was similar at near and distance, which supported the hypothesis.

9.3 PROMs

All participants (n=27) completed the three different questionnaires (AS-20 (Appendix J), VFQ-25 (Appendix K) and the additional study questions (Appendix L)) at both visits. The PROMs results are firstly presented for each questionnaire (sections 9.3.1, 9.3.2 and 9.3.3) and then secondly presented by theme (vision, task performance, physical symptoms and confidence and emotions) (section 9.3.4). The additional study questions results are additionally presented by question as this questionnaire was developed specifically for the study and the results would inform the later refinement of the questionnaire for future study (section 9.3.3).

9.3.1 AS-20 questionnaire

The AS-20 results are presented (0-100) in Figure 9-5 with a lower score representing worse HRQoL. The median and IQR overall, psychosocial subscale and function subscale scores for the surgery and the control group at visit one and visit two are shown. The surgery group appeared to have lower scores than the control group in all of the AS-20 measures at visit one. From visit one to visit two the control group had slightly higher AS-20 scores and the surgery group had much higher AS-20 scores. At visit two the AS-20 scores in the surgery group appeared higher than the control group.

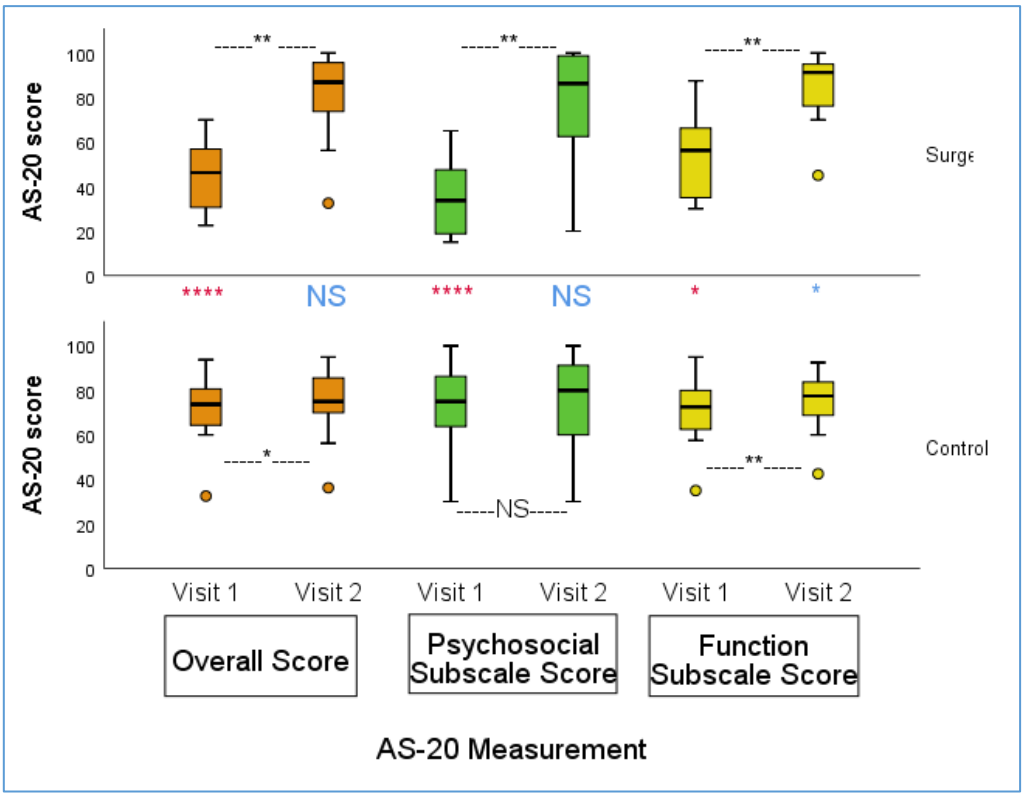


Figure 9-5. Boxplot showing the AS-20 overall and subscale scores, at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001), **** = significant (p≤0.0001).
Text in black = within groups analysis (Wilcoxon signed rank test).
Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test)

Non-parametric analysis of AS-20 scores

Visit one

The visit one between group results are shown in Figure 9-5 in red text (Independent Samples Mann-Whitney U test). At visit one the surgery group had significantly lower AS-20 scores than the control group. The AS-20 overall score in the surgery group (Mdn=46.25) was significantly lower than the control group (Mdn=73.75), U=16.5, z=3.59, p<0.0001, r=0.69. The AS-20 psychosocial subscale score in the surgery group (Mdn=33.75) was significantly lower than the control group (Mdn=75), U=14.5, z=3.69, p<0.0001, r=0.71. The AS-20 function subscale score in the surgery group (Mdn=56.25) was significantly lower than the control group (Mdn=72.50), U=40.5, z=2.42, p=0.014, r=0.47

Visit one to visit two

The visit one to visit two within group results are shown in Figure 9-5 in black text (Wilcoxon signed rank test).

Overall AS-20 scores

For the control group, the overall AS-20 scores were slightly higher on visit two (Mdn=75.00) than visit one (Mdn=73.75), and this difference (1.25) was statistically significant, T=84, p=0.048, r=0.36. For the surgery group, the overall AS-20 scores were much higher on visit two (Mdn=86.87) than visit one (Mdn=46.25), and this difference (40.62) was statistically significant, T=78, p=0.002, r=0.62.

AS-20 Psychosocial Subscale Scores

For the control group, the AS-20 psychosocial subscale scores were slightly higher at visit two (Mdn=80) compared to visit one (Mdn=75), but this difference (5) was not significant, T=55.50, p=0.19, r=0.24. For the surgery group, the AS-20 psychosocial subscale scores were higher on visit two (Mdn=86.25) than visit one (Mdn=33.75). This difference (52.5) was statistically significant, T=78, p=0.002, r=0.62.

AS-20 Function Subscale Scores

For the control group, the AS-20 function subscale scores were slightly higher on visit two (Mdn=77.50) than visit one (Mdn=72.50), and this difference (5) was significant, T=80.50, p=0.01, r=0.45. For the surgery group, the AS-20 function subscale scores were markedly higher at visit two (Mdn=91.25) than visit one (Mdn=56.25), and this difference (35) was significant, T=78.00, p=0.002, r=0.62.

Visit two

The between group results are shown in Figure 9-5 in blue text (Independent samples Mann-Whitney U test). The AS-20 overall scores at visit two were higher in the surgery group (Mdn=86.87) than the control group (Mdn=75), but were not significantly different, U=117.5, z=1.34, p=0.18, r =0.26. The AS-20 psychosocial subscale scores at visit two were higher in the surgery group (Mdn=86.25) than the control group (Mdn=80), but again they were not significantly different, U=100.5, z=0.51, p=0.61, r=0.10. The AS-20 function subscale scores at visit two were significantly higher in the surgery group (Mdn=91.25) than the control group (Mdn=77.50), U=133.5, z=2.13, p=0.03, r=0.41.

Limits of agreement

The change in AS-20 score from visit one to visit two and the number of participants reporting a change in AS-20 score greater than the 95% limits of agreement is shown in Table 9-3.

Table 9-3. Change in AS-20 scores for both groups and the number achieving a change in AS-20 score greater than the 95% limits of agreement			
Change in AS-20 score (visit two – visit one)		Control group (n=15)	Surgery group (n=12)
Overall AS-20 score	Median (IQR)	3.75 (6.25)	42.50 (33.13)
	Number achieving >95% LoA (>14.3)	0/15	10/12
AS-20 psychosocial subscale score	Median (IQR)	2.50 (10.0)	47.50 (55.0)
	Number achieving >95% LoA (>17.7)	0/15	9/12
AS-20 function subscale score	Median (IQR)	2.50 (7.50)	32.50 (30.0)
	Number achieving >95% LoA (>19.5)	0/15	8/12
LoA limits of agreement (taken from Leske et al, 2010) IQR interquartile range			

Summary of AS-20 results

All the AS-20 scores (overall score, psychosocial and function subscale scores) were significantly lower in the surgery group than the control group at visit one (Figure 9-5, red text). The surgery group therefore self-reported significantly worse HRQoL prior to surgery than the control group using the AS-20 PROM. All of the AS-20 scores increased significantly, in both groups, at visit two compared to visit one, except the psychosocial subscale score, which did not improve in the control group (Figure 9-5, black text). The increase in AS-20 scores were much greater in the surgery group than the control group. After surgery, all the AS-20 scores had increased in the surgery group to higher than the control group, although only the function subscale score was significantly higher in the surgery group than the control group at visit two (Figure 9-5, blue text). Self-reported HRQoL after strabismus surgery, using the AS-20 PROM, was significantly better than before surgery and increased to a similar or significantly better HRQoL than the control group. These results supported the hypotheses of an effect of surgery on HRQoL.

9.3.2 VFQ-25

The median (IQR) VFQ25 composite scores in both groups at visit one and visit two are shown in Figure 9-6. A lower score indicates worse visual function. The surgery group appeared to have lower scores than the control group at visit one. From visit one to visit two the surgery group appeared to have higher scores and the control group appeared to have similar scores. At visit two the surgery group appeared to have higher scores than the control group, close to the maximum score of 100.

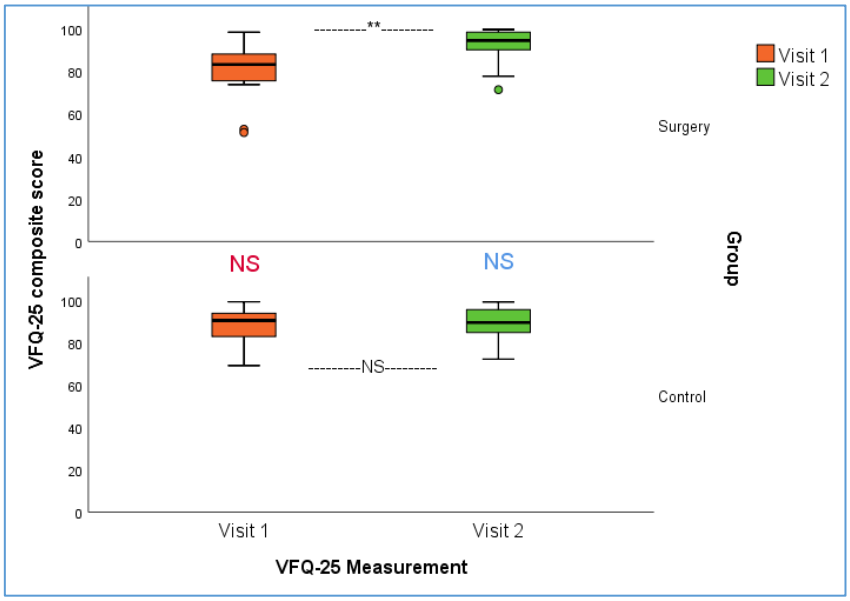


Figure 9-6. Boxplot showing the VFQ-25 composite scores at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01). Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test)

Non-parametric analysis of VFQ-25 scores

At visit one the VFQ-25 score in the control group (Mdn = 90.42) was not significantly different from the surgery group (Mdn=83.49), U=59, z=1.51, p=0.14, r =0.29. Shown in red in Figure 9-6. The within groups results from visit one to visit two are shown in black in Figure 9-6. In the surgery group the VFQ25 score at visit one (Mdn=83.49) was significantly higher at visit two (Mdn=94.81), T=78, p=0.002, r=0.62. In the control group, there was no significant difference between the VFQ25 composite score at visit one (Mdn=90.42) and visit two (Mdn=89.47), T=84, p=0.17, r=0.25. At visit two the VFQ-25 score in the surgery group (Mdn=94.81) was higher than the control group (Mdn=89.47) but the difference was not statistically significant, U=115.5, z=1.25, p=0.22, r=0.24. Shown in blue in Figure 9-6.

Limits of agreement

The change in VFQ-25 scores and the number of participants reporting a change in VFQ-25 score greater than 95% limits of agreement are shown in Table 9-4. None of the control group (0/15) but seven of the surgery group (7/12) reported improvements in VFQ-25 score greater than 95% limits of agreement.

Table 9-4. Change in VFQ-25 score for both groups and the number achieving a change in VFQ-25 score greater than the 95% limits of agreement		
Change in VFQ-25 overall score (visit two – visit one)	Control group (n=15)	Surgery group (n=12)
Median (IQR)	0.63 (5.24)	13.43 (14.97)
Number achieving >95% LoA (>11.1)	0/15	7/12
LoA limits of agreement (taken from Leske et al, 2010) IQR interquartile range		

Summary of the VFQ-25 results

The VFQ-25 composite score was not significantly different between the control group and the surgery group at visit one. Only the surgery group had a significant increase in their VFQ-25 score at visit two compared to visit one. Despite the significant increase in VFQ-25 score in the surgery group, the visit two VFQ-25 scores in the surgery and control groups were not significantly different (Figure 9-6). Whilst self-reported visual function, as measured by the VFQ-25, increased in the surgery group from worse than the control group, to better than the control group, the final self-reported visual function was not significantly better than the control group. These results supported the hypothesis of an effect of surgery on self-reported visual function, although the effect was small.

9.3.3 Additional study questions

Additional study questions – overall score

The mean (SD) overall score in the surgery and control groups at visit one and visit two are shown in Figure 9-7. A higher additional study questions score represents better functioning and QoL. On inspection of Figure 9-7 the surgery group appeared to have a lower overall score compared to the control group at visit one. From visit one to visit two the surgery group had an increased overall score, whereas the control group score appeared similar. At visit two the score appeared larger in the surgery group compared to the control group.

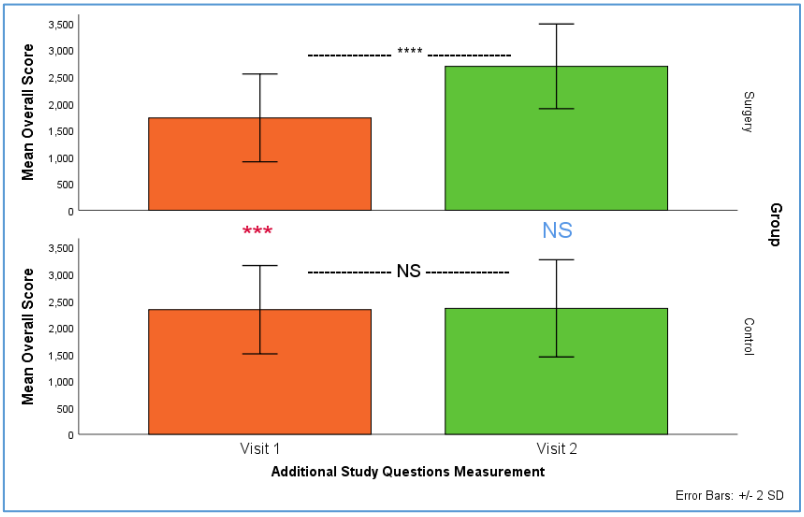


Figure 9-7. Bar chart showing the additional study questions mean overall scores at visit one and visit two, for both groups.

Visit one shown in orange. Visit two shown in green.
Statistical significance reported as NS = not significantly different (p>0.05), *** = significant (p≤0.001), **** = significant (p≤0.0001). Results from the two-way mixed ANOVA are shown.
Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Table 9-5 additionally shows the median and IQR results for comparison to the subscale results, which are presented later in this section (Table 9-6).

Table 9-5. Additional study questions overall scores at visit one and visit two, for both groups

Group		Visit one additional study questions overall score	Visit two additional study questions overall score
Surgery	Mean (SD)	1729.17 (410.77)	2693.75 (394.16)
	Median (IQR)	1825 (831)	2737.5 (394)
Control	Mean (SD)	2331.67 (413.12)	2356.67 (454.78)
	Median (IQR)	2375 (450)	2375 (725)
Additional study questions (0-3300) Higher score = better QoL and functioning			

<u>Parametric analysis of additional study questions overall score.</u>

Two-way mixed ANOVA

The between subjects factor was group (surgery or control group). The within subjects factor was visit (visit one and visit two). These factors were used in all the reported two-way mixed ANOVAs, unless stated otherwise.

The two-way mixed ANOVA results showed there was no significant main effect of group ($p=0.38$), showing that group did not affect the overall score. The main effect of visit was significant ($p<0.0001$) such that the scores were higher at visit two. The two-way interaction group x visit was significant ($p<0.0001$) such that the change in the overall score from visit one to visit two was significantly greater in the surgery group than the control group.

Planned comparisons showed that the difference in the mean overall score between the groups at visit one (control group=2332, surgery group=1729) was significantly lower in the surgery group ($p=0.0009$) and the difference between the groups at visit two (control group=2357, surgery group=2694) was not significantly different ($p>0.05$). These between groups results are shown in Figure 9-8 in red and blue text respectively.

The change in the additional questions overall score for the surgery group and control group were compared separately using a post-hoc paired samples t-test. The control group overall score did not change significantly from visit one (mean=2332) to visit two (mean=2357), $t(14)=0.69$, $p=0.50$, $r=0.18$. The surgery group overall score increased significantly from visit one (mean=1729) to visit two (mean=2694), $t(11)=0.67$, $p<0.0001$, $r=0.90$. This remained significant ($p<0.0001$) after Bonferroni correction for two tests.

ANCOVA

When the covariate (baseline horizontal distance deviation) was included there was no significant main effect of the covariate or group ($p>0.05$), indicating no overall difference between the groups. The main effect of visit remained significant ($p=0.04$); scores were significantly higher at visit two than visit one. The two-way interactions group x covariate and visit x covariate were not significant ($p>0.05$). The two-way interaction group x visit remained significant ($p=0.03$). The three-way interaction visit x group x covariate was not significant ($p>0.05$). The change in additional study questions overall score was significantly greater in the surgery group from visit one to visit two, compared to the control group, but baseline deviation size was not a significant factor.

Summary of the additional study questions overall results

The additional study questions overall scores were lower in the surgery group compared to the control group at visit one, equating to them reporting worse functioning and QoL. Postoperatively the surgery group reported higher (improved) scores, that were then higher than, but not significantly different to, the control group at visit two (shown in Figure 9-7). These results supported the hypotheses. The effect of the covariate, size of deviation at visit one, was not significant and it did not have a significant interaction with any of the other factors.

Additional study questions – subscales scores

The 33 study questions were subcategorised into the following categories: vision (n=11), task performance (n=6), physical symptoms (n=6) and confidence and emotions (n=10). To investigate the categories individually, each subscale was converted to a score out of 100 (maximum 100 and minimum 0), with 100 equating to better functioning or QoL. The median (IQR) of each of the subscale scores, for both groups, at visit one and visit two, are shown in Figure 9-8. On inspection of Figure 9-8 the surgery group appeared to have lower scores in all subscales than the control group at visit one, with all subscale score improving from visit one to visit two. Of note were the large improvements in the vision and confidence and emotions subscale scores. At visit two the surgery group subscale scores appeared similar to or higher than the control group.

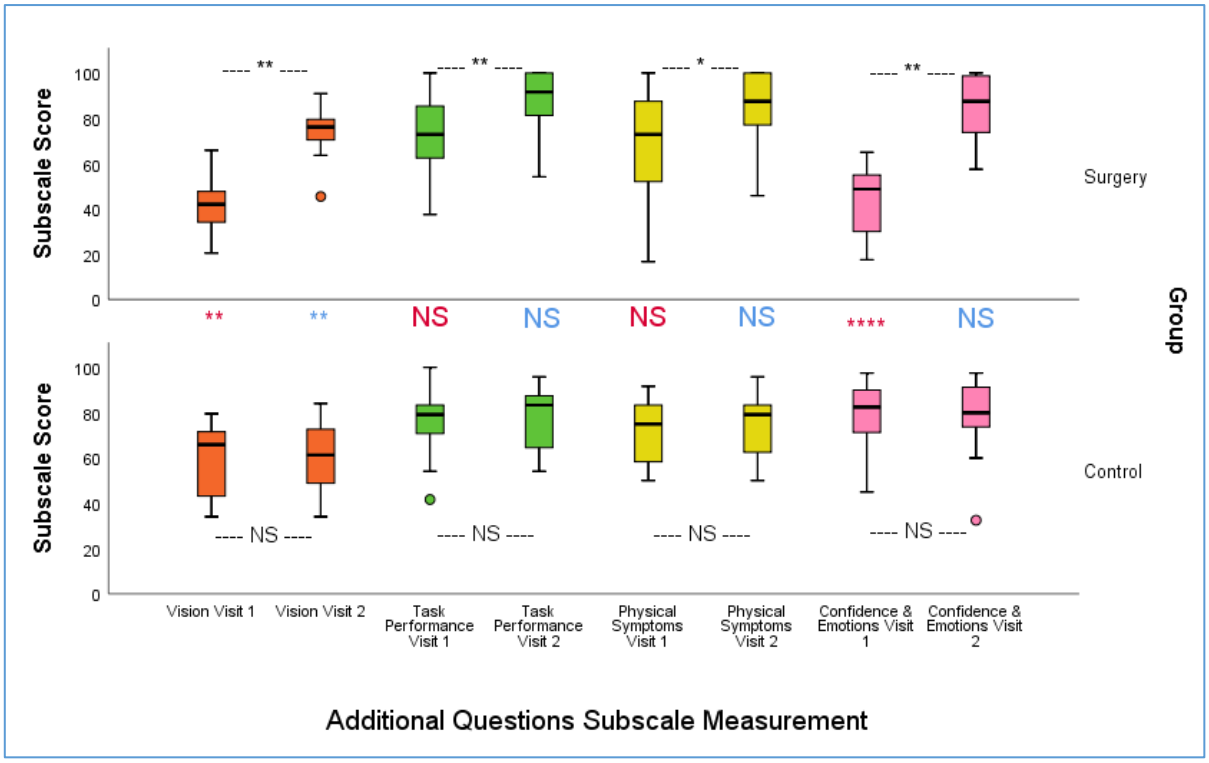


Figure 9-8. Boxplot showing the additional study questions scores for each subscale, at visit one and visit two, for both groups

Vision subscale shown in orange. Task performance subscale shown in green. Physical symptoms subscale shown in yellow. Confidence and emotions subscale shown in pink.
Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), **** = significant (p≤0.0001).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Table 9-6 additionally presents the mean results of the additional study questions subscales for comparison to the overall scores presented earlier in this section (Table 9-5).

Table 9-6. Additional study questions scores for each subscale, at visit one and visit two, for both groups									
Group		Study questions - visit 1				Study questions – visit 2			
	Subscales	Vision	Task performance	Physical symptoms	Confidence & emotions	Vision	Task performance	Physical symptoms	Confidence & emotions
Surgery	Mean	41.10	72.22	66.67	44.37	74.43	87.15	84.72	84.37
	SD	12.16	20.44	27.24	16.10	11.67	16.33	16.70	15.45
	Median	**42.04****	**72.92** ^{NS}	**72.92** ^{NS}	**48.75******	**76.14**** $^{(**)}$	**91.67** ^{NS} $^{(**)}$	**87.50** ^{NS} $^{(*)}$	**87.50** ^{NS} $^{(**)}$
	IQR	14.77	26.04	38.54	27.50	10.23	21.88	26.04	28.75
Control	Mean	59.39	75.83	71.39	79.50	59.85	77.50	73.61	79.17
	SD	15.58	14.62	15.18	15.59	15.69	14.33	15.96	17.08
	Median	**65.91**	**79.17**	**75.00**	**82.50**	**61.36**^{NS}	**83.33**^{NS}	**79.17**^{NS}	**80.00**^{NS}
	IQR	31.82	12.50	29.17	22.50	27.27	25.00	20.83	20.00

Median results used in the analysis shown in bold.

Results of statistical analysis shown in superscript. NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), **** = significant (p≤0.0001).

Text in black = within groups analysis, visit one to visit two (Wilcoxon signed rank test). Results shown in brackets next to the surgery group results where between groups analysis also reported.

Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of additional study questions subscales

Visit one

The vision subscale score in the control group was significantly higher than the surgery group, U=37, z=2.59, p=0.009, r=0.50. The confidence and emotions subscale score in the control group was significantly higher than the surgery group, U=11.5, z=3.83, p<0.0001, r=0.74. The task performance subscale score in the control group was not significantly different from the surgery group, U=79, z=0.54, p=0.61, r=0.10. The physical symptoms subscale score in the control group was not significantly different from the surgery group, U=89, z=0.05, p=0.98, r=0.009. These between groups results at visit one are shown in Figure 9-8 and Table 9-6 in red text (Independent samples Mann-Whitney U test).

Visit one to visit two

In all of the subscales, from visit one to visit two the control group subscale scores did not change significantly (p>0.05). The surgery group subscale scores all increased significantly.
The vision subscale score at visit one was significantly higher at visit two, T=78, p=0.002, r=0.62. The task performance subscale score at visit one was significantly higher at visit two, T=54, p=0.007, r=0.55. The physical symptoms subscale score at visit one was significantly higher at visit two, T=50.5, p=0.019, r=0.48. The confidence and emotions subscale score at visit one was significantly higher at visit two, T=78, p=0.002, r=0.62. The within groups results are shown in Figure 9-8 and Table 9-6 in black text (Wilcoxon signed rank test).

Visit two

At visit two the vision subscale score in the surgery group was significantly higher than the control group, U=142, z=2.54, p=0.01, r=0.49. There was no significant difference between the groups in the task performance, physical symptoms and confidence and emotions subscale scores (p>0.05). These between groups results at visit two are shown in Figure 9-8 and Table 9-6 in blue text (Independent samples Mann-Whitney U test).

Additional study questions – individual vision questions

The results from the individual questions relating to vision are shown in Table 9-7. The additional study questions and scoring method used are shown in Appendix L.

1. *My eyes have been turning*
2. *I have been able to control my eye position*
3. *I have been able to focus my eyes*
4. *I have been confused by my vision*
5. *I have been able to swap to look with each eye separately*
6. *My vision has looked central*
7. *When I have been in a busy place, I have found it difficult to see (for example, a shopping centre or train station)*
8. *I have found it difficult to move my eyes to look around*
9. *One of my eyes has been working much harder than the other eye*

10. *I have been able to look through both eyes at the same time*
11. *I have been able to use my eyes together*

Table 9-7. Additional study questions scores, for each of the questions relating to vision, at visit one and visit two, for both groups

		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 1	Visit one	25 (50)	12.5 (25)	NS (p=0.07)
	Visit two	25 (75) ^{NS}	100 (18.75) **	**** (p≤0.0001)
Question 2	Visit one	50 (50)	0 (18.75)	*(p=0.04)
	Visit two	25 (75) ^{NS}	75 (25) **	*(p=0.04)
Question 3	Visit one	75 (25)	87.5 (68.75)	NS (p=0.87)
	Visit two	75 (25) ^{NS}	75 (25) ^{NS}	NS (p=0.61)
Question 4	Visit one	75 (25)	75 (50)	NS (p=0.49)
	Visit two	75 (50) ^{NS}	100 (0) ^{NS}	NS (p=0.18)
Question 5	Visit one	75 (50)	0 (25)	**(p=0.002)
	Visit two	75 (25) ^{NS}	37.5 (75) ^{NS}	*(p=0.02)
Question 6	Visit one	100 (25)	25 (50)	**(p=0.004)
	Visit two	75 (25) ^{NS}	100 (25) **	NS (p=0.40)
Question 7	Visit one	100 (25)	75 (50)	NS (p=0.46)
	Visit two	100 (25) ^{NS}	100 (18.75) ^{NS}	NS (p=0.87)
Question 8	Visit one	100 (0)	100 (56.25)	NS (p=0.52)
	Visit two	100 (25) ^{NS}	100 (0) ^{NS}	NS (p=0.49)
Question 9	Visit one	25 (50)	0 (0)	NS (p=0.053)
	Visit two	25 (50) ^{NS}	25 (50) ^{NS}	NS (p=0.46)
Question 10	Visit one	0 (50)	62.5 (100)	NS (p=0.15)
	Visit two	0 (50) ^{NS}	100 (43.75) ^{NS}	**(p=0.005)
Question 11	Visit one	0 (50)	0 (43.75)	NS (p=0.90)
	Visit two	0 (50) ^{NS}	87.5 (87.5)**	*(p=0.02)

NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001), **** = significant (p≤0.0001).
Significant results highlighted in yellow.
Text in black = within groups analysis (Wilcoxon signed rank test).
Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Visit one

At visit one there was no statistically significant difference between the surgery group and control group for questions 1, 3, 4, 7, 8, 9, 10 and 11 (p>0.05). The score for questions 2 (p=0.04) (control of eye position), 5 (p=0.002) (ability to swap fixation) and 6 (p=0.004) (my vision has looked

central) was significantly lower in the surgery group than the control group. Results shown in red in Table 9-7 (Independent samples Mann-Whitney U test).

Visit one to visit two

From visit one to visit two none of the question responses changed significantly in the control group ($p>0.05$). In the surgery group, there was a significant increase (improvement) in the responses to questions 1 ($p=0.002$) (my eyes have been turning), 2 ($p=0.004$) (control of eye position), 6 ($p=0.003$) (my vision has looked central) and 11 ($p=0.01$) (using eyes together). There was no significant change in the responses to questions 3, 4, 5, 7, 8, 9 and 10 ($p>0.05$). Results shown in black in Table 9-7 (Wilcoxon signed rank test).

Visit two

At visit two there was no statistically significant difference between the surgery group and control group for questions 3, 4, 6, 7, 8 and 9 ($p>0.05$). The score for questions 1 ($p<0.0001$) (my eyes have been turning), 2 ($p=0.04$) (control of eye position), 10 ($p=0.005$) (looking through both eyes at the same time) and 11 ($p=0.02$) (using eyes together) were significantly higher in the surgery group than the control group. The score for question 5 ($p=0.02$) (ability to swap fixation) was significantly lower in the surgery group than the control group. Results shown in blue in Table 9-7 (Independent samples Mann-Whitney U test).

Additional study questions – individual task performance questions

The results from the individual questions relating to task performance are shown in Table 9-8.

12. *My eyes have limited my ability (for example, at work, in activities I enjoy or in undertaking day-to-day tasks)*
13. *I have had difficulty with eye hand coordination*
14. *I have made mistakes when performing day-to-day tasks*
15. *I have been able to perform day-to-day tasks quickly*
16. *I have had difficulty looking at screens (for example, computer, mobile phone or tablet screens)*
17. *I have had problems with my balance*

Table 9-8. Additional study questions scores, for each of the questions relating to task performance, at visit one and visit two, for both groups				
		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 12	Visit one	100 (25)	50 (50)	NS (p=0.053)
	Visit two	100 (25) NS	87.5 (43.75) NS	NS (p=0.65)
Question 13	Visit one	75 (25)	100 (43.75)	NS (p=0.15
	Visit two	75 (25) NS	100 (0) NS	**(p=0.005)
Question 14	Visit one	75 (50)	87.5 (43.75)	NS (p=0.49)
	Visit two	75 (25) NS	100 (18.75) NS	NS (p=0.22)
Question 15	Visit one	75 (0)	75 (43.75)	NS (p=0.52)
	Visit two	75 (25) NS	100 (25)*	NS (p=0.24)
Question 16	Visit one	75 (50)	62.5 (43.75)	NS (p=0.32)
	Visit two	75 (50) NS	100 (25)*	NS (p=0.35)
Question 17	Visit one	75 (50)	100 (37.50)	NS (p=0.37)
	Visit two	75 (50) NS	100 (37.50) NS	NS (p=0.12)
NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), Significant results highlighted in yellow. Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).				

Visit one

At visit one there was no significant difference between the control group and the surgery group for questions 12, 13, 14, 15, 16 and 17 (p>0.05). Results shown in red in Table 9-8 (Independent samples Mann-Whitney U test).

Visit one to visit two

From visit one to visit two none of the question responses changed significantly in the control group (p>0.05). In the surgery group, there was no significant change in the responses to questions 12, 13, 14 and 17 (p>0.05). However, there was a significant increase (improvement) in the responses to questions 15 (p=0.01) (performing tasks quickly) and 16 (p=0.02) (difficulty looking at screens). Results shown in black in Table 9-8 (Wilcoxon signed rank test).

Visit two

At visit two there was no significant difference between the control group and the surgery group for questions 12, 14, 15, 16 and 17 (p>0.05). For question 13 the surgery group had a significantly higher score (less difficulty with hand eye coordination) than the control group (p=0.005). Results shown in blue in Table 9-8 (Independent samples Mann-Whitney U test).

Additional study questions – individual physical symptoms questions

The results from the individual questions relating to physical symptoms are shown in Table 9-9.

18. I have had headaches

19. I have needed to take breaks because of my eyes (for example, breaks from work, from activities I enjoy or when performing day-to-day tasks)

20. I have felt my eyes pulling

21. My eyes have felt tight

22. I have felt tired

23. I have felt my eyes turning

Table 9-9. Additional study questions scores, for each of the questions relating to physical symptoms, at visit one and visit two, for both groups

		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 18	Visit one	75 (50)	75 (50)	NS (p=0.35)
	Visit two	75 (50) NS	100 (25) **	NS (p=0.09)
Question 19	Visit one	75 (25)	100 (50)	NS (p=0.30)
	Visit two	100 (25) *	100 (18.75) NS	NS (p=0.46)
Question 20	Visit one	75 (50)	75 (50)	NS (p=0.58)
	Visit two	100 (25) NS	100 (25) *	NS (p=0.83)
Question 21	Visit one	75 (50)	100 (43.75)	NS (p=0.61)
	Visit two	75 (50) NS	100 (25) NS	NS (p=0.55)
Question 22	Visit one	50 (0)	50 (43.75)	NS (p=0.61)
	Visit two	50 (25) NS	100 (50) *	NS (p=0.06)
Question 23	Visit one	75 (50)	62.5 (75)	NS (p=0.58)
	Visit two	75 (25) NS	100 (25) *	NS (p=0.35)

NS = not significantly different (p>0.05), * = significant (p≤0.05).

Significant results highlighted in yellow.

Text in black = within groups analysis (Wilcoxon signed rank test).

Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Visit one

At visit one there was no significant difference between the control group and the surgery group for any of the scores in the individual physical symptom's questions, 18 to 23 (p>0.05). Results shown in red in Table 9-9 (Independent samples Mann-Whitney U test).

Visit one to visit two

From visit one to visit two there was no significant change in the response to question 21 in either group ($p>0.05$). In the control group there was a significant increase in the response to question 19 ($p=0.03$) (taking breaks), but no significant change in the surgery group. In the surgery group there was a significant increase in the responses to questions 18 ($p=0.008$) (headaches), 20 ($p=0.04$) (eyes pulling), 22 ($p=0.05$) (tired) and 23 ($p=0.03$) (feeling eyes turning). For questions 18, 20, 21, 22 and 23 there was no significant change in the responses in the control group ($p>0.05$). Results shown in black in Table 9-9 (Wilcoxon signed rank test).

Visit two

At visit two there was no significant difference between the groups in their responses to questions 18-23 ($p<0.05$). Results shown in blue in Table 9-9 (Independent samples Mann-Whitney U test).

Additional study questions – individual confidence and emotions questions

Results from the individual questions relating to confidence and emotions are shown in Table 9-10.

24. I have been able to talk to people

25. I have avoided face-to-face situations

26. People have treated me differently

27. I have had confidence in my vision

28. I have had confidence in my abilities (for example, at work, to take part in activities I enjoy or to undertake day-to-day tasks)

29. I have had self-confidence

30. I have liked the way my eyes look

31. I have felt anxious

32. I have felt embarrassed about my eyes

33. I have felt happy

Table 9-10. Additional study questions scores, for each of the questions relating to confidence and emotions, at visit one and visit two, for both groups

		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 24	Visit one	100 (0)	75 (25)	* (p=0.03)
	Visit two	100 (0) NS	100 (0)*	NS (p=0.90)
Question 25	Visit one	100 (25)	50 (25)	**** (p≤0.0001)
	Visit two	100 (0) NS	100 (18.75)**	NS (p=0.65)
Question 26	Visit one	100 (25)	50 (25)	*** (p=0.001)
	Visit two	100 (25) NS	100 (75)**	NS (p=0.79)
Question 27	Visit one	75 (25)	50 (68.75)	* (p=0.02)
	Visit two	75 (25) NS	100 (25) *	NS (p=0.32)
Question 28	Visit one	75 (25)	62.5 (43.75)	* (p=0.02)
	Visit two	75 (25) NS	87.5 (25) NS	NS (p=0.65)
Question 29	Visit one	75 (25)	50 (50)	** (p=0.005)
	Visit two	75 (50) NS	75 (43.75) *	NS (p=1.0)
Question 30	Visit one	50 (75)	0 (0)	*** (p=0.001)
	Visit two	75 (25) NS	100 (25) **	*(p=0.02)
Question 31	Visit one	75 (25)	37.5 (62.5)	** (p=0.004)
	Visit two	75 (50) NS	100 (43.75) **	NS (p=0.43)
Question 32	Visit one	75 (50)	12.5 (68.75)	** (p=0.005)
	Visit two	100 (50) NS	100 (0) **	NS (p=0.43)
Question 33	Visit one	75 (0)	75 (68.75)	NS (p=0.07)
	Visit two	75 (0) NS	75 (25) *	NS (p=0.24)

NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001), **** = significant (p≤0.0001).

Significant results highlighted in yellow.

Text in black = within groups analysis (Wilcoxon signed rank test).

Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Visit one

At visit one, the surgery group had a significantly lower score than the control group for questions 24 – 32 ($p<0.05$). There was no significant difference between the surgery group and the control group for question 33 ($p=0.07$) (feeling happy). Results shown in red in Table 9-10 (Independent samples Mann-Whitney U test).

Visit one to visit two

There were no significant changes in the scores reported by the control group for questions 24-33 ($p>0.05$). The surgery group reported significantly higher scores at visit two compared to visit one for questions 24-27 and questions 29-33 ($p\leq0.05$). The surgery group reported no significant

change in the score for question 28 (p=0.06) (confidence in abilities). Results shown in black in Table 9-10 (Wilcoxon signed rank test).

Visit two

At visit two there was no significant difference between the surgery group and the control group for questions 24-29 and questions 31-33. For question 30 the surgery group reported a significantly higher score than the control group (p=0.02) (I have liked the way my eyes looked). Results shown in blue in Table 9-10 (Independent samples Mann-Whitney U test).

Summary of the results from the additional questions subscales and individual questions

At visit one the surgery group had lower vision subscale and confidence and emotion subscale scores than the control group, although the task performance and physical symptoms subscales were not significantly different (shown in red in Figure 9-8). From visit one to visit two all the subscale scores improved significantly in the surgery group only (shown in black in Figure 9-8). At visit two the surgery group reported subscale scores that were higher than the control group, however only the vision subscale score was significantly higher (shown in blue in Figure 9-8). Having strabismus surgery led to improvements in all of the additional questions' subscales to the extent that the surgery group reported similar task performance, physical symptoms and confidence and emotions, and better vision than the control group. These results supported the hypothesis of an effect of surgery on QoL and functioning.

Analysis of the responses to individual questions showed that surgery led to patients reporting improvements in the perception of their eye turning, being able to control their eye position, their vision looking central, using their eyes together, eye hand coordination, performing tasks quickly, looking at screens, headaches, the eyes pulling, feeling tired, feeling the eyes turning and all of the questions relating to confidence and emotions (except confidence in abilities).

9.3.4 PROMs results presented by theme

Vision

The PROMs results relating specifically to the theme of vision are the VFQ-25 results (section 9.3.2) and the additional study questions, vision subscale results (section 9.3.3).

Task performance

The PROMs results specifically relating to the theme of task performance are the AS-20 function subscale results (section 9.3.1), the additional study questions, task performance subscale results (section 9.3.3) and the VFQ-25 driving subscale results (Table 9-11).

Table 9-11. VFQ-25 driving subscale results at visit one and visit two, for both groups			
	Control group (n=9) Median (IQR)	Surgery group (n=9) Median (IQR)	Statistical significance (between groups)
Visit one	83.33 (29.17)	91.67 (20.84)	NS (p=0.39)
Visit two	91.67 (37.50) NS	100 (12.50) NS	NS (p=0.16)
VFQ-25 driving subscale scores (0-100) (Questions 15c, 16, 16c) Non drivers not included. Higher score = better self-reported visual function related to driving NS = not significantly different (p>0.05). Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).			

There was no significant difference in VFQ-25 driving subscale score between the groups, at visit one or visit two, and no significant difference within the groups, from visit one to visit two (p<0.05).

Physical symptoms

The PROMs results specifically relating to the theme of physical symptoms are the additional study questions, physical symptoms subscale results (section 9.3.3), AS-20 results for questions 15 and 20 and the VFQ-25 results for questions 4 and 19.

AS-20 questions - physical symptoms

The results of AS-20 questions 15 and 20, relating to physical symptoms, are shown in Table 9-12.

Question 15 'My eyes feel strained'

Question 20 'I need to take frequent breaks when reading because of my eyes'

Table 9-12. AS-20 scores at visit one and visit two, for both groups, for questions 15 and 20 relating to physical symptoms				
		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 15	Visit one	50 (50)	50 (43.75)	NS (p=0.40)
	Visit two	50 (25) NS	100 (25) **	*(p=0.04)
Question 20	Visit one	75 (50)	75 (68.75)	NS (p=0.43)
	Visit two	75 (25) NS	100 (25) **	NS (p=0.20)
NS = not significantly different, (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001), **** = significant (p≤0.0001). Yellow highlighting statistically significant results. Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test)				

The surgery group gave similar responses to the control group to the AS-20 physical symptoms questions at visit one (p>0.05). Postoperatively the surgery group reported significantly improved physical symptoms (p<0.01). At visit two the surgery group reported significantly less eye strain compared to the control group (p=0.04) and less need to take breaks. However, the surgery group response to question 20 was not significantly different from the control group at visit two (p>0.05). These results broadly supported the hypothesis (section 8.5.1).

VFQ-25 questions - physical symptoms

The individual VFQ-25 question results relating to physical symptoms (question 4 and 19) are reported in Table 9-13.

Question 4 'how much pain or discomfort have you had in and around your eyes?'

Question 19 'how much does pain or discomfort in or around your eyes keep you from doing what you'd like to be doing?'

Table 9-13. VFQ-25 results at visit one and visit two, for both groups, from questions 4 and 19, relating to physical symptoms

		Control group Median (IQR)	Surgery group Median (IQR)	Statistical significance (between groups)
Question 4	Visit one	75 (25)	87.5 (25)	NS (p=0.68)
	Visit two	75 (50) ^{NS}	100 (0) ^{NS}	*(p=0.02)
Question 19	Visit one	100 (25)	87.5 (25)	NS (p=1.0)
	Visit two	100 (25) *	100 (0) *	NS (p=0.35)

NS = not significantly different (p>0.05), * = significant (p≤0.05).
Yellow highlighting significant results.
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

At visit one there was no significant difference between the control group and the surgery group in the responses to VFQ-25 questions 4 and 19 (p>0.05). From visit one to visit two there was no significant change in the scores reported by either group, for question 4 (p>0.05). For question 19, both groups reported a significantly higher score at visit two (p<0.05). At visit two the surgery group reported significantly higher responses to question 4 than the control group (p=0.02). There was no significant difference between the control group and the surgery group in the responses to question 19 (p=0.35).

Confidence and emotions

The PROMs results specifically relating to the theme of confidence and emotions are the AS-20 psychosocial subscale results (section 9.3.1) and the additional study questions, confidence and emotions subscale results (section 9.3.3).

Summary of the PROMs results presented by theme

Vision PROMs improved postoperatively in the surgery group using the additional study questions, vision subscale and to a smaller extent using the VFQ-25. Task performance PROMs improved postoperatively in the surgery group using the AS-20 function subscale and the additional study questions, task performance subscale. However, the VFQ-25 driving subscale did not change after surgery. Physical symptoms PROMs improved postoperatively in the surgery group using the additional study questions, physical symptoms subscale and the AS-20 and VFQ-25 questions that related to physical symptoms. Confidence and emotions PROMs markedly improved postoperatively in the surgery group using both the AS-20 psychosocial subscale and the additional study questions, confidence and emotions subscale.

9.4 Vision

9.4.1 Visual acuity

Distance VA in the better and the worse seeing eye was reported (logMAR) in the control group (n=15) and the surgery group (n=12). Near VA both eyes open was reported in the control group (n=15) and the surgery group (n=11). Near VA was omitted from the visit one assessment of participant 5. No near VA data was included for this participant at visit two.

Distance VA

The median (IQR) distance VA results are presented in Figure 9-9. Distance VA in the better seeing eye was around 0.00 logMAR in both the surgery and control group. Distance VA in the worse seeing eye appeared worse in the surgery group. From inspection of Figure 9-9, there was no apparent change in VA from visit one to visit two, in either group.

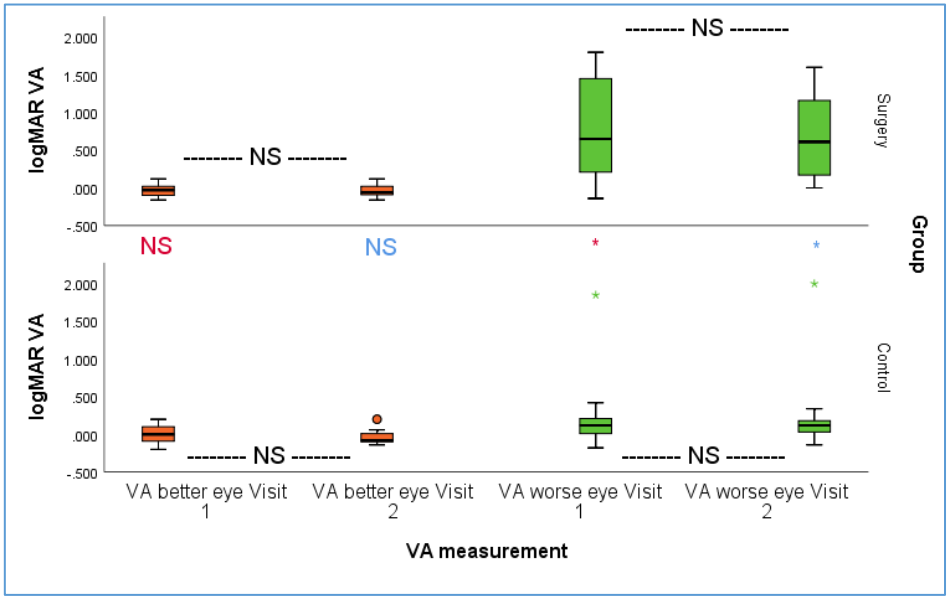


Figure 9-9. Boxplot showing the distance VA (logMAR) at visit one and visit two, for both groups

VA in the better eye shown in orange and in the worse eye shown in green.
Statistical significance reported as NS = not significantly different ($p>0.05$), * = significant ($p\leq0.05$). Text in black = within groups analysis (Wilcoxon signed rank test).
Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of distance VA

Visit one

Distance VA in the better eye at visit one in the surgery group (Mdn=-0.03) and the control group (Mdn=0.00) was not significantly different, U=74.5, z=0.76, p=0.46, r=0.15. Distance VA in the worse eye at visit one in the surgery group (Mdn=0.65) was significantly poorer than in the control group (Mdn=0.12), U=133.5, z=2.13, p=0.03, r=0.41. These between groups results are shown in Figure 9-9 (red text).

Visit one to visit two

Distance VA in the better eye in the control group at visit one (Mdn=0.00) was not significantly different to visit two (Mdn=-0.08), T=23, p=0.21, r=0.23. In the surgery group distance VA in the better eye at visit one (Mdn=-0.03) was not significantly different to visit two (Mdn=-0.06), T=21.5, p=0.91, r=0.02. Distance VA in the worse eye in the control group, at visit one (Mdn=0.12) was not significantly different to visit two (Mdn=0.12), T=44.5, p=0.30, r=0.19. Distance VA in the worse eye in the surgery group at visit one (Mdn=0.65) was not significantly different to visit two (Mdn=0.61), T=10, p=0.07, r=0.36. Within groups results are shown in Figure 9-9 (black text).

Visit two

At visit two, distance VA in the better eye, in the control group (Mdn=-0.08) was not significantly different to the surgery group (Mdn=-0.06), U=96, z=0.30, p=0.79, r=0.06. Distance VA in the worse eye, at visit two, in the surgery group (Mdn=0.61) was significantly poorer than in the control group (Mdn=0.12), U=135.5, z=2.23, p=0.03, r=0.43. These between group results are shown in Figure 9-9 (blue text).

Near VA

Mean (SD) near VA, measured with both eyes open, in the control group (visit one=0.11, visit two=0.08) and the surgery group (visit one=0.13, visit two=0.17) are shown in Figure 9-10. On inspection of Figure 9-10 there appeared to be a minimal difference in near VA in the two groups at visit one and visit two. Between visits there appeared to be a small improvement in near VA in the control group and a small worsening in the surgery group.

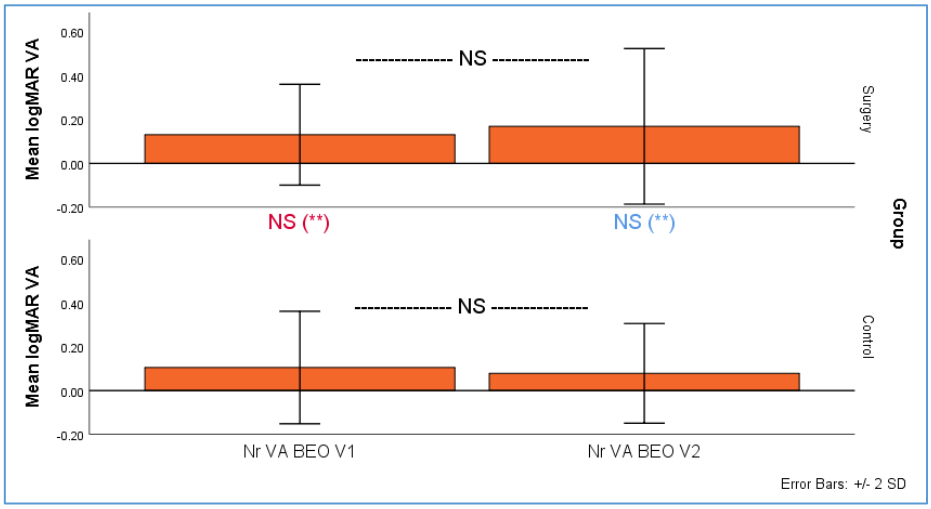


Figure 9-10. Bar chart showing near VA (logMAR) at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01). Two-way mixed ANOVA analysis reported in text on the figure. ANCOVA results additionally reported in brackets.
Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Parametric analyses of near VA

Two-way mixed ANOVA

Levene's test was significant (near VA visit two) (p<0.05), indicating that there was a significant difference in the error variance across the groups. The results are reported here for comparison to the ANCOVA below. There was no significant main effect of group or visit on near VA (p>0.05). There was no significant interaction group x visit (p=0.09) indicating stability of near VA across groups and visits. Planned comparisons showed that the difference in the near VA between the groups at visit one and visit two were not significant (p>0.05) (shown on Figure 9-10 in red and

blue text respectively). The change in near VA for the surgery group and control group were compared separately using a post-hoc paired samples t-test. The control group and the surgery group did not change significantly from visit one to visit two (p>0.05) (shown in Figure 9-10 in black text).

ANCOVA

Levene's test was not significant (p>0.05), indicating that there was no significant difference in the error variance of the near VA results across the groups, after the covariate (baseline size of deviation at visit one) was included. With the covariate included in the model, there was a significant main effect of group (p=0.004) with the surgery group having worse near VA than the control group. There was no significant main effect of visit or the covariate (p>0.05) on near VA. There was a significant two-way interaction group x covariate (p=0.004). Figure 9-11 shows that those with larger baseline deviations had better near VA at visit two in the surgery group, but worse near VA in the control group. There were no other significant interactions (group x visit, visit x covariate, or visit x covariate x group) (p>0.05).

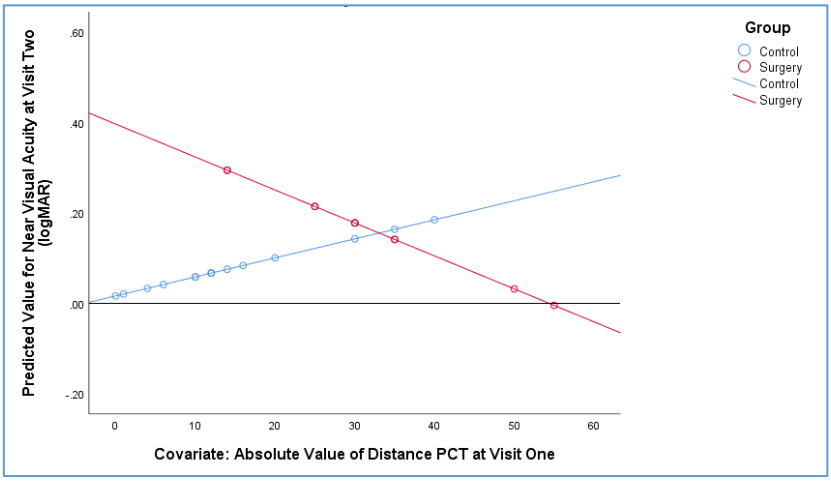


Figure 9-11. Scatter plot showing the interaction effect of the covariate on near VA at visit two

Near VA between the groups, at visit one and visit two, including the covariate were compared (planned comparison). At visit one the difference between the groups was significant (p=0.009) and the interaction with the covariate was significant (p=0.003). At visit two the difference between the groups was significant (p=0.005), the effect of the covariate was significant (p=0.03) and the interaction with the covariate was significant (p=0.01). These results are shown in Figure 9-10 (in brackets) for comparison to the two-way mixed ANOVA results.

For comparison to the distance VA analysis, non-parametric analysis of the near VA measurements was performed. This showed no significant between group differences (at either visit) and no significant within group differences (in either group) in near VA (p<0.05).

Summary of VA results

In the distance, there was no difference in the VA in the better eye between the two groups, although the surgery group had worse VA in the worse eye (Figure 9-9). At near the surgery group had worse near VA than the control group (after the covariate was included in the analysis) (Figure 9-10). There was no change in near or distance VA from visit one to visit two in either group (Figures 9-9 and 9-10). Surgery had no effect on near or distance VA and hence the hypothesis was rejected.

9.4.2 Cover test and fixation grade

The ability to take up fixation with the strabismic eye was graded (1-7 scale). The median (IQR) fixation grade in the control group (n=15) and surgery group (n=12), at visit one and visit two are shown in Figure 9-12. There appeared to be a higher (better) fixation grade in the control group compared to the surgery group, but no apparent difference (in either group) from visit one to visit two.

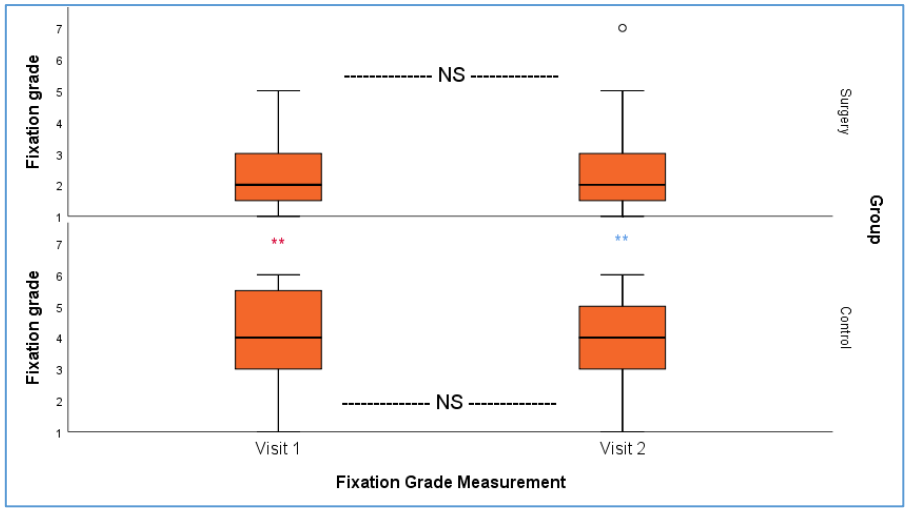


Figure 9-12. Boxplot showing the fixation grade results at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01). Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of fixation grade

Fixation grade in the surgery group was significantly lower (worse) than the control group at visit one, U=33, z=2.83, p=0.004, r=0.54 (Figure 9-12, red text). Fixation grade in the control group at visit one was not significantly different to visit two, T=7, p=0.41, r=0.15. Fixation grade in the surgery group at visit one was not significantly different to visit two, T=3, p=0.18, r=0.27 (Figure 9-

12, black text). Fixation grade in the surgery group was significantly lower (worse) than the control group at visit one, U=46, z=2.19, p=0.03, r=0.42 (Figure 9-12, in blue text).

Summary of the fixation grade results

At visit one and two the surgery group had a significantly lower fixation grade than the control group, which equated to a poorer ability to take up fixation with the strabismic eye. There was no significant change in fixation grade from visit one to visit two, in either group. Surgery had no effect on fixation grade and the hypothesis was rejected.

9.4.3 Ocular movements

OM were tested in all participants (n=27) at both visits. The restrictions and underactions of each horizontal rectus muscle and the total of all the horizontal recti combined, are presented in Figure 9-13. The right eye results (LR and medial rectus (MR)) are shown, followed by the left eye results (LR and MR) and the total horizontal recti results. There appeared to be little difference between the groups at visit one or two. In the surgery group the total restrictions appeared to increase from visit one to visit two.

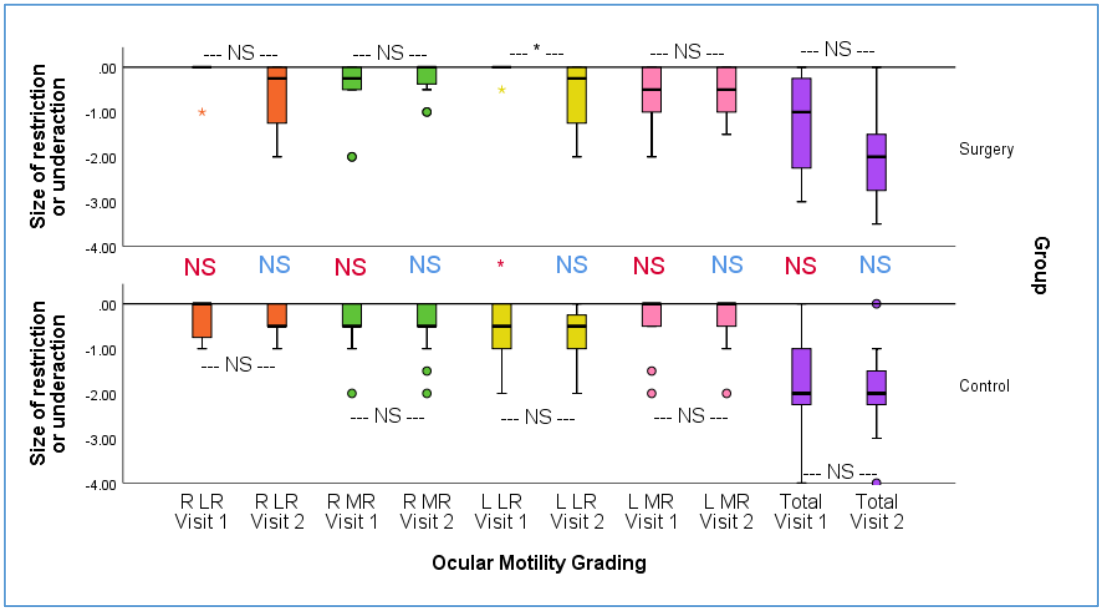


Figure 9-13. Boxplot showing the OM restriction and underaction results at visit one and visit two, for both groups

R LR results shown in orange. R MR results shown in green. L LR results shown in yellow. L MR results shown in pink. The total horizontal recti results shown in purple.

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05). Text in black = within groups analysis (Wilcoxon signed rank test).
Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of OM grading

There was no significant difference in the size of the individual horizontal recti restrictions, or the total restriction between the groups control group and the surgery group ($p>0.05$), except for the left LR. The left LR restrictions in the control group were significantly larger than the surgery group, $U=145.5$, $z=3.06$, $p=0.005$, $r=0.59$ (shown in Figure 9-13, red text). From visit one to visit two there was no significant change in OM grading in the control group ($p>0.05$). In the surgery group, only the increase in the left LR limitation was significant, from visit one to visit two, $T=2$, $p=0.04$, $r=0.42$ (shown in Figure 9-13, black text). At visit two there was no significant difference in the OM grading (individual horizontal recti restrictions, or the total restriction) between the groups ($p>0.05$) (shown in Figure 9-13, blue text).

Parametric analysis of the OM total restrictions

The total OM restriction and underaction data was normally distributed, therefore additional parametric analysis was performed (two-way mixed ANOVA and ANCOVA). There was no significant main effect of group or visit on total OM restriction and underaction and no interaction group x visit ($p>0.05$). With the covariate baseline size of deviation included in the ANCOVA, there remained no significant main effects and no significant interactions ($p>0.05$).

Summary of the OM results

At visit one the control group had significantly larger left LR limitations than the surgery group. Only the left LR limitation increased significantly in the surgery group postoperatively, leading to there being no significant differences between the control group and surgery group at visit two. These results supported the hypothesis of an effect of surgery on OM.

9.4.4 Binocular summation

Binocular summation results were reported for the control group (n=15) and the surgery group (n=12), at visit one and visit two, using the letter scoring method and the categorisation method.

Binocular summation – letter scoring method

The binocular summation results are shown in Figure 9-14. The surgery group and control group appeared to have similar binocular summation scores at visit one. From visit one to visit two the surgery group appeared to have an increase in 100% binocular summation score, leading to a higher score than the control group at visit two, at 100% contrast.

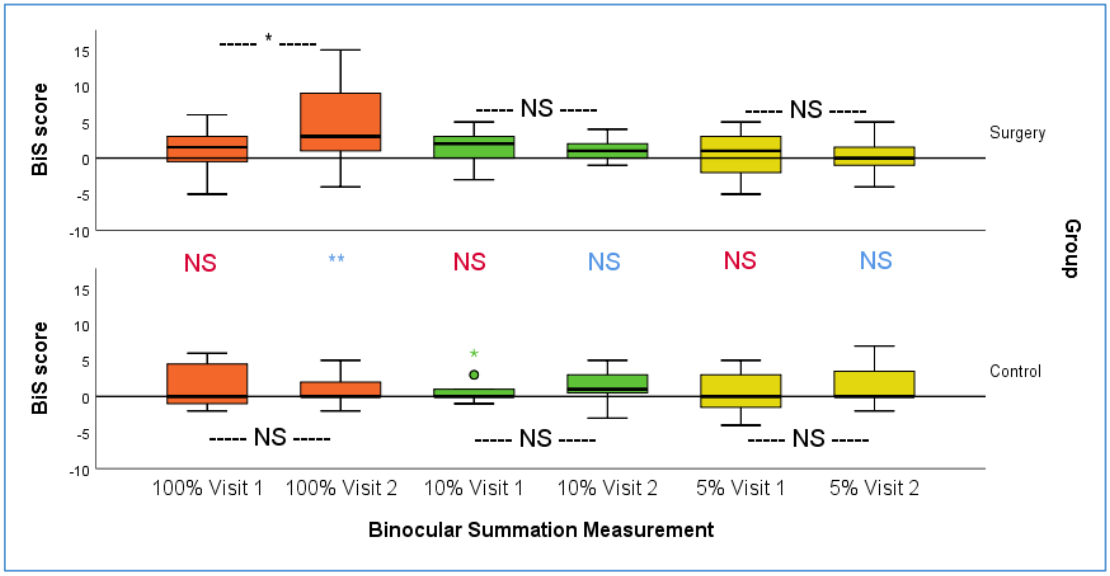


Figure 9-14. Boxplot showing the binocular summation results at 100%, 10% and 5% contrast, at visit one and visit two, for both groups

100% contrast shown in orange. 10% contrast shown in green. 5% contrast shown in yellow.
Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of binocular summation results - letter scoring method

Visit one

At visit one there was no significant difference in the 100%, 10% or 5% binocular summation scores between the control and the surgery group (p>0.05) (shown in Figure 9-14, red text).

Visit one to visit two

From visit one to visit two the binocular summation scores were not significantly different in the control group at 100%, 10% and 5% (p>0.05). For the surgery group, the 100% binocular summation score at visit one (Mdn=1.5) was significantly higher at visit two (Mdn=3), T=49.5, p=0.03, r=0.46. The 10% and 5% binocular summation scores at visit one (10% Mdn=2, 5% Mdn=1) were lower at visit two (10% Mdn=1, 5% Mdn=0), but these differences were not statistically significant (p>0.05). These within groups results are shown in Figure 9-14 (black text).

Visit two

At visit two the 100% binocular summation score in the surgery group (Mdn=3) was significantly higher than the control group (Mdn=0), U=142.5, z=2.59, p=0.009, r=0.50. The 10% and 5%

binocular summation scores at visit two were not significantly different in the surgery and the control group (p>0.05). These between groups results are shown in Figure 9-14 (blue text).

Parametric analysis of the change in binocular summation score – letter scoring method

The change in binocular summation score (letter scoring method) was calculated by subtracting the visit one score from the visit two score. A positive change equated to increased binocular summation or reduced binocular inhibition. A negative change equated to reduced binocular summation or increased binocular inhibition. A zero-change equated to no change in binocular summation.

The mean (±2SE) change in binocular summation scores are displayed in Figure 9-15. The surgery group appeared to have a positive change in binocular summation score at 100% contrast, compared to the control group who appeared to have a negative change. The control group had positive changes in binocular summation at 10% and 5% contrast, compared to the surgery group who had negative changes at these contrast levels.

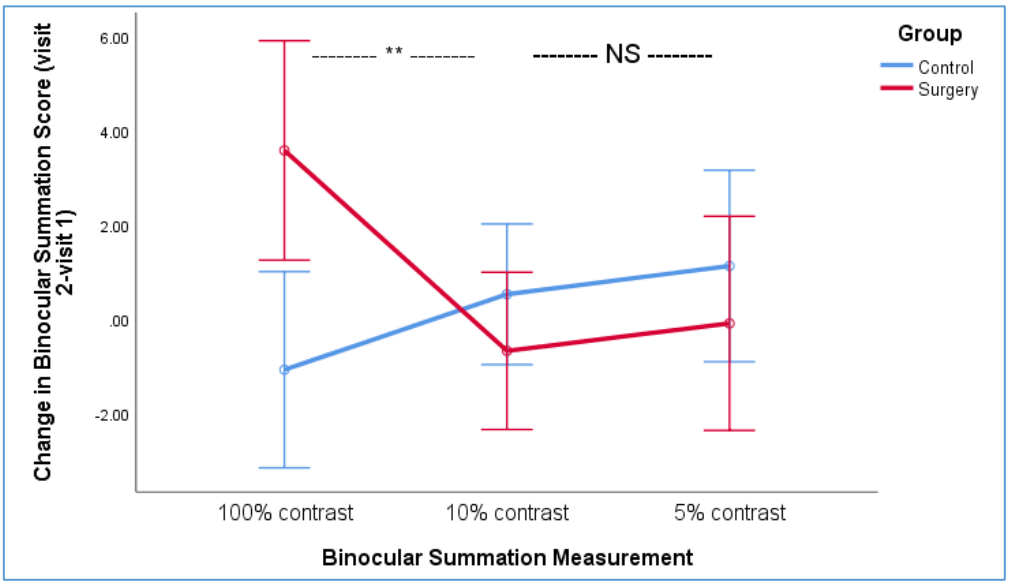


Figure 9-15. Mean (±2SE) change in binocular summation score, from visit one to visit two, in the surgery and control groups at 100%, 10% and 5% contrast.

Control group shown in blue. Surgery group shown in red.
Two-way mixed ANOVA results shown. Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01).

Two-way mixed ANOVA

The main effects of group and contrast level were not significant (p>0.05). The contrast level x group interaction was significant F(2,50)=5.16, p=0.009, r=0.31, indicating that the change in binocular summation score at the different levels of contrast was different in the control group and

the surgery group. Planned contrasts compared the change in binocular summation scores at different contrast levels. There was no main effect of contrast level for any of the comparisons (100% to 10%, or 10% to 5%) ($p>0.05$). The contrast level x group interaction was significant only when comparing the 100% and 10% levels of contrast $F(1,25)=10.12$, $p=0.004$, $r=0.54$. The contrast level x group interaction was not significant when the 10% and 5% levels of contrast were compared, $F(1,25)=0$, $p=0.99$, $r=0$ (shown in Figure 9-15, black text).

Binocular summation – categorisation method

The categorisation method used the binocular summation score to categorise individuals as having binocular summation (score ≥ +5), binocular indeterminate (score -4 to +4) and binocular inhibition (score ≤ -5). The number of participants in the surgery group (n=12) and the control group (n=15) categorised as having binocular summation, binocular indeterminate and binocular inhibition scores, at visit one and visit two, are displayed in histograms in Figure 9-16. The upper histogram shows the results at 100% contrast, the middle histogram shows the results at 10% contrast and the lower histogram shows the results at 5% contrast.

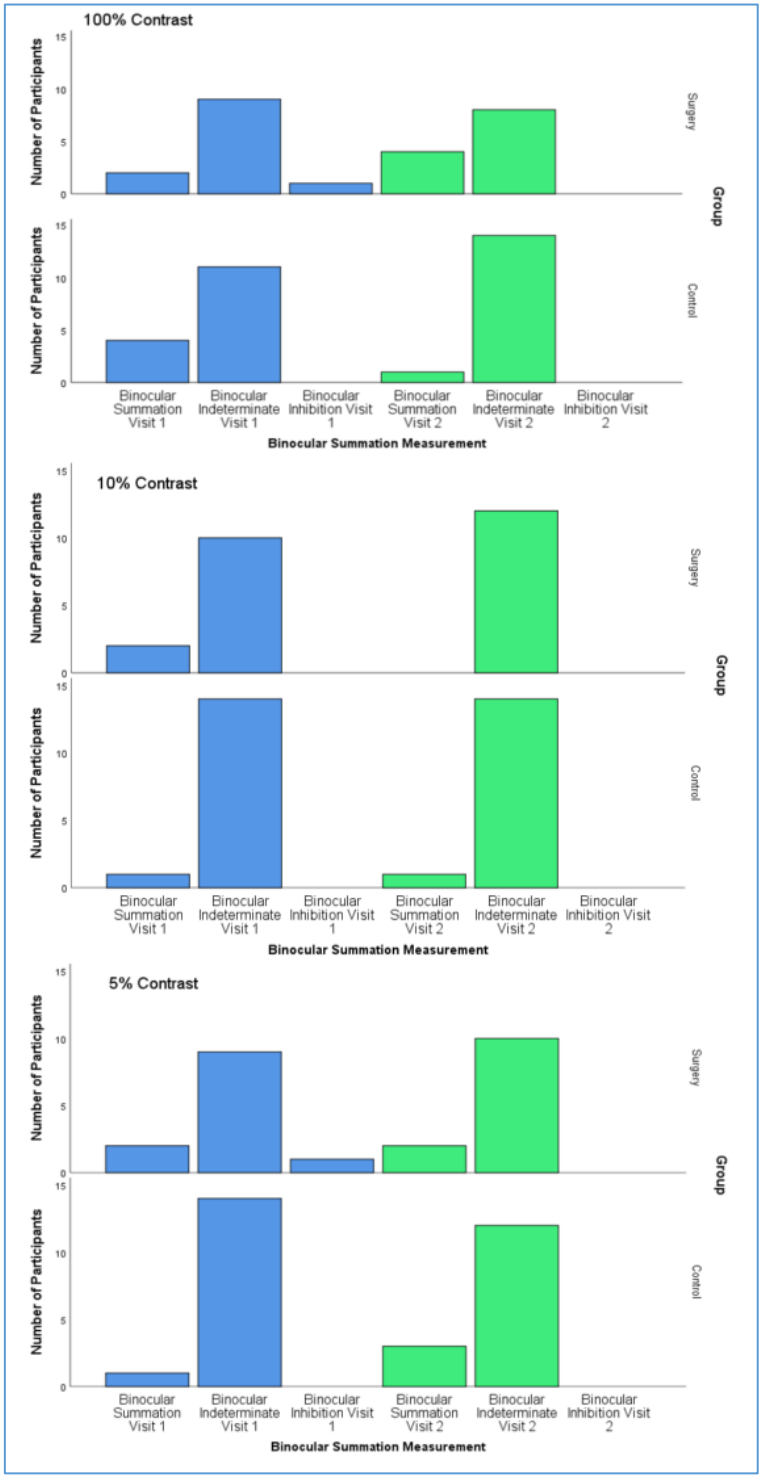


Figure 9-16. Histogram showing the number of participants with binocular summation, binocular indeterminate and binocular inhibition at each level of contrast, at visit one and visit two, for both groups.

100% contrast (upper histogram), 10% contrast (middle histogram) and 5% contrast (lower histogram). Visit one results shown in blue. Visit two results shown in green.

Summary of binocular summation results

At visit one there was no difference in binocular summation score between the groups. Strabismus surgery improved binocular summation at 100% contrast which equated to improved performance

both eyes open, compared to monocular performance with the better eye. Strabismus surgery reduced binocular summation at 10% and 5%, but this was not statistically significant. Postoperatively the surgery group had significantly better binocular summation than the control group at 100% contrast (Figure 9-14). When the change in binocular summation was analysed, there was a significant interaction between contrast level and group between 100% and 10% contrast. The surgery group had an increase in binocular summation (or a decrease in binocular inhibition) at 100% and a decrease in binocular summation (or an increase in binocular inhibition) at 10%. The control group had a decrease in binocular summation (or an increase in binocular inhibition) at 100% and an increase in binocular summation (or a decrease in binocular inhibition) at 10% (Figures 9-15). These results supported the hypothesis of an effect of surgery on binocular summation at 100% contrast, but not at 10% or 5% contrast.

9.4.5 Contrast sensitivity

Contrast sensitivity was measured in all participants (n=27) at visit one and visit two. The median (IQR) log contrast sensitivity score results for the surgery group (n=12) and control group (n=15) are shown in Figure 9-17. On inspection of Figure 9-17 there appeared to be little difference between the control group and the surgery group at visit one. From visit one to visit two the surgery group had minimal change in contrast sensitivity and the control group had a small improvement in contrast sensitivity. At visit two the control group appeared to have a slightly better contrast sensitivity.

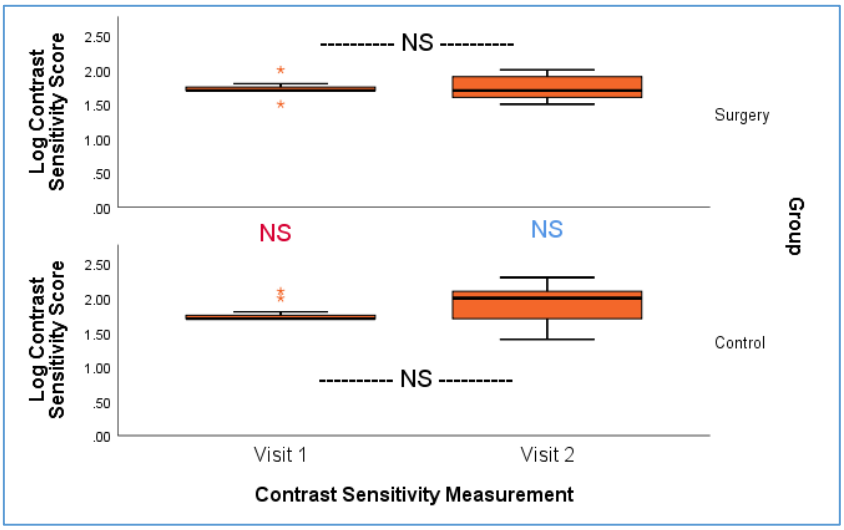


Figure 9-17. Boxplot showing the log contrast sensitivity scores, at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of contrast sensitivity results

At visit one the contrast sensitivity results in surgery group (Mdn=1.7) were not significantly different to the control group (Mdn=1.7), (p=0.68) (shown in Figure 9-17, red text). From visit one to visit two there was no significant change in the contrast sensitivity results in either the control group or the surgery group (p>0.05) (shown in Figure 9-17, black text). At visit two, the contrast sensitivity threshold results in surgery group (Mdn=1.7) were not significantly different to the control group (Mdn=2.0), (p=0.11) (shown in Figure 9-17, blue text).

Summary of contrast sensitivity results

There was no difference in contrast sensitivity between the surgery and control groups at either visit. Surgery made no difference to contrast sensitivity. The hypothesis of an effect of surgery on contrast sensitivity was rejected.

9.4.6 Visual field both eyes open

The visual field was plotted in all participants (n=27) at visit one and visit two. In the control group (n=15), the horizontal deviation was ET (n=10) and XT (n=5) at visit one and visit two. In the surgery group, all had an XT at visit one (n=12) and at visit two, six had an ET and six had an XT or X (see Table 9-1). The median (IQR) results of the visual field area for the surgery group and control group are shown in Figure 9-18. The visual field area monocularly appeared smaller than both eyes open at visit one, in both groups. The both eyes open visual field area appeared similar in both groups at visit one and visit two, with no obvious change within the groups.

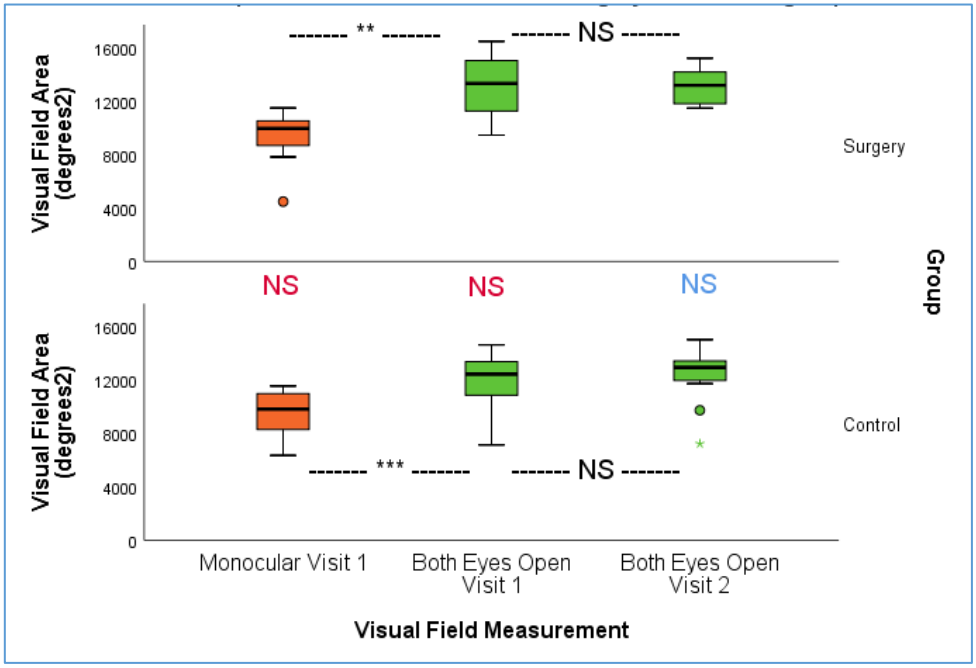


Figure 9-18. Boxplot showing the visual field area (degrees^2) tested monocularly at visit one and both eyes open at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01), *** = significant (p≤0.001).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of the total visual field area (degrees^2)

The total visual field area (degrees^2) was analysed non-parametrically as the √visual field area (degrees) remained non-normally distributed.

Total visual field area both eyes open

Visit one

The visual field area both eyes open and monocularly were not significantly different in the control group and the surgery group (p>0.05), shown in Figure 9-18 (red text).

As the surgery group had only XT at visit one, the groups were compared using XT only (control group n=5, surgery group n=12). In those with XT, the visual field area both eyes open remained not significantly different in the control group and the surgery group (p>0.05). When all participants were combined, visual field area both eyes open at visit one was larger in XT (Mdn=13399.47 degrees^2) than ET (Mdn=12482.29 degrees^2), but the difference was not significant (p>0.05).

Visit one to visit two

Visual field area, both eyes open, at visit one was not significantly different to visit two in the control group ($p>0.05$) or in the surgery group ($p>0.05$) (shown in Figure 9-18, black text).
The surgery group (n=12) all had surgery for XT, postoperatively the deviations were ET (n=6) and XT (n=6). When the surgery group were subdivided by deviation at visit two, there remained no significant difference between the visual field area at visit one and visit two in those with ET and those with XT postoperatively ($p>0.05$).

Visit two

The visual field area in the control group and the surgery group were not significantly different ($p>0.05$), shown in Figure 9-18 (blue text).

Visual field size monocularly compared to both eyes open

At visit one, when all participants were combined, the visual field area both eyes open (Mdn=12655.20 degrees^2) was significantly larger than monocularly (Mdn=9853.47 degrees^2), T=0, $p<0.0001$, r=0.87 (Wilcoxon signed rank test).

When the groups were analysed separately, the visual field area for both eyes open remained significantly larger than monocularly in the control group (both eyes open Mdn=12486.45 degrees^2, monocular Md=9853.47 degrees^2, T=0, p=0.001, r=0.88) and surgery group (both eyes open Mdn=13361.97 degrees^2, monocular Mdn=9984.55 degrees^2, T=0, p=0.002, r=0.88). Wilcoxon signed rank test, shown on Figure 9-18 (black text). When participants were split by deviation type (ET and XT) the visual field area both eyes open remained significantly larger than monocularly (ET, $p<0.01$; XT, $p<0.0001$).

Visual field size in esotropia compared to exotropia

There was no significant difference in the visual field area, measured both eyes open at visit one, between ET and XT ($p>0.05$).

Visual field measured by horizontal meridian

Parametric analysis (two-way mixed ANOVA) was used to analyse the horizontal meridian of the visual field data (degrees). The results from the analysis of the visual field size measured by horizontal meridian confirm the findings of the analysis of the visual field total area. There was a significant difference in the visual field size between monocular and both eyes open viewing at visit one. There was no significant main effect of group or visit and no significant interaction group x visit ($p>0.05$).

Summary of visual field results

Using both measures of the size of the visual field (total area (degrees^2) and horizontal meridian (degrees)), there was no significant difference between the control group and the surgery group at

visit one, from visit one to visit two, or at visit two. Surgery to reduce XT did not significantly enlarge or reduce the size of the visual field. The hypothesis of an effect of surgery on visual field size was rejected. In all participants the size of the visual field both eyes open was significantly larger than monocularly (Figure 9-18), regardless of whether they were in the control or the surgery group, and regardless of their deviation (ET or XT). The strabismic eye contributed to both eyes open viewing, by significantly enlarging the area of peripheral vision, including across the horizontal meridian. However, the visual field size was not significantly different in the ET and XT at visit one.

9.4.7 Gross binocular single vision (BSV) or stereopsis

Using Bagolini glasses, none of the control group (n=15) demonstrated evidence of BSV at visit one or visit two, as all had suppression. The surgery group (n=12) all had suppression and no potential BSV at visit one. Three participants had evidence of BSV at visit two, as described in section 9.1. The CST was used to measure both eyes open and monocular performance. The median (IQR) results are presented for CST accuracy, response time and the rate correct score (RCS) in the surgery group and the control group at both visits in Figure 9-19. Inspection of Figure 9-19 suggests that both eyes open performance was slightly better than monocular, but there may have been a ceiling effect for CST accuracy as many of the results were around the maximum result of 10. Faster response time appeared with repeated performance at visit two, although this was apparent for monocular as well as both eyes open performance. RCS appeared to improve in the surgery group more than the control group and both eyes open performance appeared better than monocular.

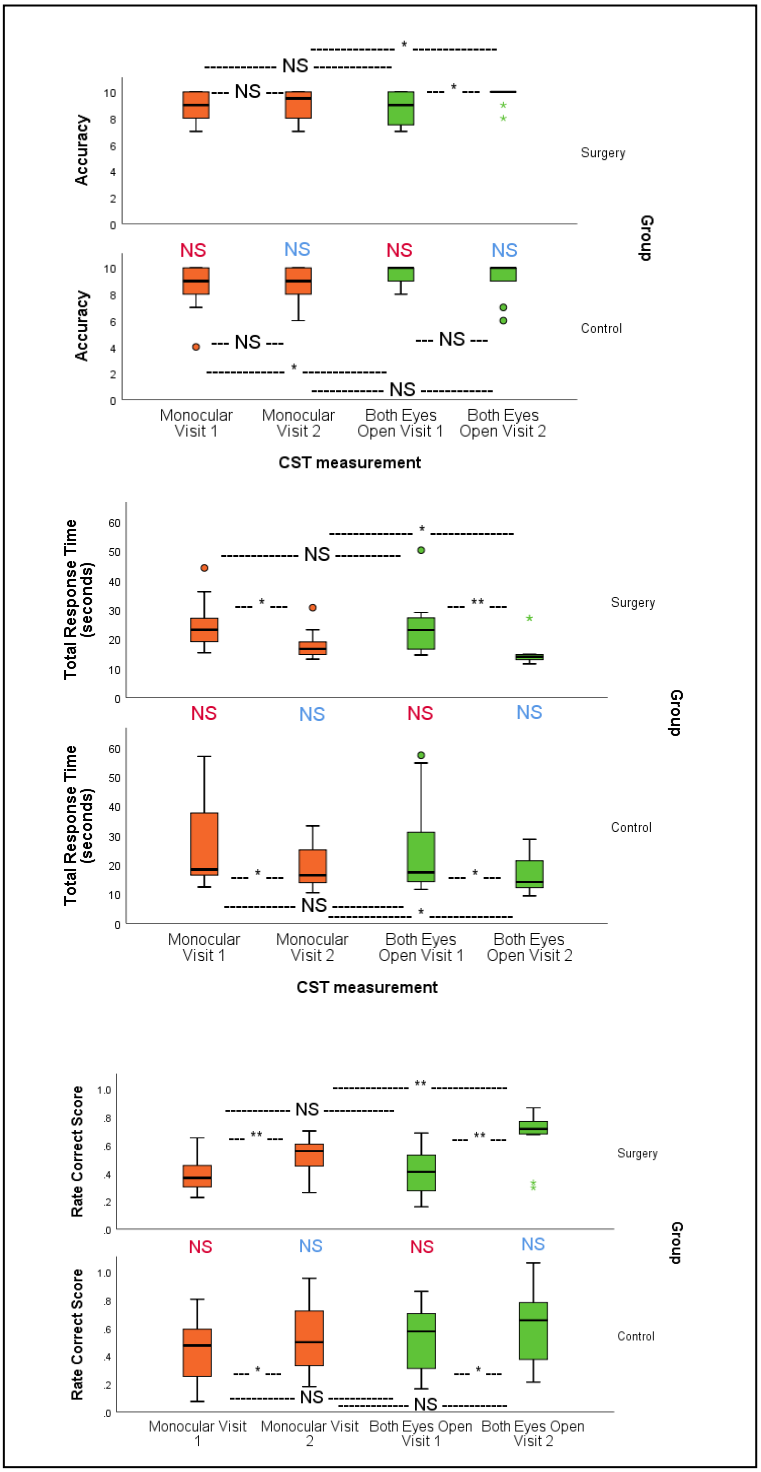


Figure 9-19. Boxplot showing the CST results at visit one and visit two, for both groups.

Accuracy shown in the top figure, response time shown in the middle figure and RCS shown in the bottom figure.

Statistical significance reported as NS = not significantly different ($p>0.05$), * = significant ($p\leq0.05$), ** = significant ($p\leq0.01$).

Text in black = within groups analysis (Wilcoxon signed rank test).

Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of the CST results

Visit one

At visit one, performance of the CST (both eyes open and monocularly) was not significantly different between the groups for any of the measures ($p>0.05$). Independent samples Mann-Whitney U test results are shown in Figure 9-19 (red text). At visit one there was no significant difference between both eyes open and monocular performance of the CST ($p>0.05$), except in the control group where the accuracy was significantly better both eyes open than monocularly, $T=2.5$, $p=0.048$, $r=0.36$. Wilcoxon signed rank test, shown on Figure 9-19 (black text).

Visit one to visit two

Both eyes open performance of the CST improved in the surgery group. Accuracy (Mdn visit one=9, visit two=10), response time (Mdn visit one=23.08, visit two=13.97) and RCS (Mdn visit one=0.65, visit two=0.72) all improved significantly ($p<0.05$). Both eyes open performance improved in the control group, but only in response time, ($p=0.02$) and RCS ($p=0.02$), not accuracy ($p=0.16$). Monocular performance of the CST improved in both the surgery group and the control group, in both response time and RCS ($p<0.05$). Accuracy, monocularly, did not improve in either group ($p>0.05$). These within groups results (Wilcoxon signed rank test) are shown in Figure 9-19 (black text).

Visit two

At visit two, both eyes open and monocular performance of the CST was not significantly different between the groups for any of the measures ($p>0.05$). Independent samples Mann-Whitney U test results are shown in Figure 9-19 (blue text). At visit two, the surgery group had better both eyes open than monocular performance of the CST, using all of the measures ($p<0.05$). Wilcoxon signed rank test, shown on Figure 9-22 (black text). The control group only had better both eyes open than monocular performance, when measured by response time ($p=0.047$), shown on Figure 9-19 (black text).

Change in CST results

To visualise the change in all the different measures of CST performance on the same scale, the change in each measure was calculated as a ratio (visit two result / visit one result) and are shown in Figure 9-20. A change of 1 equated to no change from visit one to visit two. Inspection of Figure 9-20 showed that accuracy had very little change but did improve slightly with both eyes open in the surgery group. Response time appeared to improve for monocular and both eyes open testing in both groups. The RCS seemed to improve in both groups, but the largest improvement was apparent in the surgery group with both eyes open.

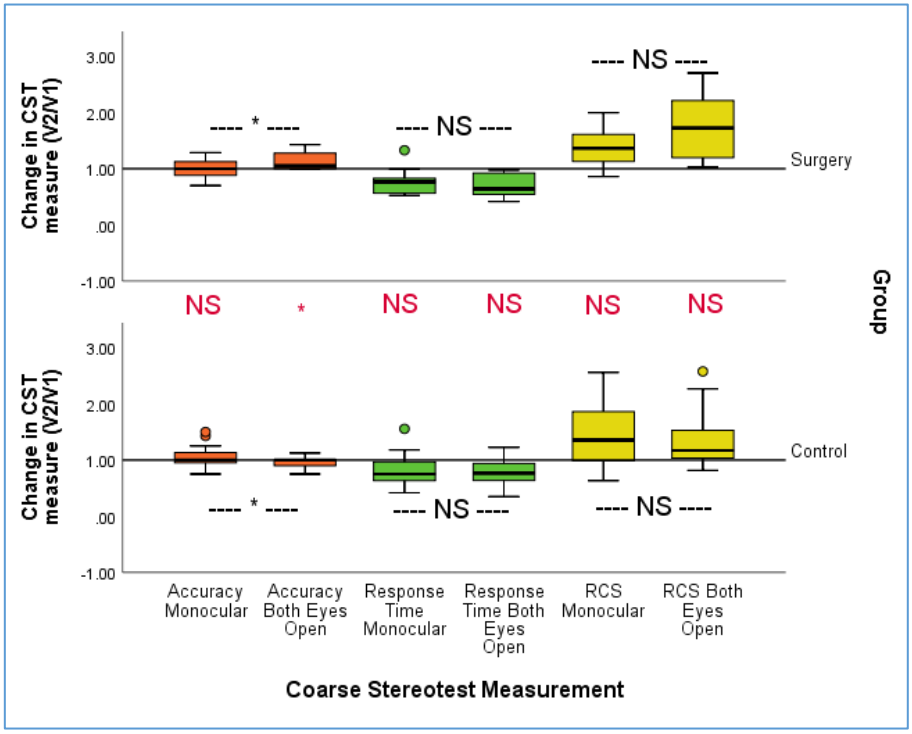


Figure 9-20. Boxplot showing the change in CST performance from visit one to visit two as a ratio, for both groups. CST accuracy, response time and RCS are shown for both monocular and both eyes open viewing.

Change=1.0 reflects the same result at visit one and visit two.
Change >1 reflects a larger / higher result at visit two (better performance for accuracy and RCS, but worse performance for response time).
Change <1 reflects a smaller / lower results at visit two (worse performance for accuracy and RCS, but better performance for response time).
Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05).
Between groups analysis shown in red text (Independent samples Mann-Whitney U test). Within groups analysis shown in black text (Wilcoxon signed rank test).

Non-parametric analysis of the change in CST results

The change in CST response time and RCS were not significantly different for both eyes open and monocular performance, in either the control group or the surgery group (p>0.05). The change in accuracy for both eyes open and monocular performance was significantly different in both the control group (p=0.03) and the surgery group (p=0.028). In the control group accuracy improved monocularly more than both eyes open (both eyes open Mdn=1.0, monocular Mdn=1.0), T=40.5, z=2.14, p=0.03, r=0.40. In the surgery group accuracy improved both eyes open more than monocularly (both eyes open Mdn=1.06, monocular Mdn=1.0), T=4, z=2.20, p=0.028, r=0.45. These within group results (Wilcoxon signed rank test) are shown in Figure 9-23 (black text).

There was no significant difference in the change in CST measures between the control group and the surgery group in any of the CST results ($p>0.05$), except the accuracy both eyes open. The change in accuracy both eyes open was significantly higher in the surgery group (Mdn=1.06) than in the control group (Mdn=1.0), $U=140.5$, $p=0.01$, $r=0.51$. These between groups results (Independent samples Mann-Whitney U test) are shown in Figure 9-20 (red text).

Summary of CST results

The CST was used as a measure of gross BSV. At visit one there was no difference between the groups (Figure 9-20, red text), and no difference between both eyes open and monocular performance in the surgery group. The control group had slightly (but significantly) better accuracy both eyes open than monocularly (Figure 9-20, black text). A large learning effect was evident. Performance across both groups improved significantly at visit two, particularly in response time and RCS (Figure 9-20, black text). From visit one to visit two, the amount of change in the measures was similar for monocular and both eyes open performance, for all measures of the task. However, there was a slight, but significant, increase in accuracy in the surgery group (Figure 9-20, red text). At visit two there was no difference in performance between the groups (Figure 9-20, blue text). When comparing monocular and both eyes open performance within each group at visit two, the control group has faster response time with both eyes open and the surgery group had better performance with both eyes open in all of the measures (Figure 9-20, black text). There was a significant improvement in CST results, just from repeating the test at visit two, even without surgery. Performance was not significantly different between the groups at visit one or visit two. Despite this, the surgery group had significantly better both eyes open performance of the task after surgery. These results supported the hypothesis of an effect of surgery on CST performance.

9.4.8 Sequential stereopsis

Sequential stereopsis was recorded in all participants in the control group (n=15) and the surgery group (n=12). The sequential stereotest results are presented as the mean absolute error percentage (%) for the unfiltered stimuli and the high pass filtered stimuli, under both eyes open and monocular viewing conditions, at both visit one and visit two (Figure 9-21). On inspection of Figure 9-21 both groups appeared to have slightly higher mean absolute percentage error (worse performance) at visit two. Performance with the high pass filtered stimuli appeared slightly worse than the unfiltered stimuli. Performance appeared similar for monocular and both eyes open testing.

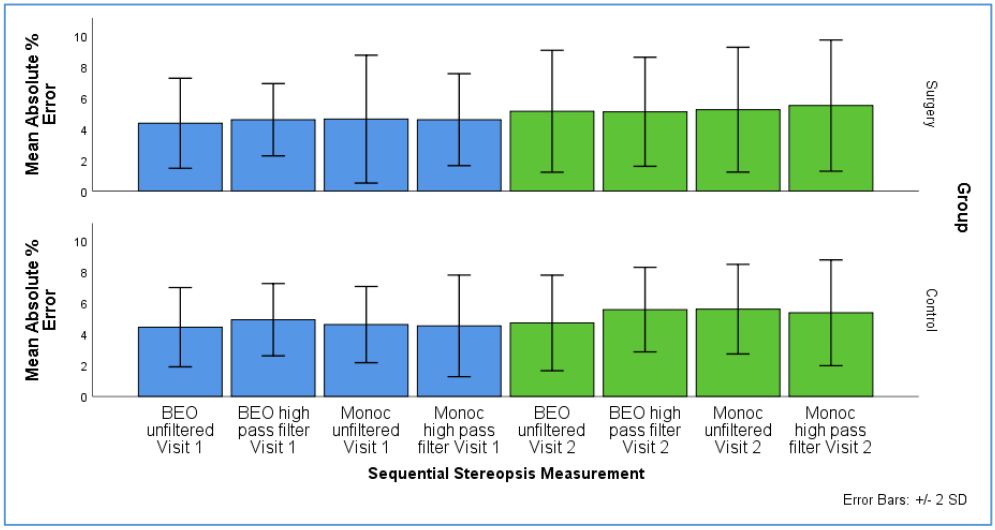


Figure 9-21. Bar chart of the sequential stereopsis mean absolute percentage error results at visit one and visit two, for both groups. Both eyes open and monocular viewing results are shown for both the unfiltered and the high pass filtered stimuli.

BEO = both eyes open viewing, Monoc = monocular viewing. Visit one results shown in blue. Visit two results shown in green. Significance not shown on the figure, none of the factors were significant.

Parametric analysis of the sequential stereopsis results

Two-way mixed ANOVA

There was a significant main effect of visit, $F(1,25)=8.27$, $p=0.008$, $r=0.50$, as the mean absolute error percentage increased from visit one (M=4.58) to visit two (M=5.27). Performance of sequential stereopsis task worsened significantly at visit two, across both groups. None of the other main effects (group, stimulus or viewing condition) or any of the interactions were significant ($p>0.05$).

ANCOVA

Using the covariate, baseline size of deviation at visit one, in the ANCOVA, none of the main effects or interactions were significant ($p>0.05$). The main effect of visit was no longer significant ($p=0.09$).

Summary of the sequential stereopsis results

Strabismus surgery did not affect sequential stereopsis performance. The hypothesis of an effect of surgery on sequential stereopsis performance was rejected. Performance of the sequential stereopsis task worsened from visit one to visit two in all participants (two-way mixed ANOVA), however this was no longer significant when the covariate was included (ANCOVA). Both eyes open and monocular viewing did not affect sequential stereopsis performance. The different stimuli had no effect on performance (Figure 9-21).

9.5 Task performance

9.5.1 Grooved pegboard

Twenty-seven participants completed the grooved pegboard at visit one. One participant (018) was unable to complete the task at visit two due to suffering with joint pain secondary to arthritis. All data from participant 018 were removed prior to analysis, leaving twenty-six remaining participants (control group (n=15) and surgery group (n=11)). The median (IQR) time to complete the grooved pegboard (seconds) in bright and dim lighting conditions, in both the surgery and the control groups, at visit one and visit two are shown in Figure 9-22. On inspection of Figure 9-22 there appeared to be no difference between the groups at either visit. The appeared to be little difference in performance from visit one to visit two in either group. Performance seemed better in the bright lighting condition compared to the dim lighting condition, however this difference was more apparent in the control group.

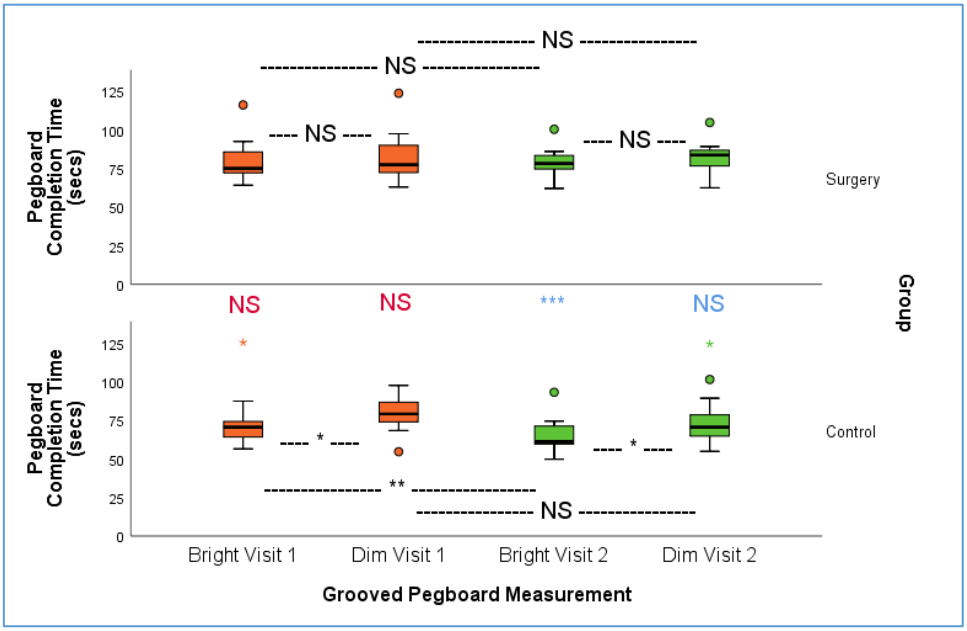


Figure 9-22. Boxplot showing the grooved pegboard completion time (seconds) under bright and dim lighting conditions, at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of the grooved pegboard results

Visit one

There was no significant difference in grooved pegboard performance between the surgery group and control group, in bright or dim lighting (p>0.05). These between groups results are shown in

Figure 9-22 (red text). Grooved pegboard completion time was faster in bright light compared to dim lighting. With all participants combined, they were on average 5.72 seconds faster, T=275, p=0.012, r=0.50. Spilt by group, the control group were on average 8.48 seconds faster in bright light, T=103, p=0.015, r=0.45, and the surgery group were on average 2.40 seconds faster in bright light, T=44, p=0.33, r=0.21. These within group results are shown in Figure 9-22 (black text).

Visit one to visit two

From visit one to visit two, in bright light, the control group significantly improved their grooved pegboard completion time, by 9.51 seconds on average (T=14, p=0.009, r=0.48), and the surgery group worsened their performance by 3.12 seconds on average, but this was not significant (T=34, p=0.93, r=0.02). In dim lighting, the differences were also not significant. The control group were on average 8.49 seconds faster at visit two (T=29, p=0.08, r=0.32) and the surgery group were on average 6.25 seconds slower at visit two (T=38, p=0.66, r=0.09). These within groups results are shown in Figure 9-22 (black text).

Visit two

At visit two, in bright lighting, the control group were significantly faster (on average 17.2 seconds faster) than the surgery group (U=143, z=3.14, p=0.001, r=0.62). In dim lighting the control group were on average 13.2 seconds faster than the surgery group, but this was not significant (U=119, z=1.89, p=0.06, r=0.37). These between groups results are shown in Figure 9-22 (blue text). At visit two, grooved pegboard completion time was significantly faster in bright light compared to dim lighting conditions. With all participants combined, pegboard completion was on average 3.85 seconds faster in bright light (T=307.5, p=0.001, r=0.47). When split by group, the control group were on average 9.5 seconds faster and this was significant (T=113, p=0.003, r=0.55). The surgery group were on average 5.5 seconds faster and this was not significant (T=52, p=0.09, r=0.36). These within groups results are shown in Figure 9-22 (black text).

Summary of grooved pegboard results

At visit one there was no difference between the control group and the surgery group in grooved pegboard performance, in either of the lighting conditions (Figure 9-22, red text). Performance was better under bright lighting conditions compared to dim lighting conditions, for both groups at visit one, although this was only significant in the control group. From visit one to visit two grooved pegboard completion time worsened slightly, but not significantly, in the surgery group, under both lighting conditions. In comparison the control group showed significantly improved performance in bright light, or a practice effect (Figure 9-22, black text). At visit two the surgery group performed the grooved pegboard worse than the control group, however this was only significant for the bright lighting condition (Figure 9-22, blue text). At visit two performance under bright conditions remained significantly better than dim conditions for the control group, but the difference between performance under bright and dim conditions was not significant for the surgery group. Strabismus

surgery led to worsened grooved pegboard performance in the surgery group, particularly in bright lighting. The results support the hypothesis of an effect of surgery, but a negative effect, or worsening of grooved pegboard performance.

9.5.2 Purdue pegboard

The Purdue pegboard was completed by eighteen participants at visit one. Participant 018 was unable to complete the task at visit two due to suffering with joint pain secondary to arthritis. All data from participant 018 were removed prior to analysis. Of the remaining participants (n=17), eight were in the control group and nine were in the surgery group. The median (IQR) peg insertion task score (number of pegs inserted in 30 seconds) and the assembly task score (number of assemblies completed in 60 seconds) for both groups at both visits are shown in Figure 9-23. On inspection of Figure 9-23, the peg insertion score appeared similar in the surgery group and the control group, and there appeared to be little change from visit one to visit two in either group. The assembly task score appeared higher in the control group. The control group appeared to have similar performance at visit one and visit two. The surgery group appeared to have slightly lower scores (worse performance) at visit two compared to visit one.

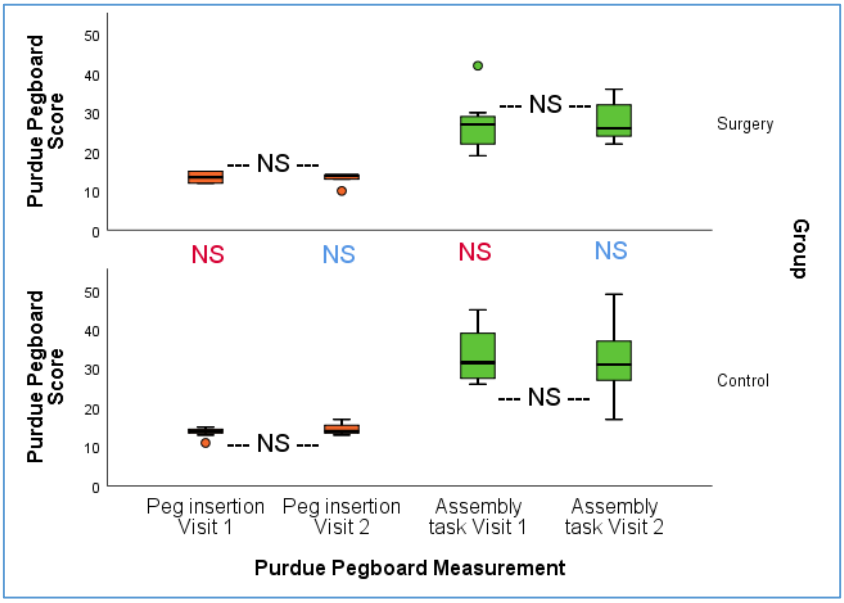


Figure 9-23. Boxplot showing the Purdue pegboard pin insertion and assembly task scores at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of the Purdue pegboard results

At visit one there was no significant difference between the control group and the surgery group, in either the peg insertion or the assembly task ($p>0.05$) (between groups results shown in Figure 9-23, red text). From visit one to visit two there was no significant difference in the peg insertion and assembly task, in either group ($p>0.05$). These within groups results are shown in Figure 9-23 (black text). At visit two there was no significant difference between the control group and the surgery group, in either the peg insertion or the assembly task ($p>0.05$) (between groups results shown in Figure 9-23, blue text).

Summary of Purdue pegboard results

Surgery did not affect Purdue peg insertion or assembly performance. The hypothesis of an effect of surgery was rejected. There was no significant difference between the groups, on either of the Purdue pegboard tasks, at either visit one or visit two. Performance did not change significantly from visit one to visit two.

9.5.3 Bead threading

Bead threading time (seconds) was measured for large and small beads. Eighteen participants completed bead threading at visit one, but one participant (018) was unable to complete the task at visit two due to suffering with joint pain secondary to arthritis. All data from participant 018 were removed prior to analysis. Of the remaining participants (n=17), eight were in the control group and nine were in the surgery group. The median (IQR) bead threading time (seconds) are displayed in Figure 9-24. Bead threading time appeared shorter (faster) for the large beads than for the small beads, in both groups and at both visits. Both groups appeared to have similar performance with both sizes of beads, at both visits. From visit one to visit two the control group appeared to have a slight improvement in bead threading from visit one to visit two, but the surgery group appeared to have similar performance.

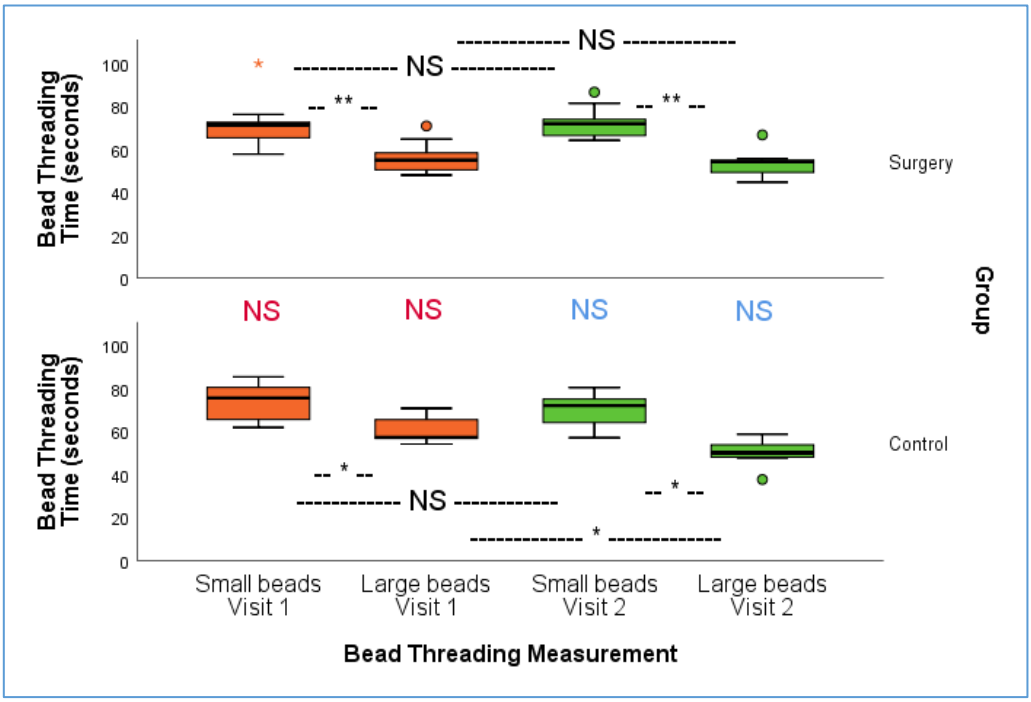


Figure 9-24. Boxplot showing the bead threading time (seconds) for small and large beads, at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of bead threading time

Visit one

There was no significant difference between the surgery group and the control group at visit one, for the large or small beads (p>0.05). These between groups results are shown in Figure 9-24 (red text). The time to thread the large beads was significantly faster than the small beads, in the control group (Mdn time large=57.41, small=75.72) (T=36, p=0.01, r=0.63) and the surgery group (Mdn time large=54.90, small=71.37) (T=45, p=0.008, r=0.63). These within groups results are shown in Figure 9-24 (black text).

Visit one to visit two

Large bead threading in the control group was on average 7.26 seconds faster (T=0, p=0.01, r=0.63) and in the surgery group was on average 0.77 seconds faster (T=11, p=0.17, r=0.32) at visit two compared to visit one. Small bead threading in the control group was on average 3.55 seconds faster and in the surgery group was on average 0.52 seconds faster at visit two, but these

differences were not significant (p>0.05). These within groups results are shown in Figure 9-24 (black text).

Visit two

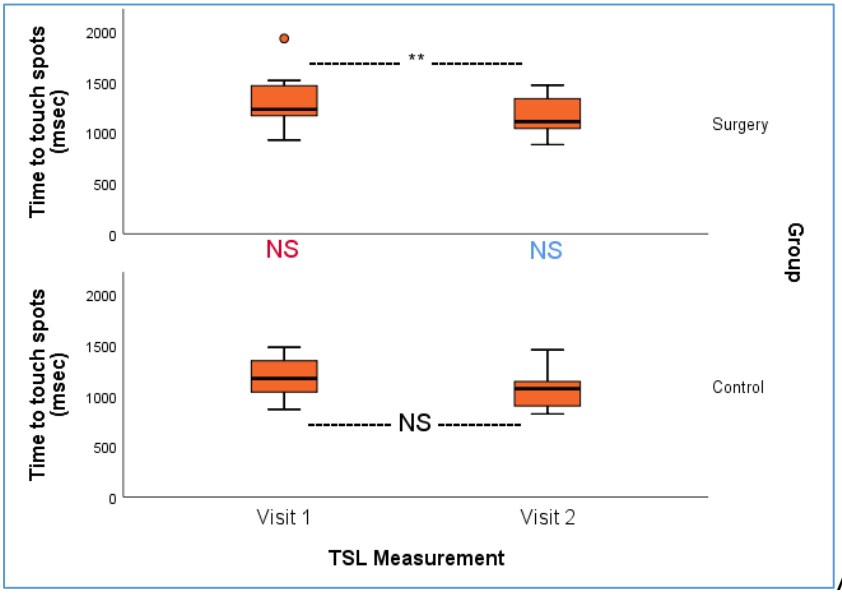
At visit two there was no significant difference between the groups, in large or small bead threading time (p>0.05). These results are shown in Figure 9-24 (blue text). Bead threading time remained significantly faster for large beads compared to small beads, in the control group (Mdn time large=50.15, small=72.17) (T=36, p=0.01, r=0.63) and in the surgery group (Mdn time large=54.13, small=71.89) (T=45, p=0.008, r=0.63). These within groups results are shown in Figure 9-24 (black text).

Summary of bead threading results

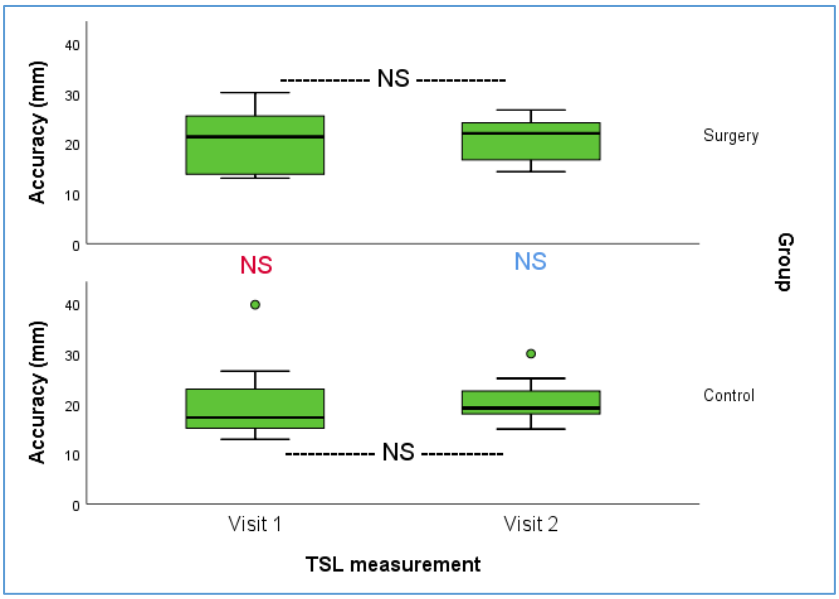
There was no significant difference in bead threading time between the control group and the surgery group, at visit one or visit two. There was a small practice effect such that bead threading time tended to improve at visit two, compared to visit one, however only large bead threading in the control group improved significantly. Bead threading was significantly faster with large beads than with small beads, at both visit one and visit two (Figure 9-24). The control group had a greater improvement in bead threading than the surgery group was interpreted as strabismus surgery slightly worsening bead threading performance. The hypothesis of an effect of surgery on bead threading was accepted, but the effect was worsened performance.

9.5.4 Touch screen spatial localisation (TSL) task

Seventeen participants completed the TSL task at visit one. Participant 018 was unable to complete the task at visit two due to suffering with joint pain secondary to arthritis. All data from participant 018 were removed prior to analysis. Of the remaining participants (n=16), eight were in the control group and eight were in the surgery group. The median (IQR) time taken to touch the spot (ms) and accuracy (mm), distance between the point touched on the screen and the centre of the spot (mm), are shown in Figures 9-25 A and B respectively. On inspection of Figure 9-25A the time to touch the spots (ms) in the control group and the surgery group appeared similar. The time to touch the spots appeared to improve at visit two, in both groups. On inspection of Figure 9-25B TSL accuracy (mm) appeared similar in both groups, with little change from visit one to visit two, in either group.



A



B

Figure 9-25 A and B. Boxplots showing the TSL results at visit one and visit two, for both groups. A: TSL time to touch the spots during the TSL task (ms). B: TSL accuracy (mm).

Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01). Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two. (Independent samples Mann-Whitney U test).

Non-parametric analysis of the TSL data (time and accuracy)

At visit one there was no significant difference between the surgery group and the control group in the TSL time or accuracy (p>0.05) (shown in Figure 9-25, red text). From visit one to visit two, a practice effect was evidence such that the time to touch the spots (ms) reduced and accuracy (mm) increased in both groups. The control group 101ms reduction was not significant (T=6, p=0.09, r=0.42) but the surgery group 120.8ms reduction was significant (T=0, p=0.01, r=0.63). TSL accuracy increased in both groups (control group = 0.02mm, surgery group = 0.25mm) but

these differences were not significant (p>0.05). These within groups results are shown in Figure 9-25 (black text).
At visit two there was no significant difference in TSL time or accuracy, between the surgery group and the control group (p>0.05). These between groups results are shown in Figure 9-25 (blue text).

Summary of TSL results

There was no significant difference in TSL task performance between the groups at visit one or visit two. The surgery group performed the TSL task significantly faster postoperatively, although this improvement was small (121ms) (Figure 9-25, black text). These results supported the hypothesis of an effect of surgery on TSL performance time, but not accuracy. This improvement may have been a practice effect or a small effect of strabismus surgery.

9.5.5 CKAT

The CKAT task was attempted in 20 participants, in both the control group (n=10) and the surgery group (n=10). The CKAT tasks included tracking (with guide and without guide), aiming and steering (using shape A and shape B) (Appendix R). Technical difficulties were experienced when using the CKAT, these are documented in Table W-1 (Appendix W). If technical difficulties prevented completion of the CKAT task at visit one, testing was not repeated at visit two. Data were removed if both visit one and visit two data was not available. There was no significant difference between the steering results for shape A and shape B, (paired samples t-test, $t(10)=0.66$, $p=0.52$, $r=0.20$) so data for the two shapes were combined into one value for steering (pPA).

Median (IQR) tracking results (RMSE mm) are shown in Figure 9-26. There appeared to be an accuracy, speed trade off such that tracking accuracy worsened as tracking target speed increased. The results without guide appeared similar to with guide. The control group results appeared similar to the surgery group, except the fast speed with guide at visit two, which appeared less accurate in the control group.

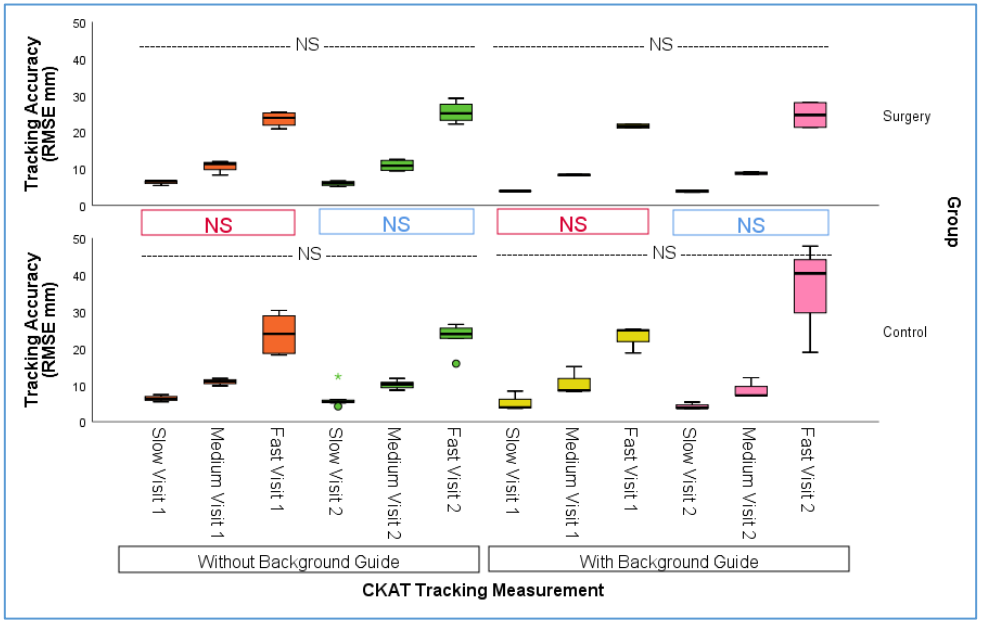


Figure 9-26. Boxplot showing the CKAT tracking accuracy results (RMSE mm) for slow, medium and fast target speeds, with and without the background guide, at visit one and visit two, for both groups

Statistical significance reported as NS = not significantly different (p>0.05).
Text in black = within groups analysis (Wilcoxon signed rank test) for each speed and background type.
Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two (Independent samples Mann-Whitney U test) for each speed.

Non-parametric analysis of the tracking data

Tracking

At visit one and visit two, there was no significant difference in the tracking accuracy results between the control group and the surgery group, with or without the guide, for the slow, medium or fast target speeds (p>0.05). These between group results are shown in Figure 9-26 (red text, visit one and blue text, visit two). There was no significant difference (p>0.05) in the visit one and visit two tracking accuracy results, in either of the groups, for any of the tracking measures (without guide or with guide; slow, medium or fast target speed). These within group results are shown in Figure 9-26 (black text). There was no significant difference in tracking accuracy with and without the background guide (p>0.05).

Tracking accuracy significantly reduced as target speed increased (Friedman test, $\chi^2(2)=58.07$, p<0.0001). Each of the planned comparisons was also significant, even after Bonferroni correction (for three comparisons). Tracking accuracy reduced as the target speed changed from slow to medium (z=3.62, p=0.001, r=0.66), slow to fast (z=7.62, p<0.0001, r=1.39) and from medium to fast (z=4.00, p<0.0001, r=0.73). Parametric analysis (three-way mixed ANOVA using the between subjects factor of group and the within subjects factors of visit and target speed) confirmed the non-parametric analysis, as only target speed was significant (p<0.0001) such that accuracy

decreased as target speed increased. Group and visit were not significant factors and there were no significant interactions (p>0.05).

Parametric analysis of the aiming and steering data

Aiming

Mean (SD) aiming results (mean path length time (seconds)) for the control group (n=5) and the surgery group (n=5) are shown in Figure 9-27. On inspection of Figure 9-27 there did not appear to be a difference between groups at visit one. From visit one to visit two the aiming results appeared similar in the surgery group, but slightly worse in the control group. At visit two, the aiming results appeared slightly better in the surgery group.

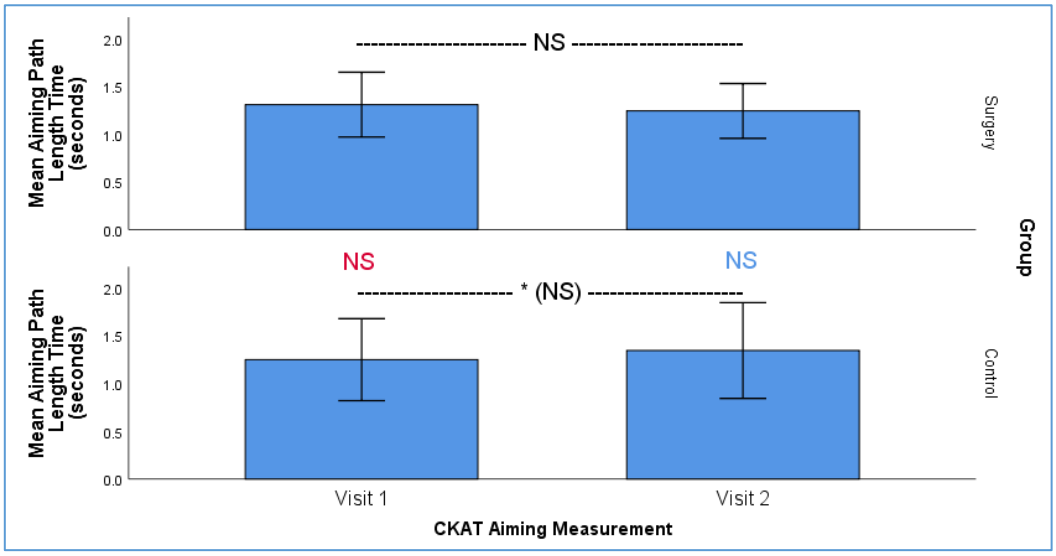


Figure 9-27. Bar chart of the CKAT mean aiming path length time (seconds) at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05). Results of the mixed ANOVA shown. Significance shown in brackets is after Bonferroni correction. Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Two-way mixed ANOVA

There was no significant main effect of visit (p=0.49) or group (p=0.86). There was a significant interaction visit x group (p=0.005), shown in Figure 9-27. Aiming improved in the surgery group (by 0.07 seconds) and worsened in the control group (by 0.10 seconds).

Planned comparisons showed that the difference in aiming between the groups at visit one and visit two were not significant (p>0.05). These between groups results are shown in Figure 9-27 in red and blue text respectively. The change in aiming from visit one to visit two for the surgery

group and control group were compared separately using a post-hoc paired samples t-test. Aiming in the control group significantly worsened from visit one (mean=1.25 seconds) to visit two (mean=1.35 seconds) (t(4)=3.37, p=0.028, r=0.86). Aiming in the surgery group was not significantly different from visit one (mean=1.31 seconds) to visit two (mean=1.24 seconds), (t(4)=2.16, p=0.10, r=0.73). This difference in the control group was no longer significant (p=0.056) after Bonferroni correction for two tests. These within groups results are shown in Figure 9-27 (black text).

Steering

Steering results (penalised path accuracy (pPA) score) for the control group (n=6) and surgery group (n=3) are shown in Figure 9-28. Inspection of Figure 9-28 revealed that both groups had similar performance of the steering task at visit one and visit two. Both groups appeared to have slightly improved steering task performance at visit two.

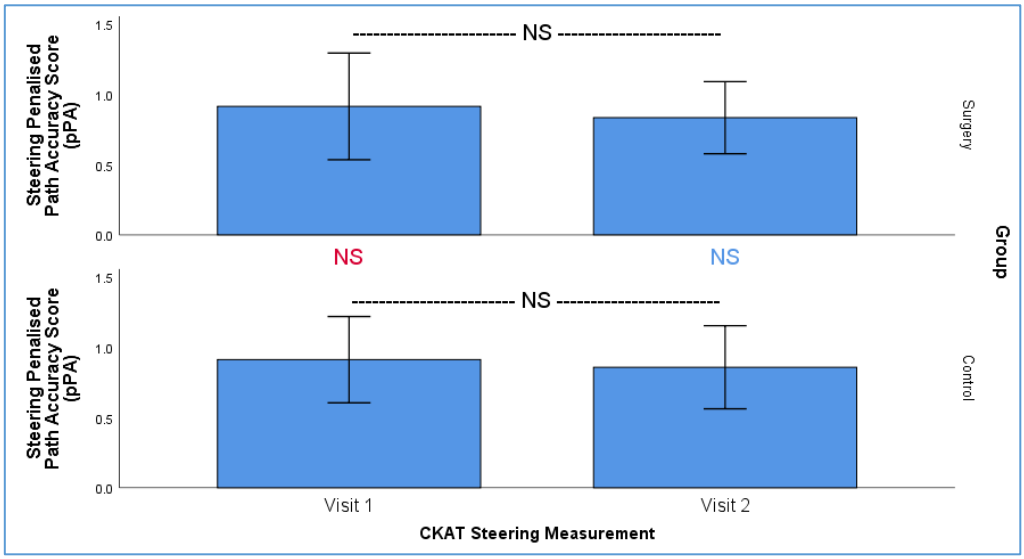


Figure 9-28. Bar chart showing the CKAT steering penalised path accuracy (pPa) scores at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different. Results of the mixed ANOVA shown. Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Two-way mixed ANOVA

There was no significant main effect of visit (p=0.25) or group (p=0.92). The two-way interaction visit x group was not significant (p=0.82). Both groups had slightly improved steering performance at visit two, compared to visit one, but this was not significant (p>0.05). The between groups results are shown in Figure 9-28 (red text, visit one and blue text, visit two). The within groups results are shown in Figure 9-28 (black text).

Summary of CKAT results

Technical difficulties using the CKAT limited the analysis of the CKAT data in both groups, particularly for steering. There was no difference in the tracking or the steering results in the two groups and no change in tracking or steering performance from visit one to visit two. The aiming results however showed there to be a significant interaction between group and visit ($p=0.005$) with a large effect size ($r=0.81$) (Figure 9-27). The surgery group improved their aiming performance from visit one to visit two (by 0.07 seconds), whilst the control group worsened (by 0.10 seconds). This difference, however, was no longer significant after Bonferroni correction for two post-hoc comparisons. Strabismus surgery had no effect on CKAT tracking and steering performance, but there was a possible improvement in aiming performance. The hypothesis of an effect of surgery on CKAT performance was rejected for tracking and steering but tentatively accepted for aiming.

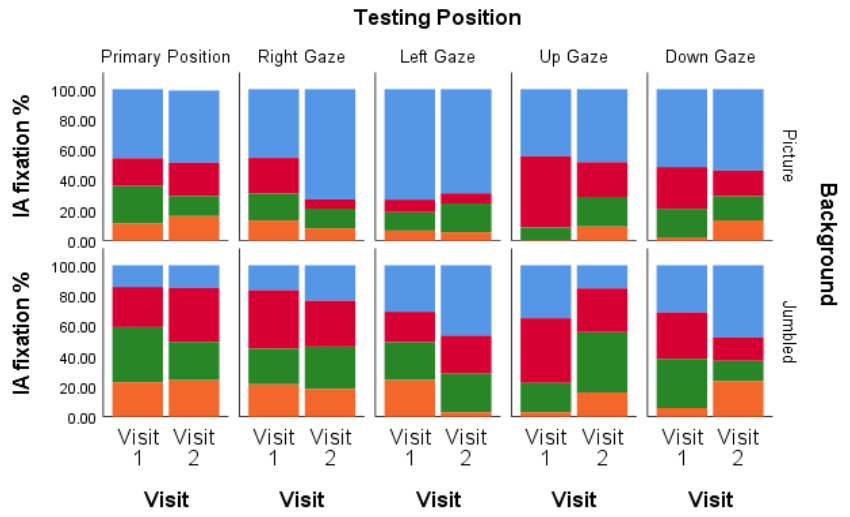
9.6 Eye movement recording

Eleven participants took part in the EMR (control group n=6 and surgery group n=5) (Appendix U). The hypotheses for the EMR tasks are described in section 8.5.1.

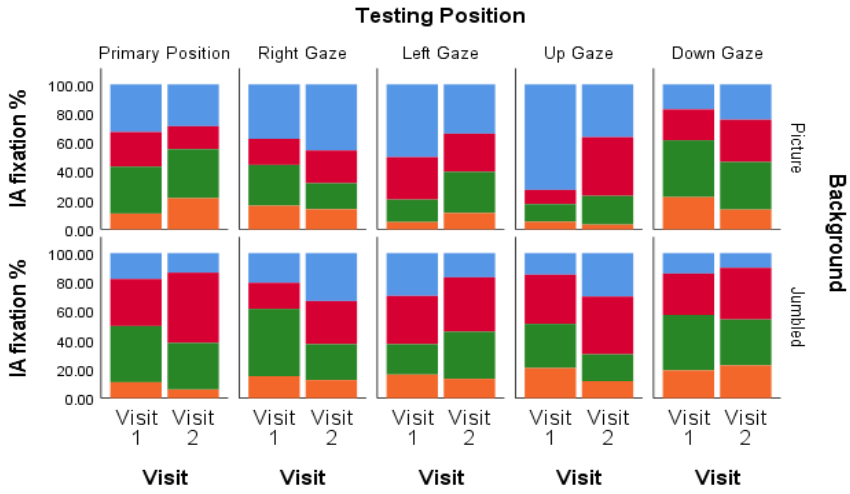
9.6.1 Fixation

Interest area fixation percentage results for the 0.5°, 1° and 2° diameter interest areas and outside the 2° interest area are shown in Figure 9-29. Figure 9-29 A (above) shows the surgery group results and B (below) shows the control group results. Fixation results are shown for each position of gaze (primary position, right, left, up and down gaze) and separately for the picture background (above) and jumbled background (below). The fixation results are shown as a stacked bar chart, with the sum of all fixation during the trial represented by each bar (100% fixation). Each bar is then divided up into orange, green, red and blue. Orange shows the percentage fixation falling within the 0.5° interest area. Green shows fixation within the 1° interest area (minus the 0.5° interest area). Red shows fixation within the 2° interest area (minus the 1° interest area). Blue shows fixation occurring on the screen, but outside the 2° interest area.

On inspection of Figure 9-29 fixation appeared better against the jumbled background compared to the picture background in both groups, as there was less fixation falling outside the interest area (blue bar). There did not appear to be a difference between the groups at either visit. Fixation performance from visit one to visit two appeared broadly similar.



A



B

Figure 9-29. Stacked bar chart showing the EMR IA fixation percentage in each of the five gaze positions, against the picture and jumbled backgrounds, at visit one and visit two, for the surgery group (A) and the control group (B)

Orange = fixation % within the inner 0.5° interest area.
Green = fixation % in the 1° interest area (minus the 0.5° interest area).
Red = fixation % in the 2° interest area (minus the 1° interest area)
Blue = fixation % occurring outside the 2° interest area.
The sum of the fixation percentage results for each of the interest areas and outside the 2° interest area = 100% (100% of fixation during the trial).

Non-parametric analysis of the fixation results

Interest area fixation percentage within the 2-degree interest area was analysed non-parametrically.

Visit one

There was no significant difference between the control group and the surgery group in the fixation task ($p>0.05$), except in downgaze against the picture background, where the control group (Mdn=85.31) had significantly better fixation than the surgery group (Mdn=55.56), ($U=1$, $z=2.56$, $p=0.009$, $r=0.77$). At visit one, there was no significant difference in fixation between the picture background and the jumbled background, in either the control group or the surgery group, for primary position, right gaze or downgaze ($p>0.05$). For fixation on left gaze, there was no significant difference in fixation between the picture background and the jumbled background, for the control group. In the surgery group, fixation was significantly better against the jumbled background (Mdn=71.43) compared to the picture background (Mdn=32) ($T=15$, $p=0.043$, $r=0.64$). For fixation on up gaze, there was no significant difference in fixation between the picture background and the jumbled background, for the surgery group. In the control group, fixation was significantly better against the jumbled background (Mdn=90.60) compared to the picture background (Mdn=18.56) ($T=21$, $p=0.028$, $r=0.63$).

Visit one to visit two

There was no significant difference between fixation performance at visit one and visit two in either the control group or the surgery group, for either the picture background or the jumbled background, and for any of the five positions of gaze ($p>0.05$).

Visit two

There was no significant difference between the control group and the surgery group in the fixation task in any of the positions ($p>0.05$), except in downgaze. In downgaze, the control group had significantly better fixation than the surgery group, against both backgrounds ($p<0.05$). At visit two, there was no significant difference in fixation between the picture background and the jumbled background, in either the control group or the surgery group, for up gaze ($p>0.05$). Fixation against the jumbled background was significantly better than against the picture background for the surgery group in primary position and on right gaze, and for the control group on left gaze and down gaze ($p<0.05$).

Summary of fixation results

Fixation ability did not change after strabismus surgery. The hypothesis of an effect of surgery on fixation was rejected. Using fixation within the 2-degree interest area as the measure of fixation performance, there was no significant difference between the control and surgery group at visit one, except for downgaze where the control group performed better than the surgery group against the picture background. The was no significant difference in fixation performance at visit two

compared to visit one. At visit two there was no significant difference in fixation performance in the control group and the surgery group in any of the positions of gaze, except again for downgaze, where fixation was significantly better in the control group against both backgrounds (Figure 9-29). At visit one and visit two, fixation with a jumbled background and a picture background was either not significantly different or was better with the jumbled background.

9.6.2 Smooth pursuit

Horizontal and vertical smooth pursuit results are presented separately and include the mean velocity gain (0-1), mean pursuit accuracy (RMSE in degrees of visual angle) and total length, or duration, of fixation (ms) during the trial. Eleven participants took part in the smooth pursuit recordings (Appendix U). Control participant 08 had their horizontal smooth pursuit data removed prior to analysis, due to them having difficulty performing the horizontal smooth pursuit task accurately. Surgery participant 09 had their vertical smooth pursuit data removed prior to analysis, due to technical difficulties experienced during the visit one recording. Data from ten participants are therefore reported for the horizontal smooth pursuit analysis (control n=5, surgery n=5) and the vertical smooth pursuit analysis (control n=6, surgery n=4).

Smooth pursuit - velocity gain

The mean (±2 SD) velocity gain for horizontal and vertical smooth pursuit, in the surgery group and the control group at visit one and visit two are shown in Figure 9-30. On inspection of Figure 9-30 vertical smooth pursuit velocity gain appeared worse than horizontal. During horizontal smooth pursuit, there appeared to be little difference in velocity gain between the groups at either visit and velocity gain appeared to reduce at visit two. During vertical smooth pursuit, velocity gain appeared similar in both groups at visit one. However, the surgery group appeared to have worse velocity gain at visit two, whilst the control group had better velocity gain at visit two.

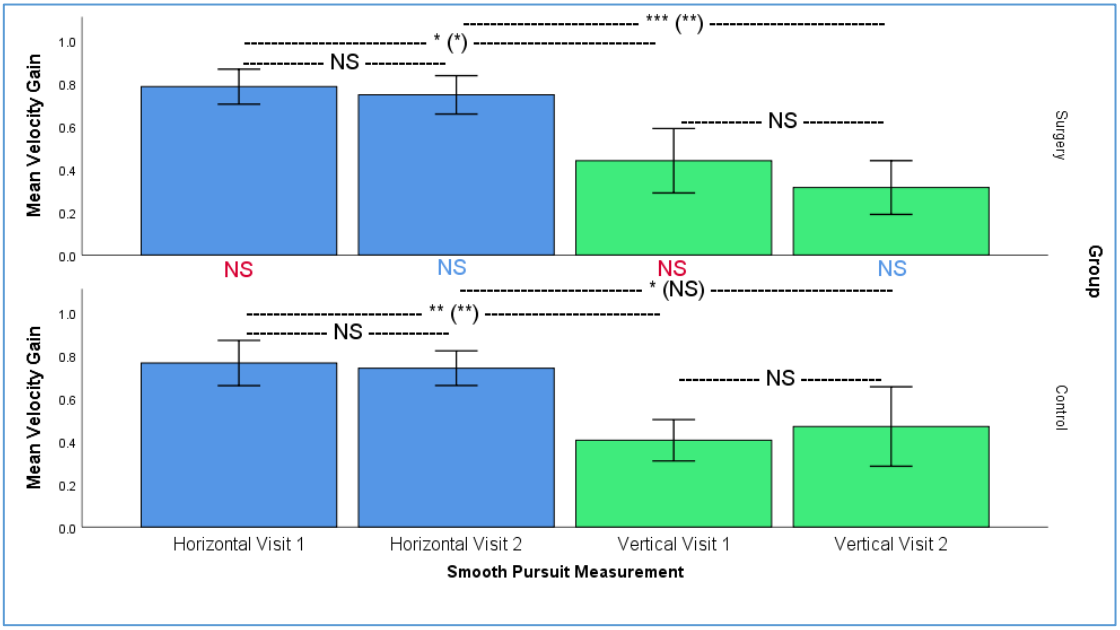


Figure 9-30. Bar chart showing the EMR smooth pursuit mean velocity gain for horizontal and vertical smooth pursuit, at visit one and visit two, for both groups.

Blue = horizontal smooth pursuit. Green = vertical smooth pursuit.
Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05), ** = significant (p≤0.01), *** = significant (p≤0.001). Mixed ANOVA results shown and results after Bonferroni correction shown in brackets.
Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

<u>Parametric analysis of smooth pursuit velocity gain</u>

Three-way mixed ANOVA

The between subjects factor was group and the within subjects factors were visit and smooth pursuit direction. There was a significant main effect of smooth pursuit direction. Horizontal velocity gain (control group M=0.753, surgery group M=0.766) was higher (better) than vertical velocity gain (control group M=0.437, surgery group M=0.377) ($F(1,7)=62.45$, $p<0.0001$, $r=0.95$). The main effects of visit and group were not significant ($p>0.05$). None of the interactions were significant ($p>0.05$). Planned comparisons between groups showed there was no significant difference between the groups at visit one or at visit two ($p>0.05$) (shown in Figure 9-30, red and blue text respectively). Post-hoc paired t-tests showed there was significantly better gain horizontally than vertically, in both groups at both visits ($p<0.05$). Although after Bonferroni correction (for three comparisons) this was no longer significant for the control group, visit two (Figure 9-30, black text, in brackets). Post hoc paired t-tests confirmed there was no significant difference between the visit one and visit two results in either group, for either direction. These within groups results are both shown in Figure 9-30 (black text).

Smooth pursuit - accuracy

The median (IQR) accuracy results for horizontal and vertical smooth pursuit, in the surgery group and the control group at visit one and visit two are shown in Figure 9-31. On inspection of Figure 9-31 vertical smooth pursuit appeared less accurate than horizontal smooth. Pursuit at visit two was less accurate than at visit one, in both groups. Smooth pursuit accuracy appeared similar in both groups.

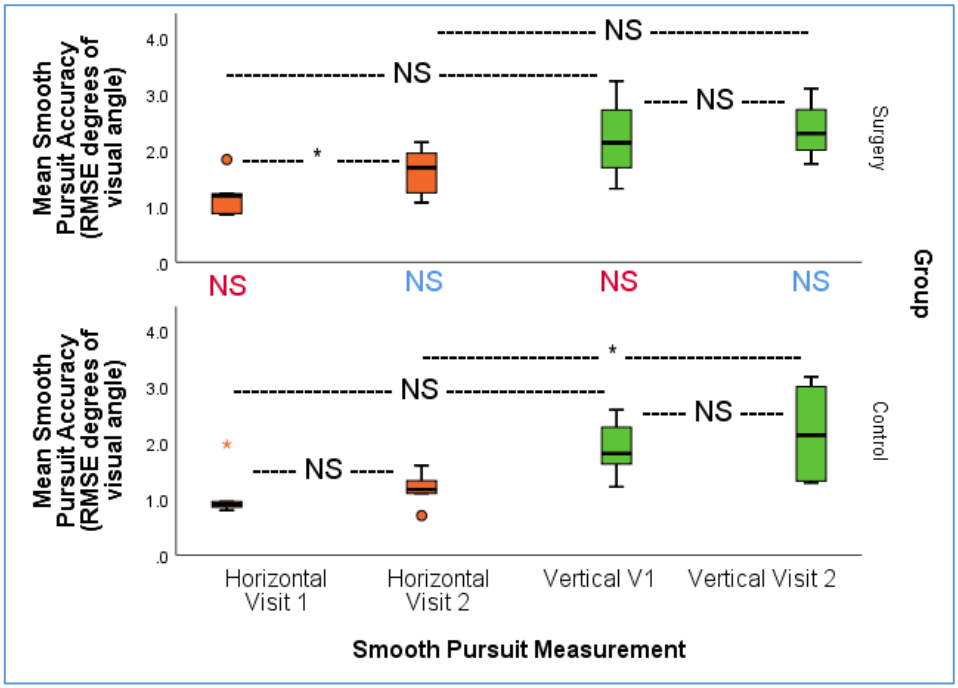


Figure 9-31. Boxplot of the EMR mean smooth pursuit accuracy (RMSE degrees of visual angle) for horizontal and vertical smooth pursuit, at visit one and visit two, for both groups.

Orange = horizontal smooth pursuit accuracy. Green = vertical smooth pursuit accuracy. Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05). Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two (Independent samples Mann-Whitney U test).

<u>Non-parametric analysis of smooth pursuit accuracy</u>

There was no significant difference in horizontal or vertical smooth pursuit accuracy between the groups at visit one or visit two (p>0.05) (between groups results shown in Figure 9-31, red and blue text respectively). From visit one to visit two horizontal smooth pursuit accuracy worsened significantly in the surgery group (from Mdn=1.19 to 1.69) (T=15, p=0.04, r=0.90) and worsened, but not significantly, in the control group (from Mdn=0.90 to 1.17) (T=10, p=0.50, r=0.30). There was no significant change in vertical smooth pursuit accuracy in either group (p>0.05). At visit one there was no significant difference in horizontal and vertical smooth pursuit accuracy (p>0.05). However, at visit two vertical smooth pursuit accuracy was significantly worse than horizontal in the control group (p=0.04). Within groups results are shown in Figure 9-31 (black text).

Smooth pursuit – fixation duration

The total duration of fixation (ms) results for horizontal and vertical smooth pursuit, in the surgery group and the control group at visit one and visit two are shown in Figure 9-32. Vertical smooth pursuit fixation duration was expected to be shorter than the fixation during the horizontal smooth pursuit as each vertical trial (8000ms) was shorter than the horizontal trial (10,000ms). On inspection of Figure 9-32 there was no apparent difference between the groups at visit one or visit two. There appeared to be little change in the duration of fixation from visit one to visit two within the surgery group or the control group.

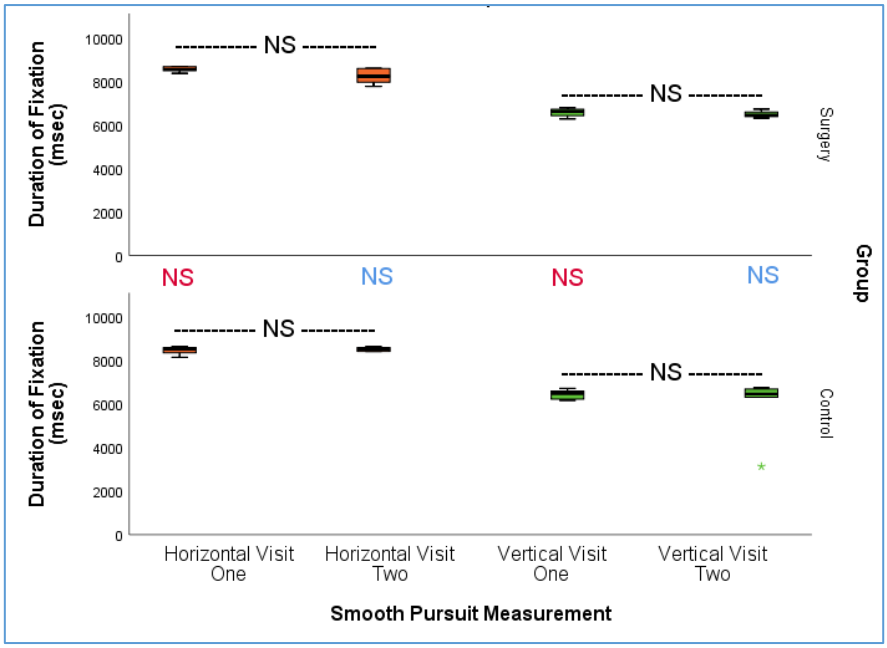


Figure 9-32. Boxplot of the EMR smooth pursuit duration of fixation (msec) for horizontal and vertical smooth pursuit, at visit one and visit two, for both groups.

Orange = horizontal smooth pursuit. Green = vertical smooth pursuit.
Statistical significance reported as NS = not significantly different (p>0.05).
Text in black = within groups analysis (Wilcoxon signed rank test). Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two (Independent samples Mann-

<u>Non-parametric analysis of smooth pursuit fixation duration</u>

There was no significant difference in duration of fixation (ms) between the groups at visit one or visit two, for either horizontal smooth pursuit or vertical smooth pursuit (p>0.05). These between groups results are shown in Figure 9-32 (red and blue text respectively). Within each group, there was no significant difference in the duration of fixation from visit one to visit two, for either horizontal smooth pursuit or vertical smooth pursuit (p>0.05). These results are shown in Figure 9-32 (black text). The difference between horizontal and vertical smooth pursuit duration was not analysed.

Smooth pursuit results – summary

Surgery had no effect on smooth pursuit velocity gain or fixation duration, but surgery worsened smooth pursuit accuracy. Accuracy worsened in both groups, but only the results in the surgery group were significantly worse (Figure 9-31). The hypothesis of an effect of surgery on smooth pursuit was rejected for velocity gain and duration of fixation but accepted for accuracy. There was no significant difference in smooth pursuit velocity gain (Figure 9-30), accuracy (Figure 9-31) and duration of fixation (Figure 9-32) during horizontal and vertical smooth pursuit between the surgery group and the control group at visit one or visit two.

9.6.3 Saccades

Horizontal and vertical mean saccade gain (0->1), peak velocity (degrees/second) and latency (ms) results are presented for eleven participants (control n=6, surgery n=5). Parametric analysis was used to analyse the saccade data.

Saccade gain

The mean (±2SD) saccade gain for the horizontal and vertical saccades, in the surgery group and the control group at visit one and visit two are shown in Figure 9-33. On inspection of Figure 9-33 vertical saccade gain appeared slightly lower than horizontal saccade gain. There appeared to be little difference between the groups at visit one or visit two. Within each group there appeared to be little change in the gain from visit one to visit two.

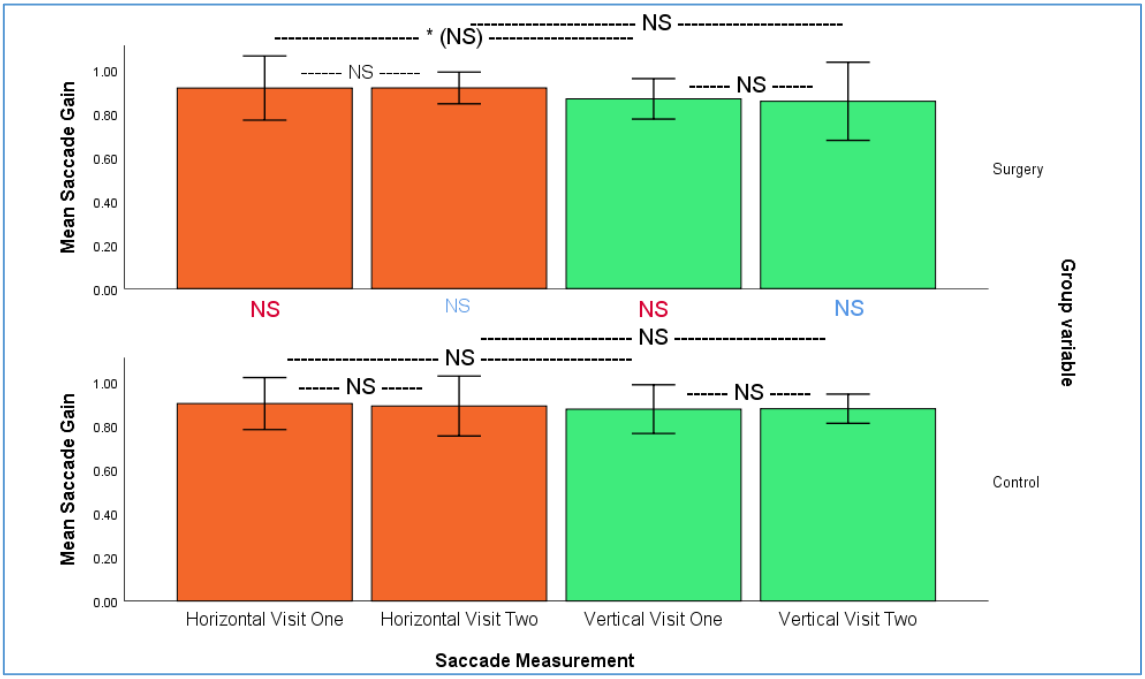


Figure 9-33. Bar chart of the EMR mean saccade gain for horizontal and vertical saccades, at visit one and visit two, for both groups

Orange = horizontal saccades. Green = vertical saccades.
Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05). Mixed ANOVA results are shown. Results after Bonferroni correction shown in brackets.
Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue =

Three-way mixed ANOVA

Levene's test showed there was significantly different variance in the control group and surgery group for vertical saccade gain at visit two (p=0.04). Thus, the assumption of equal variance in the two groups was violated. There was no significant main effect of visit, group or direction (p>0.05). Post-hoc tests of visit confirmed no significant effect for any of the groups or directions. Post-hoc analysis of direction found the difference between horizontal and vertical saccade gain in the surgery group was significant at visit one (p=0.04), but this was no longer significant after Bonferroni correction (for two comparisons) (p=0.08). These within groups analyses are shown in Figure 9-33 (black text). There were no significant two or three-way interactions (p>0.05). There was no significant difference between the groups at visit one or visit two (planned comparisons shown in Figure 9-33, in red and blue text respectively). Within the groups there was no significant difference in saccade gain, shown in Figure 9-33 (black text). Saccade gain was not significantly different in the surgery group after surgery. Horizontal saccade gain was not significantly different from vertical saccade gain.

Non-parametric analysis of saccade gain

Due to the violation of assumption of equal variances in the two groups (in the ANOVA above), non-parametric analysis of saccade gain was also undertaken and confirmed there was no effect of surgery on saccade gain. Thus, the hypothesis was rejected. At visit one and visit two there was no significant difference in horizontal and vertical saccade gain between the two groups (p>0.05) (Independent Samples Mann-Whitney U Test). Within the groups there was no significant difference in horizontal and vertical saccade gain from visit one to visit two (p>0.05) (Wilcoxon Signed Rank Test). There was no significant difference between horizontal and vertical saccade gain at either visit, except in the surgery group at visit two where the vertical saccade gain (Mdn=0.79) was significantly lower than horizontal saccade gain (Mdn=0.91) (T=0.0, p=0.04, r=0) (Wilcoxon Signed Rank Test).

Saccade peak velocity

The mean (±2SD) saccade peak velocity for horizontal and vertical saccades, in the surgery group and the control group, at visit one and visit two are shown in Figure 9-34. On inspection of Figure 9-34 the peak velocity of vertical saccades appears lower (slower) than for horizontal saccades, as expected due to the difference in saccade amplitude (horizontal=10 degrees, vertical=8 degrees). There appeared to be little difference in peak velocity between the groups at either visit and within each group, from visit one to visit two.

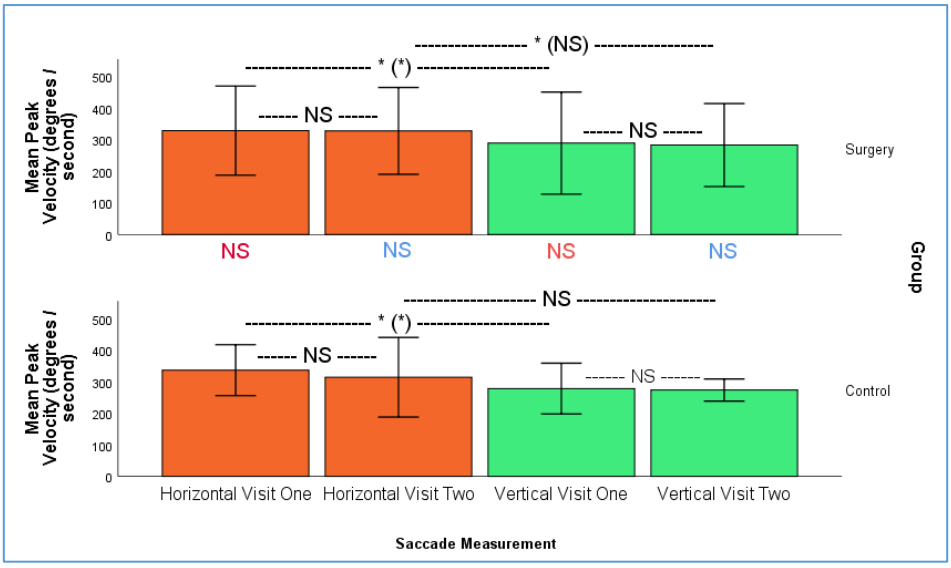


Figure 9-34. Bar chart of the EMR mean saccade peak velocity (degrees/second) for horizontal and vertical saccades, at visit one and visit two, for both groups.

Statistical significance reported as NS = not significantly different (p>0.05), * = significant (p≤0.05). Results after Bonferroni correction shown in brackets. Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Three-way mixed ANOVA

There was no significant main effect of visit or group (p>0.05). Planned between groups comparisons confirmed no significant difference in saccade peak velocity between the groups at visit one or two (p>0.05) (shown in Figure 9-34, in red and blue text respectively). The main effect of direction was significant (F(1,9)=18.12, p=0.002, r=0.82). Saccade peak velocity was significantly slower during vertical saccades (control M=275.87, surgery M=286.50) compared to horizontal saccades (control M=321.06, surgery M=328.79). Post-hoc analysis of the direction confirmed the vertical peak velocity was significantly slower than the horizontal peak velocity for both groups at visit one (but not at visit two), even after Bonferroni correction for two comparisons (p<0.05). The within groups analysis for visit and direction are shown in Figure 9-34 (black text). There were no significant two or three-way interactions (p>0.05). Saccade peak velocity was not significantly different in the surgery group after surgery.

Saccade latency

Saccade latency (ms) for horizontal and vertical saccades, in the surgery group and the control group, at visit one and visit two are shown in Figure 9-35. On inspection of Figure 9-35 the latency of vertical saccades appeared slightly longer than horizontal saccades. Saccade latency appeared to be longer in the surgery group compared to the control group, at both visits and for both directions. Saccade latency at visit two appeared longer then visit one, in both groups.

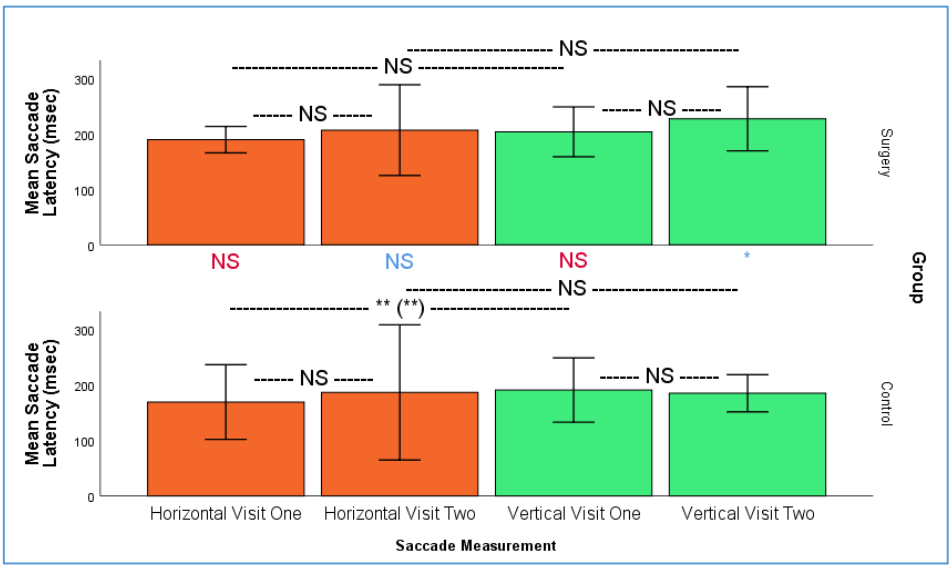


Figure 9-35. Bar chart of the EMR mean saccade latency (msec) for horizontal and vertical saccades, at visit one and visit two, for both groups.

Orange = horizontal saccades. Green = vertical saccades.
Statistical significance reported as NS = not significantly different (p>0.05), ** = significant (p≤0.01).
Results of the mixed ANOVA shown. Results after Bonferroni correction shown in brackets.
Text in black = within groups analysis. Text in red = between groups analysis, visit one. Text in blue = between groups analysis, visit two.

Three-way mixed ANOVA

There was no significant main effect of group or direction ($p>0.05$). The main effect of visit was significant, with shorter saccadic latency at visit one ($M=188.80$) than at visit two ($M=201.86$) ($F(1,9)=6.25$, $p=0.034$, $r=0.64$). None of the two or three-way interactions were significant ($p>0.05$). Planned comparisons between groups are shown in Figure 9-35 (red text, visit one and blue text, visit two). The difference between the groups was significant for the surgery group at visit two, as the surgery group had longer latency vertical saccades than the control group ($p=0.014$) (Figure 9-35, blue text). Within each group and direction there was no significant change in saccade latency from visit one to visit two ($p>0.05$) (post-hoc within groups analysis of visit). Whilst the main effect of direction was not significant ($p>0.05$), post-hoc within groups analysis of direction showed there was a significantly shorter latency for horizontal saccades than vertical, in the control group at visit one ($p<0.01$) which remained significant after Bonferroni correction for two comparisons ($p<0.01$) (Figure 9-35, black text). Saccade latency was not significantly different in the surgery group after surgery.

Summary of saccade results

Using saccade gain, peak velocity and latency as the measures of saccade performance, there was no difference between the surgery group and the control group at either visit one or visit two, except vertical saccade latency which was longer in the surgery group at visit two (compared to the control group) (Figure 9-35, blue text). There was no significant change in saccade performance in either group from visit one to visit two. Comparing horizontal and vertical saccades, there was no significant difference in saccade gain (Figure 9-33) or saccade latency (Figure 9-35), except for the control group at visit one who had significantly shorter horizontal saccadic latency. The peak velocity of vertical saccades was significantly slower than horizontal saccades (Figure 9-34), however this was expected due to the smaller amplitude vertical saccades during the task. Strabismus surgery had no effect on saccade gain, peak velocity or latency and the hypothesis of an effect of surgery on saccades was rejected. Saccade performance in the surgery group was not better, or worse, after strabismus surgery.

9.7 Further analysis by postoperative judgement of success, partial success and failure

The success, partial success and failure criteria presented in Table 8-1 (adapted from Hatt et al. (2012a)) were used to analyse the deviation results and were shown in Table 9-2. Additionally, the binocular summation, CST, grooved pegboard, AS-20, VFQ-25 and additional study questions results are presented using the same criteria and are shown in Table 9-14. These measures were selected as they were completed by all (or most) of the surgery group and were the measures most likely to measure change postoperatively. The mean change in each of the measures from

visit one to visit two are shown, additionally median change is shown for comparison to the results of Hatt et al. (2012a).

Table 9-14 Mean and median change in measures for the success, partial success and failure outcomes in the surgery group

	Success (n=7) Participants 1, 10, 13, 15, 16, 20, 21		Partial Success (n=4) Participants 3, 5, 9, 18		Failure (n=1) Participant 7
Change in measure (from visit one to visit two)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	
Binocular Summation 100%	1.43 (3.91)	2 (3)	6 (4.55)	5.5 (8.5)	9
Binocular Summation 10%	-1.14 (1.35)	-1 (2)	0.75 (4.27)	-1 (7.25)	-3
Binocular Summation 5%	1.86 (4.18)	3 (7)	-3 (2.94)	-3 (5.5)	-2
CST correctly seen BEO	1.1 (0.16)	1 (0.13)	1.13 (0.20)	1.06 (0.35)	1.43
CST response time BEO	0.73 (0.22)	0.7 (0.40)	0.67 (0.17)	0.61 (0.31)	0.54
CST RCS BEO	1.61 (0.51)	1.6 (0.84)	1.81 (0.67)	1.72 (1.23)	2.63
grooved pegboard time - bright #	1.01 (0.09)	1.01 (0.06)	0.93 (0.09)	0.89 (0)	1.05
grooved pegboard time - dim #	1.05 (0.11)	1.02 (0.18)	0.88 (0.19)	0.89 (0)	1.12
AS-20 overall score	44.29 (14.45)	46.25 (22.50)	29.69 (19.72)	28.75 (37.19)	8.75
AS-20 psychosocial subscale	54.64 (24.68)	55 (30)	33.75 (24.20)	31.25 (46.25)	2.5
AS-20 function subscale	35.36 (12.94)	35 (22.50)	25.63 (18.41)	25 (34.38)	15
VFQ-25 composite score	11.47 (9.00)	11.1 (14.58)	11.32 (8.46)	13.19 (15.57)	18.64
Additional Study Questions vision subscale	41.56 (16.63)	40.91 (22.73)	25.57 (16.22)	30.68 (28.98)	6.82
Additional Study Questions task performance subscale	19.05 (13.36)	20.83 (25)	6.25 (9.92)	6.25 (18.75)	20.83
Additional Study Questions physical symptoms subscale	26.19 (17.96)	29.17 (37.50)	1.04 (10.96)	-2.08 (19.79)	29.17

Additional Study Questions confidence and emotions subscale	46.07 (19.20)	45 (25)	31.25 (20.56)	32.5 (38.75)	32.5
# the grooved pegboard was not completed by participant 18 postoperatively and therefore the partial success results for the grooved pegboard are n=3					

9.8 Key quantitative findings

The hypothesis of an effect of strabismus surgery on each of the measures tested in phase two was stated without direction, as change was not assumed to be an improvement. The possibility of improved and worsened performance of each of the measures was thoroughly investigated. The key findings from the quantitative phase are summarised below.

9.8.1 Improvement after strabismus surgery

The size of the horizontal deviation significantly reduced after strabismus surgery. The baseline size of deviation was a significant factor in the change of deviation, as those with larger deviations at visit one who had surgery, had larger changes in their deviation.

PROMs

Postoperatively HRQoL (AS-20 overall score and both subscales) increased to the same level (overall and psychosocial) or significantly better (function subscale) than the control group. Self-reported visual function (VFQ-25) improved significantly after strabismus surgery to better than, but not significantly different to, the control group. It is possible that the high VFQ-25 scores had a ceiling effect that limited the ability of the VFQ-25 to measure change in visual function after strabismus surgery.

The additional study questions overall score improved significantly after surgery to higher than, but not significantly different to, the control group. Baseline size of deviation was not a significant factor. All patients reported improved scores postoperatively, not just those having surgery for a large deviation. All the additional study questions subscale scores improved after strabismus surgery to better than the control group (vision) or not significantly different to the control group (task performance, physical symptoms and confidence and emotions).

Vision

Binocular summation score (letter scoring method) increased significantly after strabismus surgery for 100% contrast optotypes (but not 10% or 5% contrast). This represented better performance both eyes open compared to monocular performance with the better seeing eye. The change in binocular summation from visit one to visit two showed the surgery group had a positive change in binocular summation (increased binocular summation or reduced binocular inhibition) and the

control group had a negative change (reduced binocular summation or increased binocular inhibition) at 100% contrast.

The CST (gross BSV test) had a significant practice effect. Postoperatively, the surgery group had improved CST performance with both eyes open that was greater than that experienced by the control group and greater than that measured monocularly.

Task performance

Postoperatively there was a small but significant improvement in TSL task time (time to touch the spots) and in the time to perform the CKAT aiming task. These improvements may represent a small effect of strabismus surgery or a practice effect.

9.8.2 Worsening after strabismus surgery

Postoperatively there was a significant increase in left LR restriction or underaction, which may have been a desired result of surgery, as the surgery group all had XT preoperatively. Grooved pegboard performance worsened postoperatively, by (on average) 3.12 seconds in bright light and 6.25 seconds in dim light. Whilst the worsening in the surgery group was not statistically significant, the control group improved at visit two by 9.51 seconds (bright light) and 8.49 seconds (dim light). Postoperatively the surgery group performed significantly worse than the control group in bright lighting. Bead threading time (large or small beads) did not change significantly after surgery, however compared to the control group this lack of improvement may have been a slight worsening in performance. Horizontal smooth pursuit accuracy significantly worsened after horizontal strabismus surgery by median 0.50 degrees (visual angle), compared to the control group where accuracy worsened by median 0.27 degrees, but this was not statistically significant.

9.8.3 No change after strabismus surgery

Vision

Postoperatively there was no significant change in distance VA (in either eye), near VA, contrast sensitivity threshold or fixation grade. Participants found the sequential stereopsis task difficult and there was no significant difference in sequential stereopsis after strabismus surgery. Both groups had worsened performance of the task at visit two, although this was no longer statistically significant after inclusion of the covariate in the ANCOVA.

Clinically the visual field area was expected to reduce after surgery for XT, as visual field area has been shown to increase following reduction of ET. However, visual field area did not change significantly after surgery, despite all the surgery group having XT preoperatively.

Task performance

There was no significant postoperative change in Purdue pegboard performance (peg insertion or assembly building), TSL task accuracy, CKAT tracking accuracy or CKAT steering performance.

EMR

There was no significant change in the performance of the fixation, smooth pursuit and saccade tasks after strabismus surgery, except horizontal smooth pursuit accuracy which worsened after surgery (described in section 9.7.2).

9.9 Summary

Chapter 9 has described the participants recruited to the surgery group and control group during the quantitative phase. Between groups analysis was reported at visit one and visit two. Within groups analysis was reported from visit one to visit two. Statistical analysis of the core orthoptic measures, objective measures of vision, task performance and EMR, and the PROMs have been presented. However, it is emphasised that the quantitative phase results of this feasibility study are exploratory. The numbers recruited to each group were small and the study was not powered for statistical significance. The statistical significance of the quantitative results have been, and should continue to be, interpreted with caution, especially considering the number of comparisons made.

The PROMs included questions exploring physical symptoms and confidence and emotions, as well as vision and task performance. The hypothesis of an effect of surgery was accepted for measures that significantly improved or worsened after surgery. Significant improvements were measured in the PROMs, binocular summation (100% contrast), CST performance, TSL task performance (time) and CKAT aiming performance (time). Significantly worsened OM (LR underaction or restriction), grooved pegboard performance, bead threading and smooth pursuit accuracy were also measured postoperatively. The hypothesis of an effect of surgery was rejected for VA, contrast sensitivity, fixation grade, sequential stereopsis, visual field area and EMR (fixation, smooth pursuit velocity gain, fixation duration during smooth pursuit, saccade gain, saccade peak velocity and saccade latency). In chapter 10 the qualitative findings (chapter 6) and the quantitative results (chapter 9) will be brought together and discussed as the overall study.

Chapter 10. Discussion

Chapter 10 discusses the results from both the qualitative phase (chapter 6) and quantitative phase (chapter 9) as the findings of the overall study. The semi-structured interviews and thematic analysis (qualitative phase one) explored patients' perceptions of their outcomes from strabismus surgery. The quantitative phase two prospectively measured vision, task performance, physical symptoms and confidence and emotions, before and after strabismus surgery (undertaken for psychosocial reasons) in a surgery group and a control group of adults with strabismus not undergoing surgery. Findings from these two datasets are discussed below under the themes of vision, task performance, physical symptoms and confidence and emotions, in the context of the wider evidence base. The findings are presented by what improved, what remained unchanged and what worsened following strabismus surgery, in addition to the expected postoperative changes.

10.1 Key findings

In adults undergoing strabismus surgery for psychosocial reasons, surgery can improve some aspects of vision and task performance, in addition to the known psychosocial benefits of improved eye alignment. Some aspects of vision and task performance do not change postoperatively, and some worsening of vision and task performance can also occur. Patients value strabismus surgery for psychosocial reasons highly and report significantly improved HRQoL, vision, task performance, physical symptoms and confidence and emotions postoperatively.

10.1.1 Key results summary

- After strabismus surgery improvements were measured in binocular summation (100% contrast), CST performance with both eyes open, TSL performance time and CKAT aiming time.
- There was no measurable change in visual field size postoperatively, despite an expected reduction with improved alignment in XT.
- Using PROMs, participants reported significantly improved visual function, task performance, physical symptoms, and confidence and emotions postoperatively, in addition to the expected improvement in HRQoL (AS-20).
- Postoperatively LR limitation or underaction, grooved pegboard performance, bead threading and smooth pursuit accuracy worsened.
- Postoperatively no change was measured in VA, contrast sensitivity, ability to take up fixation, sequential stereopsis, Purdue pegboard performance, CKAT tracking, CKAT steering, TSL accuracy, binocular summation (10% and 5% contrast) or EMR (fixation, smooth pursuit and saccades, except smooth pursuit accuracy which worsened).

10.1.2 Key feasibility and acceptability findings

- It was feasible to measure and quantify changes in visual function and task performance following strabismus surgery.
- It was feasible to recruit and retain adult strabismus patients to a study that required measurements before and after surgery, however a smaller number of measurements that could have been performed in the eye clinic may have improved recruitment.
- Most of the tests performed in phase two were acceptable to patients, but sequential stereopsis was commonly reported as too difficult.
- The acceptable tests could be refined for a future study to include those most likely to measure change, including PROMs, binocular summation, task performance time (TSL or CKAT), gross BSV (CST) and visual field size (with modifications).

The findings from the qualitative phase one (chapter 6) and quantitative phase two (chapter 9) will now be discussed in the context of the wider evidence.

10.2 What improved following strabismus surgery?

10.2.1 Vision

Binocular summation

In phase two improved binocular summation was measured postoperatively at 100% contrast, but not at 10% or 5% contrast. Binocular summation has been shown to be reduced in strabismus (Chang et al., 2017; Pineles et al., 2013) compared to normative values (Pineles et al., 2014). Binocular summation or improved visual performance of a task both eyes open compared to performance with the better seeing eye monocularly, is described as evidence that the strabismic eye contributes positively to both eyes open viewing, even though suppression may be detected clinically. The method of using high and low contrast visual acuity to measure binocular summation has been reported as comparable to other methods (Pineles et al., 2013).

Barrett et al. (2013) measured binocular summation across the central 50 degrees of the visual field in an adult cohort (n=10) with constant manifest strabismus and amblyopia. Six subjects had no BSV. Binocular summation existed in most (but not all) of their subjects. Dorr et al. (2019) measured binocular summation using the contrast sensitivity function in subjects with strabismus, with and without fusion. Kattan et al. (2016) used high (100%) and low contrast (2.5% and 1.25%) acuity charts to measure binocular summation, presenting the low contrast charts in a dimly lit room, although luminance was not reported. The same method was used in other studies from the same group (Chang et al., 2017; Pineles et al., 2015; Pineles et al., 2014; Pineles et al., 2013). Kattan et al. (2016) reported only postoperative measurements of binocular summation and analysed the relationship between binocular summation and stereoacuity in patients with and

without BSV and diplopia. Postoperatively stereoacuity was significantly correlated with greater binocular summation, however, correlations were greatest at 2.5% contrast. At 100% contrast the correlation was not significant. In comparison, this study found the greatest changes occurred at 100% contrast. This difference may be due to different testing methods, as Kattan et al. (2016) used physical visual acuity charts that could not be randomised. Whereas this study presented 100%, 10% and 5% contrast optotypes, randomised on the Thomson chart. 2.5% and 1.25% optotypes were not visible on the Thomson chart, therefore 10% and 5% presentations were selected. Additionally, Kattan et al. (2016) presented the low acuity charts in a dimly lit room, yet in this study room luminance was bright and constant. The effect of room luminance on binocular summation is not currently known. Kattan et al. (2016) concluded stereopsis and binocular summation may have common neural pathways but that binocular summation was more easily demonstrated in low contrast acuity. Whilst they did not specifically report binocular summation measurements in those without diplopia and BSV postoperatively, they hypothesised that low contrast acuity and binocular summation may represent visual improvements postoperatively, beyond improvements in diplopia even if stereoacuity was not achieved. The phase two results support the views of Kattan et al. (2016), as participants showed improved binocular summation at 100% contrast postoperatively, but did not gain improvements in diplopia or stereopsis, except participant 15 who unexpectedly gained stereopsis.

Pineles et al. (2015) reported improved binocular summation, at all contrast levels, 2 months postoperatively in children and adults with a range of strabismus subtypes, including with and without BSV and diplopia. Patients with a successful surgical outcome (within 10PD horizontal and 4PD vertical deviation) had the greatest improvement in binocular summation. Whilst improvements occurred, binocular summation did not reach normative levels (Pineles et al., 2014). Pineles et al. (2015) concluded that postoperative binocular summation improvements were further evidence of strabismus surgery leading to functional improvements in vision. Chang et al. (2017) measured binocular summation before and 6-10 weeks after strabismus surgery in patients with strabismus and amblyopia (n=11) and in control groups with strabismus and no amblyopia (not undergoing surgery) and with normal vision. A small number of the strabismic patients (in both groups) had diplopia. Reduced binocular summation at 2.5% and 1.25% contrast was measured in strabismus compared to normal vision, with no additional effect of amblyopia. Only those with stereopsis preoperatively had improved low contrast binocular summation postoperatively. Having a larger strabismus preoperatively and a later onset strabismus were also significantly associated with improved binocular summation at 1.25% contrast. Whilst the findings of Chang et al. (2017) appeared to be in contrast with the findings of this study, of note were the similarities to some of the patients reported. Those with no stereopsis preoperatively were most similar to the phase two surgery group. Interestingly those with no stereopsis preoperatively showed a similar pattern of results to those measured in this study, with a positive change in binocular summation at 100% contrast and a negative change in binocular summation at lower contrast levels (Figure 3 in Chang

et al. (2017) and Figure 9-15). Chang et al. (2017) also measured improved binocular summation on repeated testing in their control group, highlighting the importance of a control group.

The postoperative change in binocular summation measured in phase two may be due to some subclinical improved binocular vision or binocular cooperation between the eyes, or the effect of the change in angle of deviation, or both. Alberti and Bex (2018) investigated binocular summation in subjects with BSV, presenting stimuli on different retinal locations. Binocular summation was influenced by disparity. Binocular inhibition occurred when stimuli were on non-corresponding retinal locations and binocular summation occurred when stimuli were on corresponding retinal locations. It is therefore possible that the improved binocular summation (100% contrast) measured in this study, and at lower contrast levels measured by others (Chang et al., 2017; Pineles et al., 2015) are the result of the stimulus perceived by the deviating eye changing from a more eccentric retinal location to a more central (but not necessarily corresponding) retinal location, during both eyes open viewing. It is possible that improved binocular summation may represent the improvements in vision both eyes open described by some participants in phase one. These are discussed in detail below.

Clearer vision and better focussing

In phase one, clearer vision and better focussing with both eyes open were commonly reported, including being able to focus or see more easily, faster or more naturally. Vision was described as being sharper and clearer, more central and straighter. Better vision was also described as having more natural and more comfortable vision. The consequences of improved vision included feeling less confused, less distracted by their strabismus, or less like their strabismus was getting in the way. Having less confusion was perceived, by one participant, to be occurring because they felt they were able to look through both eyes postoperatively.

Perception of looking through both eyes at the same time

In phase one some participants described looking through both eyes at the same time, despite suppression clinically. It is not known whether this related to binocular summation or CST performance. Kushner (1994) reported ET's with a postoperative expanded visual field with both eyes open described simultaneously seeing with both eyes, however this was attributed to postoperative binocularity, as 29 of 35 participants had BSV (Bagolini lenses). Economides et al. (2012) measured suppression across the visual field in participants with XT. Suppression of peripheral temporal retina of both eyes occurred in XT and smaller deviations required less retina to be suppressed. Suppression area and dichoptic visual field perception (Economides et al., 2012) were not measured in phase two. However, it is possible that in the absence of BSV but with a smaller deviation postoperatively, less retinal area was suppressed compared to preoperatively and this was perceived as improved vision. If less retinal area was suppressed, a greater retinal area of the deviating eye may have contributed to both eyes open viewing.

The feeling of looking through both eyes at the same time, despite suppression with Bagolini glasses, may also relate to use of the apparently suppressed eye. In ten strabismic amblyopes Barrett et al. (2012) found little evidence for suppression during a visual field sensitivity task (blue stimuli presented on a yellow background using the Humphrey Visual Field Analyser). In a separate paper reporting the same cohort, Barrett et al. (2013) reported the 'apparently suppressed' strabismic and amblyopic eye was used by most (7 of 10) of their cohort to improve performance of the task when both eyes were open. A useful contribution of the apparently suppressed eye in strabismus, during both eyes open viewing, has also been demonstrated during a saccade task (Griffiths et al., 2011), a mobility and obstacle task (Buckley et al., 2010b), prehension (Grant et al., 2007) and motor skills (O'Connor et al., 2010a). It is possible that strabismus surgery may change the amount or the extent the deviating eye can contribute to both eyes open performance. Or the deviating eye contribution may remain unchanged, but it is more perceptible to the patient postoperatively. In phase two, the apparently suppressed eye contributed to both eyes open performance by expanding the visual field area, improving the number of letters read at 100% contrast (binocular summation) and improving performance of the CST. There was no measurable contribution of the suppressed eye to the sequential stereopsis task. A similar visual field sensitivity task to Barrett et al. (2012, 2013) was not performed. The PROM questions relating to looking through both eyes at the same time and being confused by vision will be discussed later (sections 10.2.3 and 10.4.3)

Perception of using the eyes together as a pair

In phase one, two participants reported a strong perception that they were using their eyes together as a pair and viewing binocularly, despite suppression (Bagolini glasses). In phase two, three participants achieved unexpected BSV. Participant 03 demonstrated abnormal peripheral fusion at near (Bagolini glasses cross with a central gap in the line and an intermittently positive Wirt fly (3552")). Participant 10 reported abnormal sensory fusion at near (Bagolini glasses cross only). Participant 15 reported normal BSV with sensory fusion, motor fusion (normal PFR with Bagolini glasses control) and stereopsis (110" FNS). In phase two additional measurements expected to capture the ability to use the eyes together included binocular summation, sequential stereopsis, the CST and the visual field. As discussed, postoperatively binocular summation (100% contrast) improved and there was some evidence of improved CST performance both eyes open. No improvement was measured in sequential stereopsis or visual field area. PROMs questions relating to using the eyes together as a pair and looking through both eyes will be discussed later (section 10.4.3).

Less eye closure

In phase one, needing to close one eye less was reported as a positive outcome from surgery, particularly when concentrating. This may have been associated with improved binocular

summation or less binocular inhibition (Chang et al., 2017). In phase two, improved binocular summation at 100% contrast was measured, as discussed earlier in this section. Using the categorisation method (Figure 9-16), less binocular inhibition (binocular summation score of ≤ -5) and more binocular summation (binocular summation score of ≥ +5) was measured postoperatively, although the numbers were small.

Perception of using the strabismic eye more postoperatively

In phase one, some participants described being able to use their strabismic eye more postoperatively. This was described as an improved ability to take up fixation with the strabismic eye or having greater peripheral vision and needing to make less head movements to see to the side. Additionally, two participants felt they gained control over their strabismus postoperatively. In phase two, greater peripheral vision was not measured, and surgery had no effect on fixation grade. Only two of the surgery group had a change in fixation grade; participant 15 (3 to 7) had an X and BSV postoperatively and participant 21 (2 to 3). PROMs questions relating to swapping fixation and control of eye position will be discussed later (section 10.2.3).

Coarse stereotest (CST)

The CST results showed a significant practice effect. All participants had improved CST performance at visit two, particularly in response time and RCS. Accuracy was consistently high for the CST. The accuracy ceiling effect and small number of trials may therefore have limited the analysis. The practice effect made interpretation of a possible effect of surgery challenging, however, the surgery group showed improved performance of the task with both eyes open to the extent that it was better than monocular performance, and this difference was larger than in the control group. However, there was no significant difference in both eyes open performance in the surgery group compared to the control group at visit two. Having improved performance of the CST with both eyes open postoperatively supported the phase one findings of improved depth perception and ability to judge object position and alignment better postoperatively. Whilst patients had no measurable BSV, some reported they were better able to see depth or judge object position postoperatively. Examples of improvements included the ability to pick up objects, thread needles or beads, judge steps and obstacle height. The CST and sequential stereopsis both measured judgements about the position and alignment of objects in phase two. While the CST showed some improvement postoperatively, there was no measurable change in sequential stereopsis (section 10.4.1).

The CST measured stereopsis at 3094” of arc at 1m. Despite all attempts to limit monocular clues during testing, it is possible some remained. All surgery group participants (n=12) had suppression at visit one and nine had suppression at visit two. Stereopsis was only tested if sensory fusion was reported (Bagolini glasses). Despite evidence of suppression, specific testing with the Wirt stereotest in all participants would have allowed greater comparison of the Wirt and CST results. It

is possible that the real depth stimuli of the CST may have been easier to detect than the Wirt fly (3552” at 40cm).

The CST was designed to be a test of gross stereopsis or gross BSV. Kitaoji and Toyama (1987) showed position stereopsis and motion stereopsis could be preserved independently and peripherally in the presence of manifest strabismus, including a small deviation postoperatively. The CST was a position stereopsis task, as no motion of the stimuli was involved, yet the stimuli used by Kitaoji and Toyama (1987) were small spots of light, rather than the black rods used in this study. It is possible that the CST was a gross enough measure of stereopsis, that it detected a similar position stereopsis peripherally compared to that measured by Kitaoji and Toyama (1987). The aspect of the CST that had the most improvement was the response time. This may have been a practice effect or related to participant descriptions of improved judgements of alignment or object position in phase one. The speed, or time taken to detect depth, has been described as an important consideration in stereo performance, but one rarely measured or captured by current testing methods (Saladin, 2005).

10.2.2 Task performance

Screen based tasks

In phase one a number of participants described tasks on screens or computers as easier or better postoperatively. In phase two the TSL and CKAT were screen-based measures of task performance. Surgery led to a small but significant improvement in TSL reaction time (time to touch the spots) in the surgery group (121ms) (Figure 9-25). Using the CKAT, surgery led to a very small but significant improvement in time to perform the aiming task (0.07seconds) whilst the control group had a very small but significant worsening (0.10seconds) in performance time. This interaction group x visit was significant, but when explored further in post-hoc comparisons with Bonferroni correction, was no longer significant (Figure 9-27). This may represent a very small effect of surgery, or a practice effect. It is possible that these small postoperative improvements in task performance time may be perceptible and associated with the phase one findings of improved task performance at near and when using screens. PROMs questions about task performance at near and using screens will be discussed later (section 10.2.4).

A touch panel test, similar to the TSL, was used by Schiller et al. (2012) to compare binocular and monocular viewing in participants with normal, reduced and nil stereoacuity. In those with no stereopsis touch panel accuracy and completion time were poorer monocularly compared to with both eyes open, however no statistical analysis was presented (Schiller et al., 2012). The CKAT was used in a pilot trial before and after neurosurgery (Raw et al., 2017). Compared to standard clinical measures of ability (function and cognition) that showed little change 6-weeks postoperatively, CKAT performance (all tasks) improved. Performance 6-weeks postoperatively

was similar to a control group; however, the performance of the control group was not reassessed, limiting analysis of a possible practice effect (Raw et al., 2017). Other tablet based measures of task performance (time and accuracy) have been used to investigate eye-hand coordination (Lee et al., 2014). Practice effects were evident in these tasks, highlighting the need to consider test-retest variability and a control group in repeated measures studies. Scott et al. (2002) found that reduced contrast sensitivity was the most significant factor associated with reduced computer task performance in AMD. In phase two all participants had good contrast sensitivity with both eyes open, with no significant change postoperatively.

10.2.3 PROMs - vision

In phase one, most participants (12 of 13) reported their vision improved postoperatively (section 6.3.2). Vision was described as clearer, focussing was better, peripheral vision and reading were improved and they were able to use the strabismic eye more. In addition to the phase two objective measures of visual function, the AS-20, VFQ-25 and additional study questions were used as PROMs for participants to subjectively report different aspects of their HRQoL, visual function, task performance, physical symptoms and confidence and emotions. The AS-20 results will be discussed later (section 10.5.2). The VFQ-25 and additional study questions results will be discussed below.

VFQ-25

The VFQ-25 composite score was not significantly different between the groups at visit one. Whilst the VFQ-25 score improved significantly after surgery, there remained no significant difference between the groups at visit two (Figure 9-6). The VFQ-25 results in this study were higher (better visual functioning) than adults with strabismus reported by Leske et al. (2010). Leske et al. (2010) also reported a slight, but significant, improvement in VFQ-25 score on retest. In comparison, in phase two there was a marginal reduction in VFQ-25 score in the control group, which was not significant. Compared to the surgery group in this study (visit one=83.49, visit two=94.81) Hatt et al. (2010a) reported a higher preoperative VFQ-25 score in patients without diplopia (n=26) (88.3), but less improvement in VFQ-25 score postoperatively (Mdn=5), even after removal of postoperative partial successes and failures. Possible reasons for a greater improvement in VFQ-25 score in this study are unclear.

In phase two median VFQ-25 scores for each group exceeded thresholds for visually normal adults (64, 66 or 71) (Hatt et al., 2009a). Hatt et al. (2009a) reported similar results in strabismic adults without diplopia (n=19) as VFQ-25 scores were rarely subnormal (0% below 64, 11% below 71). In phase two only three participants scored lower than these visually normal thresholds at visit one (control 69.12, surgery 52.95 and 51.40). All had scores exceeding the higher normal threshold (71) at visit two. Leske et al. (2010) reported a VFQ-25 score greater than 11.1 exceeded the 95%

limits of agreement. In phase two a change in VFQ-25 score exceeding 95% limits of agreement was reported by none of the control group (0/15) and seven of the surgery group (7/12) (Table 9-5), which was higher than the 21% described by Hatt et al. (2010a).

Additional study questions – vision subscale

Additional study questions (n=33) were included in phase two as some descriptions of postoperative change in phase one were not included in the AS-20 or VFQ-25. Surgery led to a significant improvement in the vision subscale score, to significantly higher than the control group at visit two (Figure 9-8). The analysis of responses to individual questions about vision (1-11) is shown in Table 9-10. Postoperatively the surgery group reported their eyes had been turning significantly less (question 1). This agreed with the PCT results, which showed a significant reduction in horizontal angle of deviation after surgery (Figure 9-2). The surgery group also perceived they had significantly improved control of their eye position postoperatively (question 2). However, there was no significant change in fixation score and only one participant had a heterophoria postoperatively (Figure 9-12). It therefore seemed likely that participants associated improved eye alignment with improved control of the deviation. Ji et al. (2020) reported that 24% of their successfully aligned cohort perceived they still had ‘some’ or ‘obvious’ deviation postoperatively. They measured perception of deviation preoperatively, so comparison with the phase two results is difficult. However, they found no significant difference in the size of deviation between those who perceived they had no deviation postoperatively and those that perceived they had some or obvious deviation postoperatively.

Participants reported more central vision postoperatively in phase one and this was also measured postoperatively in the surgery group in phase two (question 6). However, there was no significant change in visual field size in the surgery group (increase or decrease), despite improved eye alignment and the perception of more central vision postoperatively. The visual field results will be discussed further in section 10.4.1.

Postoperatively the surgery group reported a significantly greater perception of being able to look through both eyes at the same time (question 10) and a significantly improved ability to use their eyes together (question 11). These findings agreed with the phase one results and the other phase two measures of improved binocular summation at 100% and improved CST performance (section 10.2.1).

10.2.4 PROMs - task performance

Improvements in tasks and daily activities following strabismus surgery for psychosocial benefit have been reported, but not specified or explored, in interviews and using PROMs (Beauchamp et al., 2005a; Hatt et al., 2012a; Nelson et al., 2008). Following descriptions of improved task

performance in phase one, the AS-20, VFQ-25 and additional study questions, were used as PROMs in phase two. The AS-20 results will be discussed later (section 10.5.2). The VFQ-25 composite score results have already been discussed (section 10.2.3). Many of the VFQ-25 questions were related to task performance, which highlighted the difficulty of separating vision and task performance. This was also found in phase one, as participants often talked about vision and tasks together, for example reporting being able to focus and perform tasks better.

Additional study questions - task performance subscale

There was a significant improvement in the task performance subscale score in the surgery group postoperatively (Figure 9-8). At visit two, the surgery group had a higher, but not significantly different, score than the control group (Figure 9-8). The analysis of responses to individual task performance questions (12-17) is shown in Table 9-11. Postoperatively, the surgery group reported significant improvements in being able to perform tasks quickly (question 15) and difficulty looking at screens (question 16). There was no significant change in the surgery group's perception of eye hand coordination difficulty (question 13). However, a ceiling effect was evident in the surgery group at both visits, suggesting that preoperatively difficulties with eye hand coordination were not perceived. Despite this, it is possible that a small improvement in eye hand coordination was reported by the surgery group, as they reported significantly fewer difficulties with eye hand coordination compared to the control group at visit two. These reported improvements in being able to perform tasks quickly, looking at screens and eye hand coordination were also reported in phase one. Interestingly the phase two tasks where time was recorded (TSL, CKAT aiming, CST) were found to be slightly improved postoperatively. The TSL and CKAT tasks were also screen based. These slight improvements may be perceptible by patients, and hence why they were reported as improved in phase one and in phase two using the PROM questions. However, not all timed measures improved (Purdue pegboard) and some worsened compared to the control group (grooved pegboard and bead threading).

10.2.5 PROMs - physical symptoms

In phase one, seven participants described postoperatively experiencing less eye pain, strain, tiredness and discomfort, improved headaches, the strabismic eye pulling less or feeling less tight. In some cases, surgery led to a resolution of physical symptoms. Two participants described worsened physical symptoms and reported increased eye sensitivity and watering, and worsened headaches (section 6.5.2).

AS-20 – physical symptoms

In phase two surgery resulted in a significant improvement in the both the eye strain (question 15) and the need to take frequent breaks when reading (question 20) questions of the AS-20 (Table 9-

4). At visit two the surgery group reported significantly less eye strain than the control group. These results agreed with the phase one findings.

VFQ-25 – physical symptoms

Postoperatively, the surgery group reported significantly less pain or discomfort (question 4) than the control group at visit two and a significant improvement in pain or discomfort preventing them doing things they would like to do (question 19) using the VFQ-25. These results were broadly in agreement with the phase one findings.

Additional study questions - physical symptoms subscale

In phase two, the surgery group reported significantly improved physical symptoms postoperatively (Figure 9-8). The analysis of individual questions relating to physical symptoms is shown in Table 9-12. The surgery group reported postoperative improvements in headaches (question 18), the eyes pulling (question 20), feeling tired (question 22) and feeling the eyes turning (question 23). These results were in agreement with the phase one findings.

Physical symptoms due to strabismus are less commonly reported than visual or psychosocial symptoms. Hatt et al. (2009b) described patients without diplopia reporting function concerns during the development of the AS-20, for example reporting 'sometimes' or 'often' to question 20 ('I need to take frequent breaks when reading because of my eyes'). This was interpreted as evidence supporting the inclusion of function questions in a strabismus QoL measure, even for those without diplopia. Wang et al. (2018) found that strabismus without diplopia was associated with physical discomfort and eye fatigue, in addition to blurred vision. Physical symptoms have also been reported to improve following strabismus surgery. Liebermann et al. (2014) reported postoperative outcomes in adults without diplopia (n=20) and all AS-20 functional elements improved postoperatively, with the greatest improvements in stress, worry, needing to take breaks, enjoying hobbies, depth perception and eye strain items. Question 15 ('My eyes feel strained') improved from mean 46.25 to 65.79 (Mdn=50 to 50). Question 20 ('I need to take frequent breaks when reading because of my eyes') improved from mean 61.25 to 86.25 (Mdn=50 to 100) (Liebermann et al., 2014). In this study, the phase two results showed a slightly larger postoperative improvement in eye strain (question 15) and a similar improvement in taking breaks when reading (question 20) compared to Liebermann et al. (2014).

10.2.6 PROMs - confidence and emotions

The AS-20 psychosocial subscale contains questions relating to confidence and emotions. The AS-20 results will be discussed later (section 10.5.2).

Additional study questions - confidence and emotions subscale

Surgery resulted in a significantly improved confidence and emotions subscale score (Figure 9-8). Compared to the other subscales, the improvement in the confidence and emotions median score was the largest (38.75), followed by the vision (34.1), task performance (18.75) and physical symptoms (14.58). These results showed that participants reported significant improvements in a range of psychosocial aspects relating to confidence and emotions, in addition to the psychosocial improvements measured by the AS-20 (psychosocial subscale and overall score) (Figure 9-5).

Analysis of the individual questions relating to confidence and emotions is reported in Table 9-13. Preoperatively, the surgery group reported significantly worse confidence and emotions than the control group for all questions, except feeling happy (question 33). Postoperatively, the surgery group reported significant improvements in all the questions, except having confidence in abilities (question 28). This was in contrast to the phase one finding of postoperative improvements in confidence in abilities such as work or tasks. At visit two there was no significant difference between the groups, except the surgery group reported liking the way their eyes looked (question 30) significantly more than the control group with strabismus not seeking surgery.

Self-confidence and interactions with others

Strabismus can negatively affect QoL, causing low self-esteem and self-confidence (Nelson et al., 2008; Xu et al., 2012) and a higher incidence of anxiety and depression (McBain et al., 2014b). Strabismus is known to have a negative effect on face-to-face communication. Problems making eye contact and trying to hide the strabismus (Menon et al., 2002), avoiding social situations (Xu et al., 2012), avoiding activities and embarrassment about strabismus (Ghiasi et al., 2013) have all been reported. Strabismus has also been associated with problems with interpersonal relationships and surgery for psychosocial reasons has been shown to improve psychosocial functioning (Burke et al., 1997). In phase one most participants (n=12) described themselves as happier, feeling better in themselves, having improved confidence and no longer suffering with stress, anxiety or worry postoperatively. Feeling better in themselves and having more self-confidence was linked to overall improvements in participants' lives, feeling more able to do things and socialise more. All participants (n=13) reported being able to interact with and communicate with people much more. Participants described being able to go out in public, look at people, make eye contact and have face-to-face communication. In a social setting participants reported being able to meet new people and feeling more relaxed and confident when meeting new people. Having straighter eyes was associated with being able speak to people and socialise more, and not avoiding social situations or communication with people. Participants described feeling less embarrassment and more comfortable in social situations and being more able to face people and fit in socially. The phase two confidence and emotions subscale results (and the AS-20 psychosocial subscale results, section 10.5.2) supported the phase one findings. Surgery resulted in significant improvements in psychosocial aspects of HRQoL and confidence and emotions, to the extent that postoperatively the surgery group were not different to those with strabismus who had not sought

surgery (Figures 9-5 and 9-8). The phase two results from individual questions in the confidence and emotions subscale were also in agreement with the phase one findings. Postoperatively the surgery group reported significantly more self-confidence, liking the ways their eyes looked, being more able to talk to people, and feeling happier. They also reported feeling significantly less embarrassed about their eyes, avoiding face-to-face situations less, and feeling less anxious (Table 9-10).

Confidence relating to work

In addition to the phase two results of improved self-confidence, feeling less embarrassed about their eyes and more able to talk to people postoperatively, the surgery group also reported a significant improvement in the way they were treated by others because of their eyes (question 26 – Table 9-13). This was in agreement with the phase one findings of postoperatively being perceived as more friendly and less rude at work, due to being able to communicate with others better. Having increased confidence at work and being better at face-to-face communication led to participants describing themselves as better at their jobs, less stressed at work and being given more work opportunities postoperatively. In phase one, improved face-to-face communication was also an important factor in postoperatively having the self-confidence to pursue work and career development opportunities. Participants reported postoperatively having the confidence to apply for jobs, go for job interviews, apply for university and start a new career due to having improved eye alignment and greater self-confidence in face-to-face situations. Similar findings have been shown in studies using images of people with different eye positions. Strabismus was associated with being less likely to gain employment (Coats et al., 2000), less likely to be promoted at work (Goff et al., 2006), and being perceived negatively by others (Olitsky et al., 1999) compared to having straight eyes.

Confidence in eyes and vision

In phase two, despite a lack of improvement in confidence in abilities (question 28), a significant improvement in confidence in vision was measured postoperatively (question 27) (Table 9-13). Three participants in phase one participants described having greater confidence in their eyes and their vision postoperatively (section 6.6.2). This was typically described when performing a task, such as driving, and was associated with feeling able to try new activities. Menon et al. (2002) reported patients who underwent surgery for psychosocial reasons tried new activities or activities previously avoided (95%) and changed their plans for the future (37.5%). This was in addition to improvements in appearance, relationships with others, self-esteem and self-confidence (97.5%). Nelson et al. (2008) also reported teenagers and adults (n=128) who underwent surgery for psychosocial reasons described improved abilities to meet new people (65%), interpersonal relationships (27%) and abilities to try new activities (16%) postoperatively. These benefits were in addition to improved self-esteem (85%) and postoperative satisfaction with eye alignment (98%) reported during telephone interviews to complete questionnaires. Younger patients reported

greater improvements postoperatively and a larger preoperative deviation was associated with greater improvements in self-esteem and self-image postoperatively. Ghiasi et al. (2013) used a similar questionnaire to Nelson et al. (2008). Three months postoperatively all aspects of the questionnaire improved, with patients reporting improved relationships (82%), being able to meet new people (79%), being better at their job or work (76%), having improved chances of employment (53%) and being able to try new activities (36%), as well as improved self-esteem (89%) (Ghiasi et al., 2013). Xu et al. (2012) reported 36% of their cohort (n=56) had strabismus surgery for improved eye alignment. Whilst improved appearance (96%), and self-esteem or self-confidence (96%) were the most common postoperative outcomes; trying activities previously avoided (82%) and changing plans for the future (68%) were also frequently reported.

10.3 What worsened following strabismus surgery?

10.3.1 Task performance

The surgery group grooved pegboard performance worsened slightly, but not significantly, in both lighting conditions. However, as the control group grooved pegboard performance improved in both lighting conditions, this was interpreted as surgery leading to worsened grooved pegboard performance, in both lighting conditions. Jones and Lee (1981) suggested any advantage of using the two eyes may be more evident under dimmer conditions, however no such advantage was measured after strabismus surgery. Grooved pegboard performance in the surgery group was slightly (but not significantly) faster in bright conditions whereas the control group was significantly faster in bright conditions at both visits (Figure 9-22). The reason for this difference between the groups in bright conditions was unclear. Jones and Lee (1981) described the advantage of using two eyes, even in dim lighting conditions where stereopsis was used less, as binocular concordance. Binocular summation was not measured by Jones and Lee (1981) however, it is possible that binocular summation may be considered as a measure of the binocular concordance described.

Task performance under different lighting conditions has been measured by others to evaluate the effects of binocular vision and different amounts of visible information during tasks. Prehension under illuminated and non-illuminated conditions was investigated in subjects with normal vision (Loftus et al., 2004; Watt & Bradshaw, 2000) and under bright and dim lighting, with high and low contrast stimuli, in amblyopes with strabismus. The strabismus group, however, was very mixed. Some had microtropia and stereopsis and some had no stereopsis. Amblyopes had poorer prehension compared to controls, but contrast affected performance more than lighting (Grant & Conway, 2015). In phase two, contrast sensitivity both eyes open was not significantly different between the groups and did not change significantly in either group (Figure 9-17).

Similar to grooved pegboard performance, bead threading may have worsened postoperatively as the control group had much greater improvement in bead threading time than the surgery group. The within groups improvement was only significant for large bead threading in the control group (Figure 9-24). However, there was no significant difference between the two groups at visit one or visit two. Large beads were always threaded significantly faster than small beads, which was also found by O'Connor et al. (2010a), however their bead threading results were not directly comparable to phase two due to the difference in beads. There was a practice effect, as bead threading was faster at the second visit in both groups, despite a standardised practice before each measurement. If surgery had led to worsened bead threading performance postoperatively, it would have been expected to affect small bead threading more than large, as this was a potentially harder task.

Performance of the grooved pegboard and bead threading involve prehension (reach to grasp) during the task. In adults with BSV, binocular performance improved grasp accuracy (Watt & Bradshaw, 2000). In prehension studies of adult amblyopes, the effects of amblyopia, strabismus and reduced stereopsis have been difficult to disentangle. Amblyopia is associated with worse prehension but an improved ability to use monocular cues (Grant et al., 2007). The effects of amblyopia extend beyond reduced VA, as anisometropic amblyopia reduced reaching and eye hand coordination more than induced blur (Niechwiej-Szwedo et al., 2012). In children, improved stereopsis improved prehension more than improved VA after amblyopia treatment (Grant et al., 2014). Melmoth et al. (2009) reported that adults without stereopsis have poorer prehension both eyes open but monocular performance equal to those with normal vision. However, in contrast to the findings of Grant et al. (2007), Melmoth et al. (2009) found no monocular advantage from long term lack of stereopsis.

Two female participants in phase two (01 and 07) commented that task performance was more difficult postoperatively for the tasks requiring objects to be picked up (pegboards and bead threading) as they had much longer fingernails compared to preoperatively. Both remarked that improved confidence postoperatively led to them growing and painting their fingernails and socialising more. Fingernail length was not controlled in phase two. Other factors affecting task performance were also difficult to control, for example underlying health problems and individual variations in manual dexterity. Participant 18 was unable to complete the task performance measures at phase two as she suffered from arthritis and reported this was painful. Whilst all efforts were made to control the experimental conditions, it is possible that extraneous factors may have affected the results in this repeated measures study.

10.3.2 Vision

Ocular movements (OM)

In phase one, improved eye movements were reported by participants and this was most commonly described as being able to look further into different positions of gaze. Improved eye movements were associated with expanded peripheral vision, needing to make fewer head movements to see at the side and the eyes feeling better during eye movements. In phase two, an expanded visual field was not measured postoperatively, however as will be discussed later (section 10.4.1), only those with XT were included and the expected reduction of visual field area was not measured. In phase two, OM grading was mostly unchanged, however the left LR restriction or underaction increased slightly, but significantly (Figure 9-13). As the surgery group all underwent surgery for an XT, a postoperative LR limitation may have been a desired outcome to limit postoperative XT drift.

Despite the LR limitation possibly being a desired postoperative outcome, it would have been expected to cause a difficulty looking into different positions of gaze, rather than the improvement described in phase one. An improvement in sequential stereopsis may have led to a perceived improvement in eye movements, however no improvements in sequential stereopsis were measured. PROMs questions relating to eye movements and field of vision included the additional study questions 6 and 8. The surgery group reported a significant improvement in question 6 ('my vision has looked central') and no significant change in question 8 ('I have found it difficult to move my eyes to look around'). Although, as discussed (section 10.2.3), the question 8 ceiling effect may have limited the ability of this question to capture change postoperatively.

Smooth pursuit accuracy

The surgery group had significantly reduced horizontal smooth pursuit accuracy postoperatively (Figure 9-31). The surgery group all underwent horizontal strabismus surgery to reduce their horizontal strabismus. This may be the reason that horizontal smooth pursuit accuracy worsened, whereas for vertical smooth pursuit, accuracy was not significantly different. Whilst this reduction in accuracy was statistically significant, it is unclear whether this would be clinically significant or perceptible postoperatively. As discussed in the above paragraph, there was no significant change in the response to additional study question 8 ('I have found it difficult to move my eyes to look around').

10.4 What remained unchanged following strabismus surgery?

10.4.1 Vision

Visual field

In phase two of this study, all the surgery group had a significant reduction in their XT postoperatively, yet the expected reduction in visual field size both eyes open was not measured. The reason for this lack of reduction in visual field was unclear. It was possible that the I4e target

did not measure the full extent of the peripheral visual field, as a larger target (such as the V4e) may have done. However, the I4e target was selected to be comparable to studies that have shown an expanded visual field postoperatively in ET Wortham and Greenwald (1989) (n=10), Kushner (1994) (n=35) and Murray et al. (2007) (n=13). It is possible that the visual field both eyes open measured using this Goldmann method did not relate to the peripheral visual field perceived by patients in phase one. This was supported by the significantly improved postoperative response to additional study question 6 ('my vision has looked central'). It is also possible that surgery to reduce XT causing psychosocial symptoms, does not reduce the peripheral visual field in every patient. This was supported by the phase one finding of improved peripheral vision postoperatively (n=7), three of whom had XT preoperatively and reported their peripheral vision was larger or better postoperatively. One participant described losing his panoramic vision made his vision more normal. Cooper and Feldman (1979) investigated five subjects with intermittent XT and measured an increased peripheral field of vision (panoramic vision) when subjects were manifest and a smaller field of vision when they were aligned. However, no other studies have compared field of vision in patients with constant XT pre and postoperatively.

Using the Goldmann perimeter to measure visual field both eyes open in strabismus and an expanded visual field related to the change in ET was reported by Wortham and Greenwald (1989) (n=10) and Kushner (1994) (n=35). Both reported the expanded visual field was unaffected by amblyopia, however one patient with poor unilateral VA and retinal abnormalities had an unchanged visual field (Kushner, 1994). Some patients reported subjective improvements in peripheral vision or seeing with both eyes simultaneously, but subjective reports were not documented in all patients (Kushner, 1994; Wortham & Greenwald, 1989). Seeing with both eyes simultaneously was suggested to be related to postoperative binocularity, as most (29 of the 35) reported BSV with Bagolini lenses (Kushner, 1994). Murray et al. (2007) reported the outcomes of seventeen infantile ETs who were surgically aligned within 8PD with 'late' surgery (age 8 years or older). Most (88%) had BSV (Bagolini lenses) and 76% had an expanded visual field postoperatively. All who had an expanded visual field, had BSV. Wortham and Greenwald (1989) did not report BSV with Bagolini lenses but reported stereopsis in three (of 10) patients (80" of arc, 400" of arc and 'fly' only) using the Titmus stereotest.

In phase two, the monocular visual field at visit one was significantly smaller than both eyes open, even when split by group (surgery or control) or deviation (ET or XT). The deviating eye contributed to the visual field measured both eyes open and the technique (and target) used were able to measure this difference. This was similar to the findings of Economides et al. (2012) (Figure 1), who used a Goldmann perimeter (size V target, with undisclosed brightness) to measure a larger visual field with both eyes open in XT (n=5). There was no significant difference in the both eyes open visual field area between those with ET and XT. This suggested the XTs did not have as large an expanded visual field as expected, the difference between the visual field in ET and XT

was smaller than expected, or the target used was not sufficient to measure the expanded visual field in XT.

Visual acuity (VA) and contrast sensitivity

In phase one, participants reported postoperatively vision was clearer, sharper, and more in focus. In phase two there was no significant change in VA at near (both eyes open), VA at distance (monocularly with either eye), or contrast sensitivity (both eyes open) postoperatively. No significant change in distance VA after strabismus surgery has been reported by others (Ganguly & Pradhan, 2011; Gupta et al., 2017; Mojon, 2007, 2008; Rajavi et al., 2013; Turan et al., 2015). Near VA and contrast sensitivity measurements pre and postoperatively are not reported in the literature. The PROMs questions relating to improved vision are discussed in section 10.2.3 and related to unchanged vision will be discussed in section 10.4.3.

Sequential stereopsis

Surgery had no effect on sequential stereopsis and performance worsened in all participants. In subjects with BSV, making eye movements to continuously look between targets was shown to improve the ability to discriminate depth (Enright, 1991, 1996). This study used the same stimuli as Frisby et al. (1997), with a similar experimental set up to Frisby et al. (1997) and Taroyan et al. (2000). The phase two monocular results were similar to the monocular results reported by Frisby et al. (1997). However, this study found no advantage of viewing with both eyes open and no improvement in performance postoperatively. Participants reported the task was difficult and they felt they were guessing, rather than making a judgement of stimuli position. Many reported that the clues they relied on to make depth and position judgements were missing and they felt unable to perform the task. This is similar to the findings reported by Frisby et al. (1997) for monocular performance of the task.

It is unclear why the sequential stereopsis results were not significantly better with the unfiltered stimuli compared to the high pass filtered stimuli, as expected. The unfiltered stimuli were selected to be visible to the fixing eye using peripheral vision and the high pass filtered stimuli were selected to be visible only during foveal viewing. It is possible that the unfiltered stimuli were not visible enough using peripheral vision in this group of strabismic adults, as this was not explicitly tested prior to the task. It is also possible that individual variation exists, as during monocular viewing, two of the four participants had better performance with the unfiltered stimuli, one had minimal difference and one had worse performance with the unfiltered stimuli (Frisby et al., 1997).

Fixation grade

At visit one the surgery group had a significantly poorer fixation grade than the control group (Figure 9-12), which was expected due to significantly poorer VA in the strabismic eye (Figure 9-9). This was in agreement with the surgery group's perception of swapping fixation (question 5) (Table

9-10). In phase one, some participants described being able to use their strabismic eye more postoperatively (section 6.3.2). However, in phase two surgery had no effect on fixation grade (Figure 9-12). Only two of the surgery group had a change in fixation grade, participant 15 (3 to 7) had BSV and an X postoperatively and participant 21 (2 to 3). Despite no measurable change in fixation, the surgery group reported an improvement in control of their eye position (question 2), ability to look through both eyes at the same time (question 5) and ability to use their eyes together (question 11) postoperatively (Table 9-10).

Busy environments

Improvements in vision in busy environments were reported in phase one (section 6.3.2).. It is possible that these perceived visual improvements in busy environments related to binocular summation or fixation, both of which are discussed below. It is also possible that these perceived improvements relate to the improved eye alignment and feeling more comfortable in social situations. Results relating to confidence and emotions are discussed in section 10.2.6. The AS-20 psychosocial subscale results will be discussed in section 10.5.2. Despite this reported improvement in phase one, in phase two the perception of being confused by vision (question 4) did not change postoperatively (Table 9-10).

Binocular summation

Individuals with strabismus have reported difficulties in visually stimulating, or busy environments. Whilst binocular summation was found to be reduced in those with strabismus (n=20) compared to normal controls, it was not significantly affected by background noise (Pineles et al., 2014). Results were not presented for those without stereopsis (n=11) separately, however, the addition of visual noise led to worse binocular summation in XT and better binocular summation in ET (Pineles et al., 2014). It is therefore possible that XTs are particularly affected by visual noise. In this study all the surgery group had XT preoperatively and improved binocular summation at 100% contrast.

In subjects with normal BSV the addition of ‘visual noise’ significantly reduced binocular summation, particularly when each eye was presented with noise in a different area of the visual field (Otto et al., 2010), yet others have shown noise increased binocular summation, particularly for stimuli presented peripherally (Wakayama et al., 2012). Binocular summation differed when stimuli were presented foveally or peripherally, with increased binocular summation of stimuli presented peripherally (Zlatkova et al., 2001). However, in manifest strabismus, viewing with both eyes open means the stimulus falls on central retina of the fixing eye and on peripheral retina of the deviating eye. It is possible that improved binocular summation postoperatively reflected the improved alignment of the deviating eye and the stimulus falling on a more central (less peripheral) retinal area of the deviating eye, which is also a more similar area of retina to the fixing eye.

Fixation (EMR)

In phase two a picture background simulated a busy environment for the fixation task. Fixation of the target against the picture background was worse than against a jumbled background for all participants, but strabismus surgery had no effect on fixation performance. In is unclear why the surgery group (n=5) performed the fixation task worse than the control group (n=6) on downgaze (8 degrees) against the busy picture background (both visits) and against the jumbled background (visit two). This may represent a difference in the groups or a difference in their visual function. The small numbers may have limited the EMR analysis. Fixation studies have reported increased drift, saccadic intrusions and nystagmus in amblyopia and strabismus (Ciuffreda et al., 1979a, 1979b). Using a similar method to this study, strabismus (in children) rather than amblyopia, was associated with taking longer to make the first fixation, fewer fixations and less time fixating in the interest area (Al-Haddad et al., 2019). Other strabismus studies have measured less stable fixation only in the strabismic eye during monocular viewing (Ciuffreda et al., 1979b; Ghasia et al., 2018) or during viewing with both eyes open (Economides et al., 2016). Ghasia et al. (2018) measured less stable fixation in larger angle strabismus and those without stereopsis, but fixation was not measured pre and postoperatively for comparison to this study. Fixation was recorded from the fixing eye during both eyes open viewing in this study due to inaccuracy recording the deviating eye in a moderate to large manifest strabismus. Both eyes open viewing without filters, was selected as a more natural viewing condition.

The fixation task included visual search, as participants were not told the location of the target and were instructed to find the cross and look directly at the centre. Visual search time has been shown to be longer in eye conditions such as age related macular degeneration (AMD) (Taylor et al., 2017; Thibaut et al., 2019) and glaucoma (Lee et al., 2020) compared to normal controls. The percentage of time the eye was fixating at different locations on the screen was reported in this study and this could be considered as an indirect measure of visual search, as longer fixation in the interest area implied faster visual search. Strabismus surgery had no effect on fixation (or visual search), but for both groups fixation was worse against the busy background.

Smooth pursuit and saccades (EMR)

In phase one, improved eye alignment was perceived as the eyes moving together more. This may have represented improved eye alignment during OM as there was no change in EMR postoperatively. In phase two surgery did not improve, or worsen, smooth pursuit or saccades, with the exception of horizontal smooth pursuit accuracy which worsened (discussed in section 10.3.3). Smooth pursuits (velocity gain, vertical smooth pursuit accuracy and duration of fixation during smooth pursuit) and saccades (gain, peak velocity and latency) were all unchanged postoperatively. Bucci et al. (2002) measured saccades in children (n=8) pre and 2 weeks–5 months post strabismus surgery. BSV was present preoperatively (n=3) and postoperatively (n=4). No significant change was measured in saccade accuracy, mean peak velocity or post saccadic drift postoperatively, however saccade disconjugacy was reduced. Saccade disconjugacy was not

measured in this study, as both eyes open recording in the presence of a large angle strabismus was considered inaccurate in the deviating eye. Bucci et al. (2002) acknowledged their finding of reduced saccade disconjugacy may be the consequence of recording saccades in eyes that have been aligned surgically, however they preferred the explanation of a central adaptive mechanism leading to improved binocularity. Bucci et al. (2009) measured individual and combined saccades and vergence eye movements in adults and children (n=9, of which n=6 had no BSV), pre and 3-10 months postoperatively. Improved saccade accuracy at near (20cm) was measured, but not at distance (150cm). Improved mean velocity was measured for convergence and saccades combined with divergence. In phase two the EMR viewing distance was 93.1cm. It is therefore possible that any changes in saccade accuracy found by Bucci et al. (2009) were not measured due to the closer testing distance in this study. Vergences and saccades combined with vergences were also not measured in this study.

10.4.2 Task performance

Postoperatively 11 of the 13 participants in phase one described task performance improvements, including driving, tasks at work, using computers or screens, near tasks, practical activities requiring eye hand coordination, reading, balance and depth perception. In phase two, improved CKAT aiming and TSL task time (discussed in section 10.2.2) but worsened grooved pegboard and bead threading performance were measured (discussed in section 10.3.1). Purdue pegboard pin insertion and assembly building, TSL accuracy, CKAT tracking and CKAT steering performances were all unchanged after surgery and there was no significant difference in performance between the groups. The lack of change in the accuracy measurements contrasts with the phase one findings where participants reported being able to perform tasks better, more accurately, or with fewer mistakes postoperatively.

The Purdue pegboard pin insertion was a simpler task than the assembly task. The more complex assembly task required two handed dexterity and was expected to detect change in performance similar to the grooved pegboard, yet the results differed. Other task performance studies have reported different results for different tasks, highlighting the difficulty selecting tasks for inclusion in studies of task performance. Schiller et al. (2012) found needle threading and a touch screen task were better than a rod insertion task at differentiating between levels of stereopsis. Sheedy et al. (1986) reported tasks requiring stereopsis and fine motor skills (bead threading and inserting pointers into straws) were better able to measure binocular advantage than throwing or counting letters within text. Bead threading and Purdue pegboard (pin insertion) were reported to be more sensitive than water pouring at measuring task performance in different levels of BSV (O'Connor et al., 2010a, 2010b; Piano & O'Connor, 2013). Purdue peg insertion and bead threading were significantly worse in those with poorer and absent BSV (O'Connor et al., 2010a, 2010b). The

mean number of pegs inserted in those with absent BSV (15.2-15.59) (O'Connor et al., 2010a, 2010b) was slightly higher than the phase two result (Mdn=14).

10.4.3 PROMs – vision

As described in section 10.2.3, most phase one participants reported their vision improved postoperatively. However, in phase two VA, CS, visual field, sequential stereopsis, fixation grade, binocular summation at 10% and 5% contrast and EMR's (except horizontal smooth pursuit accuracy) were unchanged. Some of the PROMs questions were also unchanged after surgery (Table 9-10) and these are discussed below.

Additional study questions – vision subscale

In phase two no significant change was measured in the surgery group's perception of focussing (question 3), being confused by vision (question 4), difficulty seeing in busy places (question 7), eye movements (question 8) and one eye working harder than the other (question 9) postoperatively (Table 9-10). This contrasted with the improvements reported in phase one. As there was little reported difficulty moving the eyes (question 8) and no significant change postoperatively, this suggested that the increase in LR limitation measured (section 10.3.2) was not perceived by the surgery group.

10.4.4 PROMs – task performance

VFQ-25 driving subscale

In phase one, driving was reported as improved by seven participants (7/13). Improved driving was associated with improved vision, for example vision was described as clearer, more in focus, and better when looking around. Looking around during driving was described as using the strabismic eye more postoperatively to achieve greater peripheral vision. Having greater peripheral vision postoperatively was reported to make driving easier and lead to increased confidence when driving, due to needing to make less head movements to look to the side. In phase two there were nine drivers in each group. Despite descriptions of improved driving in phase one, there was no effect of surgery on the VFQ-25 driving subscale score in phase two. The surgery group perceived their vision to be more central postoperatively (section 10.2.3), however there was no effect of surgery on visual field area (section 10.4.1).

Some participants commented that they had never driven because of their eyes. During the development of the AS-20 driving was specifically excluded from the questions, as it was considered preferable to include questions that related to all or most of the population. Despite driving being a commonly mentioned topic during interviews in patients with strabismus with diplopia (82%) and without diplopia (69%), it was recognised that not all patients drive (Hatt et al., 2009b).

In a pilot study comparing driving performance in normal BSV and ET with suppression, stereopsis led to improved driving slalom performance at intermediate distances. However, participants with suppression were able to estimate the relative position of two cars better than those with BSV (Bauer et al., 2001). Others have used driving simulation to investigate driving in participants with good VA (Brooks et al., 2017). Age was found to be a significant factor, as reaction time reduced with increasing age. In older drivers, in-car recording of driving performance (Owsley et al., 2018), measured a higher rate of crashes in those with slower visual processing speed, impaired motion perception and impaired peripheral visual field sensitivity (Swain et al., 2021). Objective measures of driving and driving simulation were beyond the scope of this feasibility study, however no significant change in self-reported visual function related to driving was measured.

Additional study questions – task performance

There was no significant change in the surgery group's perception of their eyes limiting abilities (question 12), making mistakes (question 14) and balance (question 17) postoperatively (Table 9-11). Despite reported improvements in these task related aspects in phase one, no significant changes or improvements were reported in phase two.

10.4.5 PROMs – physical symptoms

Additional study questions – physical symptoms

The analysis of individual questions relating to physical symptoms showed that despite reported postoperative improvements in the eyes feeling tight in phase one, in phase two surgery did not lead to a significant change in the response to question 21 ('My eyes have felt tight'). The control group reported a significant improvement to question 19 ('I have needed to take breaks because of my eyes'). The surgery group reported no significant change, but a ceiling effect was evident at both visits.

10.4.6 PROMs – confidence and emotions

Additional study questions – confidence and emotions

In phase one, two participants reported greater confidence in their ability to perform a task or activity postoperatively. In phase two, there was a small but not statistically significant improvement in response to question 28 ('I have had confidence in my abilities'). It is possible there was a small improvement, or any improvement was noticed by a small number of participants.

10.5 Were the expected outcomes of strabismus surgery achieved?

Strabismus surgery undertaken for psychosocial reasons was expected to reduce the size of the deviation and improve HRQoL measured by the AS-20 PROM. The extent to which these expected outcomes were achieved will be discussed below.

10.5.1 Expected reduction in deviation size

As expected, surgery significantly reduced the deviation size and those with larger deviations preoperatively had larger changes in deviation postoperatively (Figure 9-2). Postoperatively the surgery group also reported an improvement in their perception of the eyes turning (additional study question 1, section 10.2.3). Preoperatively the surgery group had a similar near deviation and a significantly larger distance deviation compared to the control group. Whilst every effort was made to recruit a comparable control group, it was not possible to match the groups by deviation size, which may have affected the control group results.

The success, partial success and failure criteria described by Hatt et al. (2012a) were adapted to use APCT measurements of the deviation. The phase two postoperative results (n=12) were categorised as success (n=7), partial success (n=4) and failure (n=1) (Table 9-2). Of the four partial successes, three had deviations of 10PD at near or distance or both. In all three cases the patients were happy with the postoperative deviation size and clinically the outcome would be considered successful. The criteria of less than 10PD being considered a success may be too strict and 10PD or less may be preferable. One partial success had diplopia ‘sometimes’, as they experienced diplopia when fixing with their strabismic eye. Postoperatively they were happy with their eye alignment but were actively trying to fix with their preferred eye to eliminate diplopia. They reported diplopia occurred more than ‘never’ or ‘rarely’. The one failure postoperatively had a large overcorrection of their deviation in the distance with diplopia and a LR restriction. Postoperatively they wore blenderm occlusion initially, then later a Fresnel prism to restore suppression. The prism strength had been gradually reduced, but the patient remained with a larger than desired consecutive distance ET. They were happy with their eye alignment postoperatively, but were not happy with their diplopia, as it was disabling without the Fresnel prism.

10.5.2 Expected improvements in HRQoL (AS-20)

AS-20 psychosocial subscale

The surgery group reported significantly worse median psychosocial subscale scores (33.75) than the control group (75) at visit one ($p<0.0001$) (Figure 9-5). This was expected due to the surgery group actively seeking treatment to realign their eyes and improve some aspect of their HRQoL. Surgery led to a significant improvement in the psychosocial subscale score (33.75 to 86.25) ($p<0.01$), with no significant change in the control group (76 to 80). Postoperatively there was no significant difference in psychosocial subscale score between the groups (surgery 86.25, control

80). Using the 95% limits of agreement, an increase of greater than 17.7 in the psychosocial subscale score was considered significant (Leske et al., 2010). Nine of the surgery group (9/12) and none of the control group (0/15) reported this amount of change (Table 9-3).

In a prospective study of adult strabismus surgery outcomes at 3 months (n=61) Koc et al. (2013) reported AS-20 overall and psychosocial scores were not significantly different in those with and without BSV. Visual benefit, or BSV postoperatively, was not required for an improvement in HRQoL. The median improvement in AS-20 psychosocial score in the subgroup without BSV was approximately 52, which was similar to the phase two results. Alam et al. (2014) reported primary strabismus surgery outcomes in older children and adults (n=30). None had diplopia; however, it is unclear whether any had BSV. Significant improvements in AS-20 overall score, psychosocial and function subscales, were measured 6 weeks and 3 months postoperatively, with greater improvement in females. Compared to the phase two results of this study, Alam et al. (2014) reported higher AS-20 scores at 3 months, but less improvement in AS-20 scores postoperatively. Glasman et al. (2013) reported AS-20 results in a prospective study of 17-76-year-olds (n=86). BSV and diplopia were not reported, therefore the cohort may have included surgery for visual benefit. Preoperative AS-20 scores were similar to the phase two results. Whilst surgery improved all aspects of the AS-20, the postoperative AS-20 scores reported by Glasman et al. (2013) (overall 73.1, psychosocial 77.5, function 70), were not as high as in phase two (overall 86.87, psychosocial 86.25, function 91.25). Glasman et al. (2013) reported greater improvements in AS-20 scores in females, those with larger changes of the deviation and those with smaller strabismus postoperatively.

Strabismus surgery has been reported to improve and normalise symptoms of anxiety and depression, HRQoL, daily functioning and psychological adjustment postoperatively (Jackson et al., 2006). Xu et al. (2016) measured improved, but not normal HRQoL postoperatively, although greater HRQoL improvements were measured in those who perceived they had no strabismus postoperatively. Kim et al. (2016) found young adult males with strabismus reported poorer self-identity compared to those who had undergone surgery in childhood and those without strabismus. Those who had undergone surgery reported no difference in self-identity compared to those with no strabismus. In a prospective study of patients aged 17-88 years (n=210), Adams et al. (2016) used the AS-20 and other QoL and psychosocial outcome measures. It is unclear how many had surgery for psychosocial reasons, however 44% had no diplopia. Postoperatively there was a reduction in patients reporting poor AS-20 HRQoL, from 85% to 68%, at 3 and 6 months postoperatively. Other measures of social anxiety and avoidance, clinical anxiety, and depression improved significantly. McBain et al. (2016b) reported the same cohort, concluding that at 3 months postoperatively there were improvements in a wide range of psychosocial domains and the AS-20. Improvements in HRQoL were not associated with clinical judgements of success.

Sim et al. (2018) reported preoperatively patients without diplopia had significantly lower AS-20 psychosocial subscale scores than patients with diplopia, but there was no difference in the AS-20 function subscale results. Patients without diplopia had larger improvements in AS-20 psychosocial subscale scores postoperatively but did not achieve scores as high as those who had diplopia (preoperatively). The preoperative results of Sim et al. (2018) (30.6) were similar to phase two (33.75), however the postoperative results (67.5) were lower than phase two (86.25). Sim et al. (2018) additionally reported lower socioeconomic status was associated with greater improvement in HRQoL postoperatively. Socioeconomic status was not measured in this study. Similar to Sim et al. (2018), Hatt et al. (2010a) found those without diplopia reported lower preoperative AS-20 psychosocial scores and a larger improvement in AS-20 overall and psychosocial subscale scores postoperatively. Unlike McBain et al. (2016b), Hatt et al. (2010a) reported a significant improvement in HRQoL (AS-20 score) was associated with a clinical judgement of success, however their low numbers with partial success (n=5) and failure (n=2) compared to success (n=19) may have limited this analysis.

A clinical judgement of 'failure' postoperatively did not preclude improved AS-20 HRQoL (Hatt et al., 2016). In a later study investigating failure of HRQoL to improve postoperatively (n=276), postoperative diplopia was associated with no change or worsening AS-20 score, in addition to a distressed personality type, worse depressive symptoms postoperatively and coexisting facial abnormalities (Hatt et al., 2018). Results of surgery for psychosocial reasons were not presented separately, however 79% had diplopia preoperatively. Surgery for psychosocial reasons may therefore have been a small proportion of this cohort. Mental health was not investigated in this study, however mental health, as well as clinical factors, may influence postoperative AS-20 outcomes (Adams et al., 2016; Hatt et al., 2018; McBain et al., 2016b).

AS-20 function subscale

Strabismus surgery resulted in a significant improvement in AS-20 function subscale score, although significant improvement was also seen in the control group (Figure 9-5). Despite this control group increase in AS-20 function subscale score, at visit two the surgery group (91.25) reported significantly higher AS-20 function scores than the control group (77.50). Improvements in the AS-20 function scale have also been reported by others, however many other studies include strabismus surgery for visual benefit, as well as psychosocial, making exact comparison to results of this study challenging. Hatt et al. (2010a) reported a subgroup without diplopia (n=26) who were most comparable to the phase two surgery group. Preoperatively the without diplopia group had higher AS-20 function subscale results (66.3) compared to the surgery group in phase two (56.25). Postoperatively depending on surgical success, Hatt et al. (2010a) reported an improvement in function subscale score of 7.5 (success), 10.0 (partial success) and a worsening of 8.8 (failure). The postoperative phase two results were significantly higher in the surgery group (from 56.25 to 91.25). The reason for this difference in results is unclear. The low numbers in the without diplopia

group (n=26) and the further subgroups may have limited the analysis. It is also possible that some patients without diplopia may have undergone surgery for visual benefit, as BSV and potential BSV were not reported.

None of the control group (0/15) and most of the surgery group (8/12) scores exceeded the 95% limits of agreement for the AS-20 function subscale (Leske et al., 2010) (Table 9-3). Hatt et al. (2012b) reported a smaller proportion achieved this amount of change in AS-20 function score. Thirty-eight surgeries were performed in patients without diplopia in their large cohort (n=171 surgical procedures in n=159 patients) and only 8 (of the 38) had a change in AS-20 function score greater than the 95% limits of agreement. It is possible that the earlier measurement of postoperative outcome used by Hatt et al. (2012b), median 7 weeks, range 4-19 weeks, may have affected their postoperative results compared to this study.

Koc et al. (2013) reported similar AS-20 function scores and postoperative improvement in AS-20 function (34) compared to phase two (32.5) in their 'without BSV' subgroup. Those without BSV had improvement in AS-20 function scores that were not significantly different from those with BSV postoperatively, until amblyopes were removed from the analysis. Koc et al. (2013) did not report those with and without diplopia separately, however 21% of their cohort had diplopia, suggesting that their without BSV subgroup was likely to include a number of patients having surgery for psychosocial reasons. Glasman et al. (2013) reported lower median AS-20 function scores than found in phase two, 46.3 preoperatively and 70 postoperatively. They did not report diplopia or BSV, so it is unclear how many of their cohort (n=86) had surgery for psychosocial symptoms. Reasons for their lower scores are unclear, however it may relate to the variable postoperative measurement, mean 91 days, range 12-362 days.

Liebermann et al. (2014) reported adult surgery for constant childhood onset strabismus without diplopia (n=20). Surgery was undertaken for psychosocial reasons (n=15), eye strain (n=3) and difficulty localising objects (n=2). One year postoperatively (Mdn=12 months, range 6-19 months) five had stereopsis at near (FNS), distance (FD2) or both. It was unclear whether visual benefit was expected as potential BSV was not investigated. Pre and postoperative AS-20 results were not reported, instead significant improvement in individual items in the function subscale was reported. Nine of the ten items (questions 12-20) significantly improved. Only question 11 relating to closing one eye did not change significantly. Using an alternative method of scoring the AS-20 (Rasch-derived scores (Leske et al., 2012)), there was significant improvement in reading function (items 12, 13, 16 and 20) and general function (items 11, 15, 17, 18) domains. The cohort reported by Liebermann et al. (2014) was most similar to the surgery group and the results support the findings of both phase one and phase two of this study. Despite strabismus surgery for psychosocial reasons, study participants reported a number of different functional and visual improvements using interviews and the AS-20.

Alam et al. (2014) reported higher median pre (81.2) and 3 months postoperative (97.5) AS-20 function scores compared to phase two. All patients underwent primary surgery for strabismus without diplopia (n=30), however BSV was not reported. Some patients may have achieved BSV postoperatively, which may be the reason for higher function scores. McBain et al. (2016a) reported QoL and psychosocial outcomes, including the AS-20, at three and six months postoperatively (n=210). A number of patients had surgery for visual benefit (123 had diplopia and 43 had previously used prisms) but the number having surgery to achieve BSV was not reported. Mean AS-20 function scores for the whole cohort were 56.44 preoperatively, 68.29 three months postoperatively and 69.42 six months postoperatively. The preoperative function score was similar to the surgery group in phase two, however the postoperative scores were lower. The reason for this difference was unclear. Similar to McBain et al. (2016a), Sim et al. (2018) reported a comparable preoperative AS-20 function score (50) but a lower postoperative score (72.5) compared to the phase two results. Their group had no diplopia (n=33) but BSV was not reported. Similar to Glasman et al. (2013), Sim et al. (2018) used a variable time to report postoperative outcome, mean 57 days, range 24-126 days. Akbari et al. (2015) also presented a large cohort who underwent surgery for a range of strabismus types. BSV was not reported, but those without diplopia (101 of 112) had a median AS-20 (Persian version) function score that significantly increased postoperatively (from approximately 70 to 82). Compared to the phase two results, the preoperative AS-20 function score was higher, however the increase in function score and the postoperative score were lower.

10.6 Success, partial success and failure criteria

The definition of a success, partial success or failure outcome postoperatively was discussed relating to deviation size in section 10.5.1. Further statistical analysis of the outcome measures most likely to measure change was limited by the small numbers in each of the outcome groups (Table 9-14). However, preliminary comparisons revealed that the improvement in binocular summation at 100% contrast was higher in the failure participant and partial success group than the success group. However, at 5% contrast the failure participant and partial success groups had a worsening of binocular summation compared to the success group who had an improvement in binocular summation. No previous evidence has compared binocular summation results by these criteria. The CST results revealed a larger improvement in response time in the success group, compared to the partial success and failure participant, however the improvement in RCS score was larger in the failure participant and partial success group. The improvement in pegboard completion time was greatest in the failure participant, followed by the success group, then the partial success group.

All the improvements in AS-20 scores were greatest in the success group, followed by the partial success group and the failure participant (Table 9-14). This was not in agreement with the results of Hatt et al. (2012a) who reported variable improvements in AS-20 scores across similarly defined groups. In a subgroup of patients without diplopia (n=17), AS-20 psychosocial and function subscale results improved at 6 weeks postoperatively and improved further at 1-year postoperatively. Results defined by success, partial success and failure postoperatively were presented for all patients combined (not just those without diplopia). Those graded as a partial success and failure tended to have an improvement in AS-20 scores at 6 weeks postoperatively, that then reduction at 1-year postoperatively.

The VFQ-25 score improved the most in the failure participant, followed by similar improvements in the success and partial success groups (Table 9-14). In a subgroup of patients without diplopia (n=26) Hatt et al. (2010a) reported reduced postoperative median VFQ-25 scores in their failure group (n=2) and slightly improved median VFQ-25 scores in their success (n=19) and partial success (n=5) groups. The improvements in VFQ-25 scores in all the groups in this study (Table 9-14) were notably higher than those reported by Hatt et al. (2010a). Using the additional study questions, all groups showed improvement in the different subscales scores, with the greatest improvements seen in the success group and the failure participant. Only the median (but not the mean) change in physical symptoms subscale score worsened slightly in the partial success group.

As only one participant was graded as a failure, this limited comparison of the difference between the failure group and other groups. However, this participant was not precluded from achieving improvement postoperatively. They showed improvement in all of the outcomes, except binocular summation at 10% and 5% contrast. Being graded a ‘failure’ and achieving postoperative results still considered clinically and subjectively positive has also been reported (Hatt et al., 2012a, 2016) highlighting the difficultly of classifying postoperative results based on clinical outcomes such as eye alignment.

10.7 Strengths and weaknesses of the study

A strength of the overall study was the mixed methods design. The combination of phase one and phase two results enabled a more comprehensive investigation of the research question, as well as providing valuable results individually. The mixed methods approach also ensured that the views of patients were integrated into the whole study and phase two was not solely based on published research. A consequence of this sequential mixed methods design was the time the overall study took, as phase one had to be completed and analysed to inform the measures selected for phase two.

Inclusion of a clearly defined group of participants having strabismus surgery for psychosocial reasons was particularly important in this study. Whilst outcomes of strabismus surgery may be reported, information about the surgical aims are often lacking. The numbers with and without diplopia preoperatively may be reported, but clear descriptions of those having surgery for visual or psychosocial benefit are far less commonly reported. In light of the need for clear evidence of the outcomes of strabismus surgery undertaken for psychosocial reasons, a clearly defined and representative cohort was required to be able to answer the research question 'what are the effects of strabismus surgery undertaken for psychosocial reasons?'

Information provided to all study participants included a description of the study, however care was taken to word the information communicated in a balanced and non-leading manner to avoid introducing bias or 'demand characteristics' into their responses (Appendices D and M).

10.7.1 Phase one

The semi-structured interviews used to explore patient perceptions of postoperative outcomes had the advantage of ensuring the study design was focussed on patient experiences. The semi-structured interviews, rather than questionnaire or PROM completion, had the advantage of gathering richer information from patients about their outcomes. During the interviews further questions could be asked to encourage patients to expand on their answers and give examples to illustrate their views. Participants were asked about positive, neutral and negative experiences postoperatively. Purposive sampling was a strength of the recruitment strategy in phase one. This ensured a mixture of males, females, younger and older participants contributed to the phase one findings. All participants were evenly spread across the four variations; however, a small number of participants were recruited (n=13). It is possible that recruiting additional participants may have enhanced the data collected, however recruitment was stopped as data saturation had been reached.

A potential limitation of the retrospective phase one design was participants were asked to recall their perception of postoperative outcome at one point in time and this was variable amongst participants (range 4.5-20 months). Participants were recruited from the Orthoptic clinic at their routine follow up appointment, which led to variation and unexpected clinical delays. A narrower range of postoperative interview time would have standardised the participant experience and interview more. During the interview's participants were invited to share their experiences and talk about a range of different aspects of their lives. It was therefore anticipated that a range of different factors were reported, and multiple opportunities were given for discussion about different postoperative experiences.

During the interviews open questions were asked in a non-leading manner and were standardised, to avoid introducing bias and demand characteristics. However, it is acknowledged that there is an implicit bias that can be introduced to interviewing techniques, especially when interviews are conducted by a person connected to the research or clinical topic. An independent person may have been perceived as an 'impartial' interviewer by participants and may have encouraged more varied responses, however it was beyond the scope of this study to have a separate person conducting the interviews.

10.7.2 Phase two

A strength of phase two was the inclusion of the control group with strabismus. Control groups have been incorporated into other studies, for example controls with normal VA and BSV to compare to strabismus (Ghasia et al., 2018) or amblyopia (Gonzalez et al., 2012). Others have used visually normal controls in addition to strabismic controls (without amblyopia) to compare to strabismic amblyopes in a binocular summation study (Chang et al., 2017). Kim et al. (2016) studied self-identity in strabismus and used controls with normal vision as well as group who had previously undergone strabismus surgery. The phase two condition specific control group was selected over a control group with normal BSV and no strabismus. Postoperative BSV was not the expected outcome of surgery and comparison to a group with 'normal vision' was considered a biased comparison. Despite not being the purpose of this study, it is acknowledged that using an additional control group with normal vision and BSV would have enabled comparison of performance with strabismus (with and without surgery) to performance with normal BSV.

The phase two control group enabled comparison of the findings of the surgery group to the control group and ensured interpretation of any effects of surgery were made in comparison to the expected practice effect in strabismic controls. However, a limitation of phase two was the difference between the groups. Whilst the control group (n=15) were considered a good representation of a cohort with strabismus not actively seeking surgery, they were not selected purposefully to match the surgery group (n=12) and most had undergone previous strabismus surgery in childhood or early adulthood (n=13). A control group would ideally include those who had not undergone any strabismus surgery, but this was difficult to achieve in this study. The control group all had manifest strabismus and suppression but a significantly smaller distance deviation (M=14.8, Mdn=12) compared to the surgery group (M=29.4, Mdn=30). The control group included ET (n=9) and XT (n=5) whereas the surgery group had XT (n=12). The range of deviations in each group differed (control 40PD ET to 30 PD XT, surgery 10PD XT to 55PD XT). This was likely to have influenced the results, as the control group had smaller deviations.

The small surgery group (n=12) was a limitation of phase two and limited further analysis. This was partly due to covid-19, however phase two recruitment was challenging prior to the covid-19

pandemic. The surgery group all had XT, which limited analysis and comparison of deviation direction. Direction of strabismus was found not to affect the postoperative outcome (Ghiasi et al., 2013), yet others have shown greater HRQoL improvements in ET compared to XT (Burke et al., 1997). Analysis and comparison of ET and XT would have enhanced the visual field results and enabled greater comparison to the literature (Kushner, 1994; Wortham & Greenwald, 1989). The small surgery group also limited further exploratory analysis of other factors such as gender, deviation size, amblyopia and postoperative outcomes (success, partial success and failure). Gender has been shown to be a significant factor in some studies of strabismus, but not all. Females have reported greater improvements in HRQoL (Akbari et al., 2015; Alam et al., 2014; Burke et al., 1997; Glasman et al., 2013), yet other studies have shown no effect of gender (Ghiasi et al., 2013).

The timing of the second visits for all participants was planned to occur at the same as the surgery group, however clinical appointment delays typically led to the surgery group having their final visit later than three months postoperatively. This difference in visit two timing may have affected the results. The control group however, were not expected to have a significant change at their second visit, therefore an earlier assessment is likely to have impacted on the results less than an earlier assessment of the surgery group. The timing between visits could be included as a covariate in future analysis to explore the significance of this variation.

The inclusion of a range of different measures of vision and task performance in phase two ensured a range of different aspects were explored in this feasibility study, which was important considering the aim of the study was to determine possible outcome measures for a future larger trial. A limitation of phase two was not all measures were ‘successful’. The sequential stereopsis task was reported as too difficult by all participants. Even using unfiltered stimuli, that should have been visible with eccentric viewing, participants reported they felt they were guessing and not performing the task as required. The CKAT was limited by technical difficulties. In future, technical difficulties would need to be resolved prior to further CKAT data collection.

The visual field testing technique using a I4e target was selected to allow comparison to other results, however the results of unchanged visual field size postoperatively in XT were unexpected considering the expansion previously measured in ET postoperatively (Kushner, 1994; Wortham & Greenwald, 1989). Further assessment of visual field size in ET and XT, and with both I4e and V4e targets would be useful to explore the most accurate measurement technique in all deviations. The small number who had EMR (control n=6, surgery n=5) was a limitation of phase two. EMR required additional time, therefore it was offered to all participants, but completed by a small cohort. During phase two planning it was felt that this approach was a compromise that allowed collection of detailed EMR data in a subgroup.

The AS-20 and VFQ-25 are established PROMs that had been used in other strabismus studies. However, limitations of these PROMs included ceiling effects and the subjective interpretation of questions. For example, a ceiling effect was evident for additional study question 13 ('I have had difficulty with eye hand coordination') in the surgery group and any improvement may have underrepresented. Anecdotally, when completing the PROMs, the surgery group typically reported no eye hand coordination difficulty. Based on their postoperative outcome, most participants were unlikely to have had any experience of BSV, therefore were unaware of any better eye hand coordination previously or having reduced eye hand coordination due to strabismus. Postoperatively, whilst some participants described improved eye hand coordination, their responses to question 13 were less able to capture this. When completing the PROMs, opportunities were given for participants to ask questions, however the subjective interpretation of questions and responses remained. Additional study question one was meant to establish whether participants thought they had a strabismus, yet even the wording of the question 'my eyes have been turning' was questioned by participants to clarify the meaning of the question.

The AS-20 has been evaluated using Rasch analysis (Leske et al., 2012). Two questions were removed due to some participants having difficulty understanding or answering the question. Question 14 ('I have problems with depth perception') and question 19 ('I can't enjoy my hobbies because of my eyes'). Similarly, in this study, some participants expressed difficulty understanding what depth perception was, and often responded that they had no difficulty, even though clinically they had suppression and no BSV. No participants expressed difficulty answering question 19, however Leske et al. (2012) reported not all participants had hobbies and could answer the question. This was considered when creating the additional study questions, as they were not aligned to one activity. The original scoring method of the AS-20 was selected for this study to allow greater comparison to other studies. Both questions 14 and 19 were retained as they related to task performance. Inclusion of 18 questions may have underrepresented task performance compared to psychosocial aspects.

A positive aspect of a PROM that can be used widely in a number of different patients, also creates a limitation. The limitation of the AS-20 specific to this study was that it was developed for use in a wide range of strabismus patients, not just those with psychosocial symptoms. During development of the AS-20, a wide range of possible HRQoL questions was refined to a shorter list of questions that were relevant to many different patients, for example congenital and acquired neurogenic palsies and mechanical conditions. This led to a number of issues that were relevant specifically to those with psychosocial concerns or symptoms, not being included in the AS-20. The additional study questions were used to try and address these gaps.

10.7.3 Study decisions and key milestones

During the study regular meetings were held with supervisors and members of the advisory group to discuss progress and make decisions at key milestones. In phase one, the decision was made to stop data collection after thirteen participants had been recruited because it was felt data saturation had been reached. Whilst continuing recruitment to phase one would have generated additional qualitative data, the additional time required for this recruitment would have impacted on the timescales of the overall study and study progression. The phase one results were required to be able to plan the phase two measures. A drawback of this sequential design was the need for completion of phase one to inform the design of phase two.

Decisions on which phase two measures to include were made following detailed discussions about the phase one findings and how these could be measured. The main concern was testing burden for participants and avoiding each of the phase two visits taking too long, impacting on recruitment and retention. Where several objective measures were available, these were refined to one measure where possible (described in section 7.1 and Appendix I). It was also decided to offer the eye movement recordings to all participants as an additional and optional part of their study visit. Due to initially slow phase two recruitment, recruitment and retention was discussed in detail, with regular input from the PPI representative. Several options were considered, including expanding recruitment to other sites, ensuring all those who met the inclusion criteria were identified and keeping up to date with clinical appointment changes. Following discussion with several potential sites, recruitment was not expanded to other centres for this feasibility study due to issues with surgical waiting list delays and anticipated delays in postoperative follow up appointments.

Several phase two study visits were cancelled due to the covid-19 pandemic. Whilst the situation was discussed, the response to the pandemic meant that few alternative options could be considered. The decision was made to submit an amendment to the ethical approval for phase two, to allow telephone collection of questionnaire data from participants and resumption of phase two study visits when study visits were considered safe to resume. This amendment was approved.

10.7 Contribution of the research to the evidence base

Methodological

This study has uniquely conducted a mixed methods study in a clearly defined group of adults with strabismus and psychosocial symptoms only. A precisely defined cohort of adults with strabismus (without diplopia or measurable potential BSV) was essential to answer questions that have been raised about this specific group of patients as they are at risk of withdrawal of funding for strabismus surgery. An original aspect of the study was the qualitative phase that explored

postoperative outcomes from the patient's perspective. Commonly reported clinical outcomes of deviation size and HRQoL (AS-20) do not capture the full picture of postoperative outcome. The qualitative findings were uniquely used to inform the measures selected for phase two and develop additional PROM questions to fill the gaps in existing PROM instruments (AS-20 and VFQ-25) when used in this specific cohort of adults with strabismus with psychosocial symptoms. The quantitative phase of the study was designed to prospectively measure vision, task performance, physical symptoms, confidence and emotions. The condition specific control group was a distinctive element of the quantitative phase of the study, and it enabled a much greater understanding of the effect of surgery and practice effect of the individual tests when performed by adults with strabismus.

The feasibility and acceptability of the study design and quantitative measures have been considered. Further consideration of the feasibility included a sample size calculation for a future larger trial (Appendix X). These results will be used to inform the planning of a future larger trial (section 10.8).

Theoretical

The findings of this research add to our understanding of the contribution of the apparently suppressed eye. A clinical finding of suppression with Bagolini glasses should not be interpreted as the strabismic eye not being used at all. Participants described being able to use their suppressed eye in a variety of ways, particularly after their deviating eye was surgically aligned to a straighter position. The suppressed eye may to contribute to the quality of vision, field of vision, and to some degree of binocular cooperation between the eyes when both eyes are open.

Contribution to policy and practice

The results of this study will be published and disseminated to clinicians and academics to inform clinical practice guidelines and recommendations, and strabismus surgery policy decisions, particularly regarding funding for strabismus surgery undertaken for psychosocial reasons. The results add new information to the existing evidence and show that adults with strabismus and psychosocial symptoms gained significant benefits from strabismus surgery and improved eye alignment. Their HRQoL improved significantly, as measured by the AS-20, but they additionally gained much greater and broader improvements in their confidence and emotions, vision, task performance and physical symptoms that were not captured by a PCT measurement and the AS-20.

Most study participants had a successful postoperative outcome and uniquely the wider impact of these postoperative outcomes to the patient were captured and measured. The wider societal consequences of improved HRQoL, confidence and emotions, vision, task performance and physical symptoms included patients having the confidence to interact and communicate with

people, taking on new opportunities, applying for work and career development opportunities, and putting themselves forwards for work, educational and social opportunities that they would not have considered preoperatively.

Subjectively patients reported their vision was better and objectively improved binocular summation at 100% contrast was measured. Subjectively patients reported improved task performance and objectively improved TSL time and CKAT aiming time were measured, which may have been small effects of surgery. Not all task performance measures yielded the same results and some aspects of task performance appeared to worsen postoperatively. Further investigation is required to analyse task performance outcomes and whether task performance improvements are related to surgery or a practice effect. Subjectively patients reported improvements in the physical symptoms caused by strabismus. Headaches, the feeling of the eyes pulling, feeling tired and feeling the eyes turning were all improved postoperatively. The improvements in the psychosocial aspects of HRQoL and confidence and emotions were significant and particularly important to participants. Improved HRQoL and confidence and emotions may have additionally influenced the subjective perception of improved vision, task performance and physical symptoms.

10.8 Directions for further research

The results of this feasibility study could inform a new core outcome set specifically for adults with strabismus and psychosocial symptoms. Additionally, the results have led to several areas for future research that will be described in the following sections.

10.8.1 Future larger trial

Using the results of this feasibility study, a future larger trial of strabismus surgery undertaken for psychosocial reasons is the next step. A multicentre trial would allow recruitment of more participants from a wider area. Incorporating a condition-specific control group into the larger trial would allow ongoing investigation of possible practice effects.

The future larger trial should be powered for statistical significance and could use the outcomes measures PCT, binocular summation and PROMs as all were acceptable to patients and were most likely to measure change postoperatively. Additional outcome measures could also include a measure of task performance such as the TSL where time and accuracy are measured, the CST and visual field area. Using the same study design and analysis method in this feasibility study, sample sizes for the future trial were calculated (Appendix X). It was considered feasible to include either the AS-20 (overall score), binocular summation (100%) or CST RCS with both eyes open as primary outcome measures as all had a sample size of 20 per group or less. However, using TSL

time as the primary outcome measure was considerably less feasible as the sample size was calculated to be 179 per group.

The additional study questions should be refined and used in addition to the AS-20 to measure the HRQoL impact of strabismus and surgery, but also the wider aspects of vision, task performance, physical symptoms and confidence and emotions that are more specific to strabismus with psychosocial symptoms.

It was possible to recruit and retain participants to phase two, however if study measures were quicker to complete in the eye clinic recruitment would be expected to improve. A larger trial would allow further investigation of strabismus subgroups and other potential factors that may affect postoperative outcome. These include timing of the postoperative visit, gender, direction of strabismus, amblyopia and clinical categorisation of the outcome (for example success / partial success / failure). Subgroup analysis could also include other factors such as those unexpectedly gaining BSV compared to those who do not. A larger cohort would allow further analysis of how much each of the clinical factors relates to, or predicts the surgical outcome, for example binocular summation.

In summary, the outcome measures used in this study that would not be included in a future larger trial would be visual acuity (near and distance), contrast sensitivity, fixation grading, ocular movements, sequential stereopsis, grooved pegboard, Purdue pegboard, bead threading, VFQ-25 and EMR of fixation and saccades. Careful consideration would be given to the potential impact of additional factors that would be difficult to control, such as fingernail length, on task performance. Additional data using the CKAT would be required before it could be considered as an outcome measure. The outcome measures that could be included as primary outcome measures include the AS-20, binocular summation and CST RCS with both eyes open. Possible secondary outcome measures could include TSL time (although there is concern about the large sample size calculated), visual field area and a bolt on the AS-20 developed from the additional study questions.

10.8.2 Research into testing techniques

Areas for further study include establishing the optimum testing technique to measure visual field area in strabismus, binocular summation using letter charts (physical charts or electronic based methods where randomisation is possible) and gross BSV using the CST. Further investigation of the practice effects that were evident for some of the measures would allow future studies to include appropriate learning time or practice trials. EMRs were used in this study to measure fixation, smooth pursuit and saccades. Further measurements of eye movements during the

postoperative recovery period would be useful to further our knowledge of healing, recovery and adaptation following strabismus surgery.

10.8.3 Future research into strabismus and strabismus surgery outcomes

Future strabismus research questions arising from this research also include how accurate are our clinical investigations of potential BSV, and whether postoperative BSV affects surgical outcome from the perspective of the clinician and patient? If postoperative BSV is shown to be affect the surgical outcome, how much does gross or subnormal BSV matter compared to normal BSV or no BSV postoperatively?

Qualitative research methods have been useful in this study to explore patient perceptions of postoperative surgical outcomes. Further investigation of patient decision-making around strabismus surgery with qualitative research methods would also be useful, in particular the reasons why people do and do not seek surgery. Further investigation of patient perceptions of treatment outcomes in different clinical definitions of success, partial success and failure postoperatively would also provide useful clinical results that would inform clinical practice. Longer term postoperative outcomes following strabismus surgery are also an under researched area. Greater evidence of the longer-term outcome using both qualitative and quantitative research methods would be useful to add to our existing knowledge of the outcomes of strabismus surgery and would also enable us to progress towards measuring the value and cost effectiveness of strabismus surgery more accurately.

10.9 Conclusions

The original aim of conducting a mixed methods feasibility study to investigate the outcomes from strabismus surgery in adults having surgery for psychosocial reasons has been achieved. Interviews during the qualitative phase one elicited participant perceptions of their outcomes from surgery and these findings were used to refine the design of the quantitative phase two. Participants were prospectively recruited to phase two, to both a surgery group and a control group of adults with strabismus, not seeking surgery for psychosocial reasons.

This study has shown that it was feasible to measure change in some aspects of visual function and task performance postoperatively. Most of the measures used in this study were acceptable to patients and these have been refined as part of the planning of a future larger trial. Overall, it was feasible to prospectively recruit and retain patients to phase two, where measures were performed before and after surgery.

The results of this study have answered the primary research question. The effects of strabismus surgery undertaken for psychosocial reasons were to improve vision (binocular summation at

100% contrast and CST performance with both eyes open) and task performance (TSL performance time and CKAT aiming time). Visual field size did not reduce following alignment of XT. Surgery also improved vision, task performance, physical symptoms and confidence and emotions from the participants perspective, using PROMs. These changes were in addition to the known and expected improvements in eye alignment (PCT) and HRQoL (AS-20). Worsening of OM, smooth pursuit accuracy and task performance (grooved pegboard and bead threading) may also occur following surgery.

Prior to this study there was concern from commissioners of NHS services in some areas of England that not enough benefit was proven in adults with strabismus without expected functional visual gains from surgery. Clinicians considered strabismus surgery for adults with psychosocial problems to be highly beneficial for patients, yet robust evidence in this specific cohort was lacking. In adults undergoing strabismus surgery for psychosocial reasons this study has captured and measured significant postoperative improvements in HRQoL, but also much broader improvements in their vision, task performance, physical symptoms and confidence and emotions that were significantly beneficial to the patient and to wider society. The findings of this study can be used to inform future clinical practice guidelines and policy decisions on strabismus surgery funding. The effects of strabismus surgery undertaken for psychosocial reasons were to significantly improve HRQoL and confidence and emotions, as well as improve some aspects of vision, task performance and physical symptoms caused by strabismus.

References

Açık, A., Onat, S., Schumann, F., Einhäuser, W., & König, P. (2009). Effects of luminance contrast and its modifications on fixation behavior during free viewing of images from different categories. *Vision Res, 49*(12), 1541-1553. doi:10.1016/j.visres.2009.03.011

Açık, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental Changes in Natural Viewing Behavior: Bottom-Up and Top-Down Differences between Children, Young Adults and Older Adults. *Front Psychol, 1*(207). doi:10.3389/fpsyg.2010.00207

Adams, G. G., McBain, H., MacKenzie, K., Hancox, J., Ezra, D. G., & Newman, S. P. (2016). Is strabismus the only problem? Psychological issues surrounding strabismus surgery. *J AAPOS, 20*(5), 383-386. doi:10.1016/j.jaapos.2016.07.221

Akbari, M. R., Masoomian, B., Ameri, A., Keshtkar Jaafari, A., & Mirmohammadsadeghi, A. (2015). Development of a Persian Version of the Adult Strabismus Questionnaire and Evaluating the Effect of Strabismus Surgery on Health-Related Quality of Life. *Strabismus, 23*(2), 66-72. doi:10.3109/09273972.2015.1025986

Al-Haddad, C., & Abdul Fattah, M. (2017). Slipped extraocular muscles: characteristics and surgical outcomes. *Can J Ophthalmol, 52*(1), 42-47. doi:10.1016/j.jcjo.2016.07.016

Al-Haddad, C., Hoyeck, S., Torbey, J., Houry, R., & Boustany, R. N. (2019). Eye Tracking Abnormalities in School-Aged Children With Strabismus and With and Without Amblyopia. *J Pediatr Ophthalmol Strabismus, 56*(5), 297-304. doi:10.3928/01913913-20190726-01

Al-Saud, L. M., Mushtaq, F., Mirghani, I., Balkhoyor, A., Keeling, A., Manogue, M., & Mon-Williams, M. A. (2017). Drilling into the functional significance of stereopsis: the impact of stereoscopic information on surgical performance. *Ophthalmic Physiol Opt, 37*(4), 498-506. doi:10.1111/opo.12393

Al-Wadaani, F. A. (2017). Outcome of strabismus surgery by nonadjustable suture among adults attending a university hospital of Saudi Arabia. *Niger J Clin Pract, 20*(3), 335-340. doi:10.4103/1119-3077.196086

Al Jabri, S., Kirkham, J., & Rowe, F. J. (2019). Development of a core outcome set for amblyopia, strabismus and ocular motility disorders: a review to identify outcome measures. *BMC Ophthalmol, 19*(1), 47. doi:10.1186/s12886-019-1055-8

Alam, D., Khan, A. A., Bani, S. A., Sharma, R., & Amitava, A. K. (2014). Gain beyond cosmesis: demonstration of psychosocial and functional gains following successful strabismus surgery using the adult strabismus questionnaire adult strabismus 20. *Indian J Ophthalmol, 62*(7), 799-803. doi:10.4103/0301-4738.138620

Alberti, C. F., & Bex, P. J. (2018). Binocular contrast summation and inhibition depends on spatial frequency, eccentricity and binocular disparity. *Ophthalmic Physiol Opt, 38*, 525-537. doi:10.1111/opo.12581

Aletaha, M., Bagheri, A., Gholipour, H. M., & Kheiri, B. (2016). Effect of Limited Tenon Capsule and Intermuscular Membranes Dissection on the Outcome of Surgery in Patients with Horizontal Strabismus. *Strabismus, 24*(1), 12-15. doi:10.3109/09273972.2015.1130065

Alkharashi, M., & Hunter, D. G. (2017). Reduced surgical success rate of rectus muscle plication compared to resection. *J AAPOS, 21*(3), 201-204. doi:10.1016/j.jaapos.2017.05.012

Almuklass, A. M., Feeney, D. F., Mani, D., Hamilton, L. D., & Enoka, R. M. (2017). Peg-manipulation capabilities during a test of manual dexterity differ for persons with multiple sclerosis and healthy individuals. *Exp Brain Res, 235*, 3487-3493.

Alpak, G., Coskun, E., Erbagci, I., Bez, Y., Okumus, S., Oren, B., & Gurler, B. (2014). Effects of corrective surgery on social phobia, psychological distress, disease-related disability and quality of life in adult strabismus patients. *Br J Ophthalmol, 98*(7), 876-879. doi:10.1136/bjophthalmol-2014-304888

Ansons, A. M., & Davis, H. (2014). *Diagnosis and Management of Ocular Motility Disorders* (4th Edition ed.): John Wiley & Sons, Incorporated.

Arditi, A. (2004). Improving Letter Contrast Sensitivity Testing. *Invest Ophthalmol Vis Sci, 45*(13), 4583-4583.

Astle, A. T., Foulsham, T., Foss, A. J., & McGraw, P. V. (2016). Is the frequency of adult strabismus surgery increasing? *Ophthalmic Physiol Opt, 36*(4), 487-493. doi:10.1111/opo.12306

Awad, A. H., Mullaney, P. B., Al-Hazmi, A., Al-Turkmani, S., Wheeler, D., Al-Assaf, M., . . . Al-Mesfer, S. (2000). Recognized globe perforation during strabismus surgery: incidence, risk factors, and sequelae. *J AAPOS, 4*(3), 150-153.

Bailey, I. L., & Lovie–Kitchin, J. E. (2013). Visual acuity testing. From the laboratory to the clinic. *Vision Res, 90*, 2-9.

Baker, D. H., Meese, T. S., Mansouri, B., & Hess, R. F. (2007). Binocular summation of contrast remains intact in strabismic amblyopia. *Invest Ophthalmol Vis Sci, 48*(11), 5332-5338. doi:10.1167/iovs.07-0194

Balcer, L. J., Raynowska, J., Nolan, R., Galetta, S. L., Kapoor, R., Benedict, R., . . . Rudick, R. (2017). Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *Mult Scler, 23*(5), 734-747. doi:10.1177/1352458517690822

Ball, A., Drummond, G. T., & Pearce, W. G. (1993). Unexpected stereoacuity following surgical correction of long-standing horizontal strabismus. *Can J Ophthalmol, 28*(5), 217-220.

Barrett, B. T., Panesar, G. K., Scally, A. J., & Pacey, I. E. (2012). A Limited Role for Suppression in the Central Field of Individuals with Strabismic Amblyopia. *PLoS One, 7*(5), e36611. doi:10.1371/journal.pone.0036611

Barrett, B. T., Panesar, G. K., Scally, A. J., & Pacey, I. E. (2013). Binocular summation and other forms of non-dominant eye contribution in individuals with strabismic amblyopia during habitual viewing. *PLoS One, 8*(10), e77871. doi:10.1371/journal.pone.0077871

Bartlett, H. L., Ting, L. H., & Bingham, J. T. (2014). Accuracy of force and centre of pressure measures of the Wii Balance Board. *Gait & Posture, 39*(1), 224-228. doi:10.1016/j.gaitpost.2013.07.010

Bauer, A., Dietz, K., Kolling, G., Hart, W., & Schiefer, U. (2001). The relevance of stereopsis for motorists: a pilot study. *Graefe's Arch Clin Exp Ophthalmol, 239*(6), 400-406. doi:10.1007/s004170100273

Bayramlar, H., & Gunduz, A. (2006). Medium term outcomes of strabismus surgery in patients with monocular dense amblyopia. *Saudi Med J, 27*(2), 276-278.

BBC News. (2018, 30 June 2018). NHS England to stop 'ineffective' treatments. *BBC News website*. Retrieved from https://www.bbc.co.uk/news/health-44665560 (Accessed 2/7/20)

Beauchamp, C. L., Beauchamp, G. R., Stager, D. R., Sr., Brown, M. M., Brown, G. C., & Felius, J. (2006). The cost utility of strabismus surgery in adults. *J AAPOS, 10*(5), 394-399. doi:10.1016/j.jaapos.2006.06.004

Beauchamp, G. R., Black, B. C., Coats, D. K., Enzenauer, R. W., Hutchinson, A. K., Saunders, R. A., . . . Felius, J. (2005a). The management of strabismus in adults--III. The effects on disability. *J AAPOS, 9*(5), 455-459. doi:10.1016/j.jaapos.2005.04.009

Beauchamp, G. R., Black, B. C., Coats, D. K., Enzenauer, R. W., Hutchinson, A. K., Saunders, R. A., . . . Felius, J. (2005b). The management of strabismus in adults--II. Patient and provider perspectives on the severity of adult strabismus and on outcome contributors. *J AAPOS, 9*(2), 141-147. doi:10.1016/j.jaapos.2004.11.015

Beauchamp, G. R., Black, B. C., Coats, D. K., Enzenauer, R. W., Hutchinson, A. K., Saunders, R. A., . . . Felius, J. (2003). The management of strabismus in adults-I. Clinical characteristics and treatment. *J AAPOS, 7*(4), 233-240. doi:10.1016/s1091-8531(03)00112-5

Beauchamp, G. R., Felius, J., Stager, D. R., & Beauchamp, C. (2005c). The Utility of Strabismus in Adults. *Trans Am Ophthalmol Soc., 103*, 164-172.

Berland, J. E., Wilson, M. E., & Saunders, R. B. (1998). Results of large (8-9 mm) bilateral lateral rectus muscle recessions for exotropia. *Binocul Vis Strabismus Q, 13*(2), 97-104.

Bez, Y., Coskun, E., Erol, K., Cingu, A. K., Eren, Z., Topcuoglu, V., & Ozerturk, Y. (2009). Adult strabismus and social phobia: A case-controlled study. *J AAPOS, 13*(3), 249-252. doi:doi:10.1016/j.jaapos.2009.02.010

Biglan, A. W., Davis, J. S., Day, R., & Landsittel, D. (1994). Prospective investigation of the effectiveness of intraoperative adjustable sutures for correction of strabismus. *Trans Am Ophthalmol Soc., 92*, 325-347.

Billington, C. (2018). The Use of Clinical Evidence in Commissioning Adult Strabismus Surgery (Poster-Gloucestershire Hospitals NHS Foundation Trust). Retrieved from https://www.gloshospitals.nhs.uk/work-for-us/training-staff/gsqia/quality-improvements/The-Use-of-Clinical-Evidence-in-Commissioning-Adult-Strabismus-Surgery/: https://www.gloshospitals.nhs.uk/work-for-us/training-staff/gsqia/quality-improvements/The-Use-of-Clinical-Evidence-in-Commissioning-Adult-Strabismus-Surgery/ (Accessed 7/11/20)

Bloch, E., Uddin, N., Gannon, L., Rantell, K., & Jain, S. (2015). The effects of absence of stereopsis on performance of a simulated surgical task in two-dimensional and three-dimensional viewing conditions. *Br J Ophthalmol, 99*(2), 240-245. doi:10.1136/bjophthalmol-2013-304517

Bradbury, J. A., & Taylor, R. H. (2013). Severe complications of strabismus surgery. *J AAPOS, 17*(1), 59-63. doi:10.1016/j.jaapos.2012.10.016

Bristol North Somerset and South Gloucestershire Clinical Commissioning Group. (2019). *Surgical Correction of Strabismus or Amblyopia in Adults Policy (version 1920.01.02)*. Retrieved from https://bnssgccg-media.ams3.cdn.digitaloceanspaces.com/attachments/Strabismus_and_Amblyopia_Policy_PA_v1920.01.02_draft.pdf (Accessed 7/11/20)

British and Irish Orthoptic Society. (2015). Guidelines for the Diagrammatic Representation of Ocular Movement. 1-12.

British Standard. (2003). Visual acuity test types. Part 1: Test charts for clinical determination of distance visual acuity - Specification. In *British Standard* (pp. 1-10).

Brooks, J., Seeanner, J., Hennessy, S., Manganelli, J., Crisler, M., Rosopa, P., . . . Tanner, S. (2017). Interactive Tools for Measuring Visual Scanning Performance and Reaction Time. *Am J Occup Ther, 71*(2), 1-6. doi:10.5014/ajot.2017.020461

Bruce, A., & Santorelli, G. (2016). Prevalence and Risk Factors of Strabismus in a UK Multi-ethnic Birth Cohort. *Strabismus, 24*(4), 153-160. doi:10.1080/09273972.2016.1242639

Bryman, A. (2016). *Social Research Methods* (Fifth ed.). Oxford, UK.: Oxford University Press.

Bucci, M. P., Brémond-Gignac, D., & Kapoula, Z. (2009). Speed and accuracy of saccades, vergence and combined eye movements in subjects with strabismus before and after eye surgery. *Vision Res, 49*(4), 460-469. doi:10.1016/j.visres.2008.12.006

Bucci, M. P., Kapoula, Z., Bernotas, M., & Zamfirescu, F. (1999). Role of attention and eye preference in the binocular coordination of saccades in strabismus. *Neuro-Ophthalmol, 22*(2), 115-126. doi:10.1076/noph.22.2.115.3736

Bucci, M. P., Kapoula, Z., Yang, Q., & Brémond-Gignac, D. (2006). Latency of saccades, vergence, and combined movements in children with early onset convergent or divergent strabismus. *Vision Res, 46*(8), 1384-1392. doi:10.1016/j.visres.2005.06.035

Bucci, M. P., Kapoula, Z., Yang, Q., Roussat, B., & Brémond-Gignac, D. (2002). Binocular Coordination of Saccades in Children with Strabismus before and after Surgery. *Invest Ophthalmol Vis Sci, 43*(4), 1040-1047.

Bucci, M. P., Soufi, H., Villeneuve, P., Colleville, L., Bui-Quoc, E., & Lions, C. (2016). Importance of Proprioceptive Information for Postural Control in Children with Strabismus before and after Strabismus Surgery. *Front Syst Neurosci, 10*(67). doi:10.3389/fnsys.2016.00067

Buckley, D., Codina, C., Bhardwaj, P., & Pascalis, O. (2010a). Action video game players and deaf observers have larger Goldmann visual fields. *Vision Res, 50*(5), 548-556. doi:10.1016/j.visres.2009.11.018

Buckley, J. G., Panesar, G. K., MacLellan, M. J., Pacey, I. E., & Barrett, B. T. (2010b). Changes to Control of Adaptice Gait in Individuals with Long-standing Reduced Stereoacuity. *Invest Ophthalmol Vis Sci, 51*, 2487-2495. doi:10.1167/iovs.09-3858

Burke, J. P., Leach, C., & Davis, H. (1997). Psychosocial implications of strabismus surgery in adults. *J Pediatr Ophthalmol Strabismus, 34*(3), 159-164.

Cameron, R. (2011). Mixed Methods Research: The Five Ps Framework. *Electron J Bus Res Methods, 9*(2), 96-108.

Carkeet, A., & Bailey, I. L. (2017). Slope of psychometric functions and termination rule analysis for low contrast acuity charts. *Ophthalmic Physiol Opt, 37*, 118-127. doi:10.1111/opo.12357

Chang, M. Y., Demer, J. L., Isenberg, S. J., Velez, F. G., & Pineles, S. L. (2017). Decreased Binocular Summation in Strabismic Amblyopes and Effect of Strabismus Surgery. *Strabismus, 25*(2), 73-80. doi:10.1080/09273972.2017.1318153

Chang, M. Y., Velez, F. G., Demer, J. L., Isenberg, S. J., Coleman, A. L., & Pineles, S. L. (2015). Quality of life in adults with strabismus. *Am J Ophthalmol, 159*(3), 539-544 e532. doi:10.1016/j.ajo.2014.12.003

Chapman, G. J., Scally, A., & Buckley, J. G. (2012). Importance of binocular vision in foot placement accuracy when stepping onto a floor-based target during gait initiation. *Exp Brain Res, 216*(1), 71-80. doi:10.1007/s00221-011-2910-x

Charmaz, K. (2014). *Constructing Grounded Theory* (Second Edition ed.): SAGE Publications Ltd.

Chung, D. C., McCague, S., Yu, Z. F., Thill, S., DiStefano-Pappas, J., Bennett, J., . . . High, K. A. (2018). Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clin Exp Ophthalmol, 46*, 247-259. doi:10.1111/ceo.13022

Cifuentes, D. L., Pineles, S. L., Demer, J. L., & Velez, F. G. (2018). Surgical success and lateral incomitance following three-muscle surgery for large-angle horizontal strabismus. *J AAPOS, 22*(1), 17-21. doi:10.1016/j.jaapos.2017.10.005

Ciuffreda, K. J., Kenyon, R. V., & Stark, L. (1979a). Fixational eye movement in amblyopia and strabismus. *J Am Optom Assoc, 50*(11), 1251-1258.

Ciuffreda, K. J., Kenyon, R. V., & Stark, L. (1979b). Saccadic Intrusions in Strabismus. *Arch Ophthalmol, 97*(9), 1673-1679. doi:10.1001/archopht.1979.01020020241012

Clark, R. A., Bryant, A. L., Pua, Y., McCrory, P., Bennell, K., & Hunt, M. (2010). Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait & Posture, 31*(3), 307-310. doi:10.1016/j.gaitpost.2009.11.012

Coats, D. K., Paysse, E. A., Towler, A. J., & Dipboye, R. L. (2000). Impact of large angle horizontal strabismus on ability to obtain employment. *Ophthalmol, 107*(2), 402-405. doi:10.1016/S0161-6420(99)00035-4

Coats, D. K., Stager, D. R., Sr., Beauchamp, G. R., Stager, D. R., Jr., Mazow, M. L., Paysse, E. A., & Felius, J. (2005). Reasons for delay of surgical intervention in adult strabismus. *Arch Ophthalmol, 123*(4), 497-499. doi:10.1001/archopht.123.4.497

COMET Initiative: Core Outcome Measures in Effectiveness Trials. *comet-initiative.org (accessed 17/9/21)*.

Connolly, J. D., & Goodale, M. A. (1999). The role of visual feedback of hand position in the control of manual prehension. *Exp Brain Res, 125*(3), 281-286.

Cooper, J., & Feldman, J. (1979). Panoramic Viewing, Visual Acuity of the Deviating Eye, and Anomalous Retinal Correspondence in the Intermittent Exotrope of the Divergence Excess Type. *Am J Optom Physiol Opt, 56*, 422-429.

Cotter, S. A., Tarczy-Hornoch, K., Song, E., Lin, J., Borchert, M., Azen, S. P., & Varma, R. (2009). Fixation preference and visual acuity testing in a population-based cohort of preschool children with amblyopia risk factors. *Ophthalmol, 116*(1), 145-153. doi:10.1016/j.ophtha.2008.08.031

Crotty, M. (1998). *The Foundations of Social Research.*: SAGE Publications.

Culmer, P. R., Levesley, M. C., Mon-Williams, M., & Williams, J. H. G. (2009). A new tool for assessing human movement: The Kinematic Assessment Tool. *J Neurosci Methods, 184*(1), 184-192. doi:10.1016/j.jneumeth.2009.07.025

Cunningham, A. C., Hill, L., Mon-Williams, M., Peall, K. J., Linden, D. E. J., Hall, J., . . . van den Bree, M. B. M. (2019). Using kinematic analyses to explore sensorimotor control impairments in children with 22q11.2 deletion syndrome. *J Neurodev Dis, 11*(1), 8-8. doi:10.1186/s11689-019-9271-3

Currie, Z. I., Shipman, T., & Burke, J. P. (2003). Surgical correction of large-angle exotropia in adults. *Eye, 17*(3), 334-339. doi:10.1038/sj.eye.6700347

Dadeya, S., Kamlesh, & Fatima, S. (2002). Preliminary results of intraoperative daunorubicin in strabismus surgery. *J Pediatr Ophthalmol Strabismus, 39*(6), 340-344.

Das, A., Hancox, J., & RCOphth Quality and Safety Group. (2017). *Strabismus surgery for adults in the United Kingdom: indications, evidence base and benefits. Commissioning Guidance.* Retrieved from https://www.rcophth.ac.uk/standards-publications-research/commissioning-in-ophthalmology/ (Accessed 6/3/21):

Dawson, E. L., Leung, H., Webster, A., & Lee, J. P. (2013). Beneficial effect of treatment for strabismus in patients with bilaterally poor vision, who cannot see their strabismus. *Strabismus, 21*(1), 33-36. doi:10.3109/09273972.2012.742119

de Jongh, E., Leach, C., Tjon-Fo-Sang, M. J., & Bjerre, A. (2014). Inter-examiner variability and agreement of the alternate prism cover test (APCT) measurements of strabismus performed by 4 examiners. *Strabismus, 22*(4), 158-166. doi:10.3109/09273972.2014.972521

Deacon, M. A., & Gibson, F. (2001). Strabismus measurements using the alternating and simultaneous prism cover tests: a comparative study. *J Pediatr Ophthalmol Strabismus, 38*(5), 267-272.

Dickmann, A., Aliberti, S., Rebecchi, M. T., Aprile, I., Salerni, A., Petroni, S., . . . Balestrazzi, E. (2013). Improved sensory status and quality-of-life measures in adult patients after strabismus surgery. *J AAPOS, 17*(1), 25-28. doi:10.1016/j.jaapos.2012.09.017

Dolven, J., Lucas, J., & West, R. (2011). Influence of the Direction of Eye Turn on the Appearance of Strabismus. *Optom Vis Sci, 88*(5), 628-634. doi:10.1097/OPX.0b013e3182107cbc

Dorr, M., Kwon, M. Y., Lesmes, L. A., Miller, A., Kazlas, M., Chan, K., . . . Bex, P. J. (2019). Binocular Summation and Suppression of Contrast Sensitivity in Strabismus, Fusion and Amblyopia. *Frontiers in human neuroscience, 13*, 234-234. doi:10.3389/fnhum.2019.00234

Dotan, G., Nelson, L. B., Mezad-Koursh, D., Stolovitch, C., Cohen, Y., & Morad, Y. (2014). Surgical outcome of strabismus surgery in patients with unilateral vision loss and horizontal strabismus. *J Pediatr Ophthalmol Strabismus, 51*(5), 294-298. doi:10.3928/01913913-20140709-02

Drover, J. R., Kean, P. G., Courage, M. L., & Adams, R. J. (2008). Prevalence of amblyopia and other vision disorders in young Newfoundland and Labrador children. *Can J Ophthalmol, 43*(1), 89-94. doi:10.3129/i07-187

Durnian, J. M., Noonan, C. P., & Marsh, I. B. (2011). The psychosocial effects of adult strabismus: a review. *Br J Ophthalmol, 95*(4), 450-453. doi:10.1136/bjo.2010.188425

Durnian, J. M., Owen, M. E., Baddon, A. C., Noonan, C. P., & Marsh, I. B. (2010). The psychosocial effects of strabismus: effect of patient demographics on the AS-20 score. *J AAPOS, 14*(6), 469-471. doi:10.1016/j.jaapos.2010.08.013

Durnian, J. M., Owen, M. E., & Marsh, I. B. (2009). The psychosocial aspects of strabismus: correlation between the AS-20 and DAS59 quality-of-life questionnaires. *J AAPOS, 13*(5), 477-480. doi:10.1016/j.jaapos.2009.06.006

Early Treatment Diabetic Retinopathy Study Research Group. (1991). Early Treatment Diabetic Retinopathy Study design and baseline patient characteristics. ETDRS report number 7. *Ophthalmol, 98*(5 Suppl), 741-756.

Economides, J. R., Adams, D. L., & Horton, J. C. (2012). Perception via the deviated eye in strabismus. *J Neurosci, 32*(30), 10286-10295. doi:10.1523/JNEUROSCI.1435-12.2012

Economides, J. R., Adams, D. L., & Horton, J. C. (2016). Variability of ocular deviation in strabismus. *JAMA Ophthalmol, 134*(1), 63-69. doi:10.1001/jamaophthalmol.2015.4486

Eino, D., & Kraft, S. P. (1997). Postoperative drifts after adjustable-suture strabismus surgery. *Can J Ophthalmol, 32*(3), 163-169.

Elkamshoushy, A., & Langue, M. A. (2019). Outcomes of bilateral lateral rectus recession in treatment of recurrent exotropia after bilateral medial rectus resection. *Eur J Ophthalmol, 29*(4), 402-405. doi:10.1177/1120672118795066

Elliott, D. B., Bullimore, M. A., & Bailey, I. L. (1991). Improving the reliability of the Pelli-Robson contrast sensitivity test. *Clin Vis Sci, 6*, 471-475.

Enright, J. T. (1991). Exploring the third dimension with eye movements: better than stereopsis. *Vision Res, 31*(9), 1549-1562.

Enright, J. T. (1996). Sequential stereopsis: a simple demonstration. *Vision Res, 36*(2), 307-312. doi:10.1016/0042-6989(95)00099-L

Escardó-Paton, J. A., & Harrad, R. A. (2009). Duration of conjunctival redness following adult strabismus surgery. *J AAPOS, 13*(6), 583-586. doi:10.1016/j.jaapos.2009.09.013

Estes, K. J., Parrish, R. K., Sinacore, J., Mumby, P. B., & McDonnell, J. F. (2020). Effects of corrective strabismus surgery on social anxiety and self-consciousness in adults. *J AAPOS, 24*(5), 280.e281-280.e284. doi:10.1016/j.jaapos.2020.05.017

Faridi, U. A., Saleh, T. A., Ewings, P., & Twomey, J. M. (2007). Factors affecting the surgical outcome of primary exotropia. *Strabismus, 15*(3), 127-131. doi:10.1080/09273970701506086

Fatima, T., Amitava, A. K., Siddiqui, S., & Ashraf, M. (2009). Gains beyond cosmesis: Recovery of fusion and stereopsis in adults with longstanding strabismus following successful surgical realignment. *Indian J Ophthalmol, 57*(2), 141-143. doi:10.4103/0301-4738.45505

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fawcett, S. L., Felius, J., & Stager, D. R. (2004). Predictive factors underlying the restoration of macular binocular vision in adults with acquired strabismus. *J AAPOS, 8*(5), 439-444. doi:10.1016/j.jaapos.2004.07.001

Felius, J., Stager, D. R., Jr., Beauchamp, G. R., & Stager, D. R. (2001). Re-recession of the medial rectus muscles in patients with recurrent esotropia. *J AAPOS, 5*(6), 381-387. doi:10.1067/mpa.2001.119784

Fieß, A., Elflein, H. M., Urschitz, M. S., Pesudovs, K., Münzel, T., Wild, P. S., . . . Schuster, A. K. (2020). Prevalence of strabismus and its impact on vision-related quality of life - results from the German population-based Gutenberg Health Study (GHS). *Ophthalmol*. doi:10.1016/j.ophtha.2020.02.026

Flatters, I., Hill, L. J. B., Williams, J. H. G., Barber, S. E., & Mon-Williams, M. (2014). Manual control age and sex differences in 4 to 11 year old children. *PLoS One, 9*(2), e88692-e88692. doi:10.1371/journal.pone.0088692

Foerster, R., M., Carbone, E., Koesling, H., & Schneider, W., X. . (2011). Saccadic eye movements in a high-speed bimanual stacking task: Changes of attentional control during learning and automatization. *J Vis, 11(7)*(9), 1-16. doi:doi:10.1167/11.7.9.

Frangouli, O., & Adams, G. G. (2013). The use of amniotic membrane for the management of fibrosis in complex strabismus surgery. *Strabismus, 21*(1), 13-22. doi:10.3109/09273972.2012.762531

Frisby, J. P., Buckley, D., & Duke, P. A. (1996). Evidence for good recovery of lengths of real objects seen with natural stereo viewing. *Perception, 25*, 129-154.

Frisby, J. P., Catherall, C., Porrill, J., & Buckley, D. (1997). Sequential stereopsis using high-pass spatial frequency filtered textures. *Vision Res, 37*(22), 3109-3116. doi:10.1016/S0042-6989(97)00113-2

Fronius, M., & Sireteanu, R. (1994). Pointing Errors in Strbismics: Complex Patterns of Distorted Visuomotor Control. *Vision Res, 34*(5), 689-707.

Fujiike, K., Mizuno, Y., Hiratsuka, Y., & Yamada, M. (2011). Quality of life and cost-utility assessment after strabismus surgery in adults. *Jpn J Ophthalmol, 55*(3), 268-276. doi:10.1007/s10384-011-0022-6

Ganesh, A., Pirouznia, S., Ganguly, S. S., Fagerholm, P., & Lithander, J. (2011). Consecutive exotropia after surgical treatment of childhood esotropia: a 40-year follow-up study. *Acta Ophthalmol, 89*(7), 691-695. doi:10.1111/j.1755-3768.2009.01791.x

Ganguly, S., & Pradhan, R. (2011). Effect of monocular surgery for large-angle horizontal deviation in adults. *Nepal J Ophthalmol, 3*(1), 27-30. doi:10.3126/nepjoph.v3i1.4275

Garip, G., & Kamal, A. (2019). Systematic review and meta-synthesis of coping with retinitis pigmentosa: implications for improving quality of life. *BMC Ophthalmol, 19*(1), 181. doi:10.1186/s12886-019-1169-z

Ghasia, F. F., Otero-Millan, J., & Shaikh, A. G. (2018). Abnormal fixational eye movements in strabismus. *Br J Ophthalmol, 102*(2), 253-259. doi:10.1136/bjophthalmol-2017-310346

Ghiasi, G., Shojaei, A., Soltan-Sanjari, M., Kosari, M., & Aslani, M. (2013). The Psychosocial Improvement after Strabismus Surgery in Iranian Patients. *Iranian J Ophthalmol, 25*(3), 211-215.

Gigante, E., Romão, R. A., & Valério, F. J. P. (2018). Monocular surgery to correct large-angle esotropia: a 10-year follow-up study. *Arq Bras Oftalmol, 81*(3), 232-238. doi:10.5935/0004-2749.20180046

Glasman, P., Cheeseman, R., Wong, V., Young, J., & Durnian, J. M. (2013). Improvement in patients' quality-of-life following strabismus surgery: evaluation of postoperative outcomes using the Adult Strabismus 20 (AS-20) score. *Eye, 27*(11), 1249-1253. doi:10.1038/eye.2013.174

Goff, M. J., Suhr, A. W., Ward, J. A., Croley, J. K., & O'Hara, M. A. (2006). Effect of Adult Strabismus on Ratings of Official U.S. Army Photographs. *J AAPOS, 10*(5), 400-403. doi:10.1016/j.jaapos.2006.04.011

Gonzalez, E. G., Wong, A. M. F., Niechwiej-Szwedo, E., Tarita-Nistor, L., & Steinbach, M. J. (2012). Eye Position Stability in Amblyopia and in Normal Binocular Vision. *Invest Ophthalmol Vis Sci, 53*(5386-5394). doi:10.1167/iovs.12-9941.

Goodwill, S., R. (2013). Check2D - camera calibration tool for 2D kinematic analysis. Technical report - C2D1. http://www.check2d.co.uk/ (accessed 22/11/18).

Goseki, T., & Ishikawa, H. (2017). The prevalence and types of strabismus, and average of stereopsis in Japanese adults. *Jpn J Ophthalmol, 61*(3), 280-285. doi:10.1007/s10384-017-0505-1

Graham, P. A. (1974). Epidemiology of strabismus. *Br J Ophthalmol, 58*(3), 224-231.

Grant, S., & Conway, M. L. (2015). Reach-to-precision grasp deficits in amblyopia: Effects of object contrast and low visibility. *Vision Res, 114*, 100-110. doi:10.1016/j.visres.2014.11.009

Grant, S., Melmoth, D. R., Morgan, M. J., & Finlay, A. L. (2007). Prehension deficits in amblyopia. *Invest Ophthalmol Vis Sci, 48*(3), 1139-1148. doi:10.1167/iovs.06-0976

Grant, S., Suttle, C., Melmoth, D. R., Conway, M. L., & Sloper, J. J. (2014). Age- and stereovision-dependent eye-hand coordination deficits in children with amblyopia and abnormal binocularity. *Invest Ophthalmol Vis Sci, 55*(9), 5687-57015. doi:10.1167/iovs.14-14745

Greenwald, H. S., Knill, D. C., & Saunders, J. A. (2005). Integrating visual cues for motor control: A matter of time. *Vision Res, 45*, 1975-1989. doi:10.1016/j.visres.2005.01.025

Griffiths, H., Whittle, J., & Buckley, D. (2011). The effect of distractors on saccades and adaptation of saccades in strabismus. *Vision Res, 51*(23), 2405-2424. doi:10.1016/j.visres.2011.09.017

Gupta, P., Dadeya, S., Kamlesh, & Bhambhawani, V. (2017). Comparison of Minimally Invasive Strabismus Surgery (MISS) and Conventional Strabismus Surgery Using the Limbal Approach. *J Pediatr Ophthalmol Strabismus, 54*(4), 208-215. doi:10.3928/01913913-20170321-01

Gusek-Schneider, G., & Boss, A. (2010). Results following eye muscle surgery for secondary sensory strabismus. *Strabismus, 18*(1), 24-31. doi:10.3109/09273971003623582

Ha, S., & Kim, S. (2016). Clinical Preoperative and Postoperative Symptoms of Exotropia. *J Pediatr Ophthalmol Strabismus., 53*, 44-51. doi:doi: 10.3928/01913913-20160113-05

Haag-Streit International. Original Goldmann Perimeter 940. Instruction for use., 1-35.

Hanson, W., E., Creswell, J., W., Plano Clark, V., L., Petska, K., S., & Creswell, J., D. (2005). Mixed Methods Research Designs in Counselling Psychology. *J Couns Psychol, 52*(2), 224-235. doi:10.1037/0022-0167.52.2.224

Harding, V. R., de C Williams, A. C., Richardson, P. H., Nicholas, M. K., Jackson, J. L., Richardson, I. H., & Pither, C. E. (1994). The development of a battery of measures for assessing physical functioning of chronic pain patients. *Pain, 58*(3), 367-375. doi:10.1016/0304-3959(94)90131-7

Hashemi, H., Nabovati, P., Yekta, A., Ostadimoghaddam, H., Behnia, B., & Khabazkhoob, M. (2017). The Prevalence of Strabismus, Heterophorias, and Their Associated Factors in Underserved Rural Areas of Iran. *Strabismus, 25*(2), 60-66. doi:10.1080/09273972.2017.1317820

Hashemi, H., Pakzad, R., Heydarian, S., Yekta, A., Aghamirsalim, M., Shokrollahzadeh, F., . . . Khabazkhoob, M. (2019). Global and regional prevalence of strabismus: a comprehensive systematic review and meta-analysis. *Strabismus, 27*(2), 54-65. doi:10.1080/09273972.2019.1604773

Hassan, M. B., Hodge, D. O., & Mohney, B. G. (2015). Prevalence of Mental Health Illness among Patients with Adult-Onset Strabismus. *Strabismus, 23*(3), 105-110. doi:10.3109/09273972.2015.1070881

Hatt, S. R., Leske, D. A., Bradley, E. A., Cole, S. R., & Holmes, J. M. (2009a). Comparison of quality-of-life instruments in adults with strabismus. *Am J Ophthalmol, 148*(4), 558-562. doi:10.1016/j.ajo.2009.05.009

Hatt, S. R., Leske, D. A., Bradley, E. A., Cole, S. R., & Holmes, J. M. (2009b). Development of a quality-of-life questionnaire for adults with strabismus. *Ophthalmol, 116*(1), 139-144 e135. doi:10.1016/j.ophtha.2008.08.043

Hatt, S. R., Leske, D. A., & Holmes, J. M. (2010a). Responsiveness of Health-Related Quality-of-life Questionnaires in Adults Undergoing Strabismus Surgery. *Ophthalmol, 117*(12), 2322-2328.e2321. doi:10.1016/j.ophtha.2010.03.042

Hatt, S. R., Leske, D. A., Kirgis, P. A., Bradley, E. A., & Holmes, J. M. (2007). The Effects of Strabismus on Quality of Life in Adults. *Am J Ophthalmol, 144*(5), 643-647. doi:10.1016/j.ajo.2007.06.032

Hatt, S. R., Leske, D. A., Liebermann, L., & Holmes, J. M. (2012a). Changes in health-related quality of life 1 year following strabismus surgery. *Am J Ophthalmol, 153*(4), 614-619. doi:10.1016/j.ajo.2011.10.001

Hatt, S. R., Leske, D. A., Liebermann, L., & Holmes, J. M. (2012b). Comparing outcome criteria performance in adult strabismus surgery. *Ophthalmol, 119*(9), 1930-1936. doi:10.1016/j.ophtha.2012.02.035.

Hatt, S. R., Leske, D. A., Liebermann, L., & Holmes, J. M. (2016). Incorporating Health-related Quality of Life Into the Assessment of Outcome Following Strabismus Surgery. *Am J Ophthalmol, 164*, 1-5. doi:10.1016/j.ajo.2015.12.029

Hatt, S. R., Leske, D. A., Liebermann, L., Philbrick, K. L., & Holmes, J. M. (2014). Depressive symptoms associated with poor health-related quality of life in adults with strabismus. *Ophthalmol, 121*(10), 2070-2071. doi:10.1016/j.ophtha.2014.04.031

Hatt, S. R., Leske, D. A., Philbrick, K. L., & Holmes, J. M. (2018). Factors Associated With Failure of Adult Strabismus-20 Questionnaire Scores to Improve Following Strabismus Surgery. *JAMA Ophthalmol, 136*(1), 46-52. doi:10.1001/jamaophthalmol.2017.5088

Hatt, S. R., Leske, D. A., Yamada, T., Bradley, E. A., Cole, S. R., & Holmes, J. M. (2010b). Development and initial validation of quality-of-life questionnaires for intermittent exotropia. *Ophthalmol, 117*(1), 163-168.e161. doi:10.1016/j.ophtha.2009.06.038

Hayhoe, M., Gillam, B. J., Chajka, K., & Vecellio, E. (2009). The role of binocular vision in walking. *Vis Neurosci, 26*, 73-80. doi:10.1017/S0952523808080838

Haymes, S. A., Roberts, K. F., Cruess, A. F., Nicolela, M. T., LeBlanc, R. P., Ramsey, M. S., . . . Artes, P. H. (2006). The Letter Contrast Sensitivity Test: Clinical Evaluation of a New Design. *Invest Ophthalmol Vis Sci, 47*(6), 2739-2745. doi:10.1167/iovs.05-1419

Health and Social Care Information Centre. Hospital Episode Statistics, Admitted Patient Care England. 2014-2015, 2015-2016, 2016-2017, 2017-2018, 2018-2019, 2019-2020. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity (Accessed 7/11/20)

Hertle, R. W. (1998). Clinical characteristics of surgically treated adult strabismus. *J Pediatr Ophthalmol Strabismus, 35*(3), 138-145.

Holmes, J. M., Leske, D. A., & Hohberger, G. G. (2008). Defining real change in prism-cover test measurements. *Am J Ophthalmol, 145*(2), 381-385. doi:10.1016/j.ajo.2007.09.012

House, R. J., Rotruck, J. C., Enyedi, L. B., Wallace, D. K., Saleh, E., & Freedman, S. F. (2019). Postoperative infection following strabismus surgery: case series and increased incidence in a single referral center. *J AAPOS, 23*(1), 26.e21-26.e27. doi:10.1016/j.jaapos.2018.09.007

Hu, B., & Knill, D. C. (2011). Binocular and monocular depth cues in online feedback control of 3D poinitng movement. *J Vis, 11*(7), 23, 21-13.

Imaizumi, S., Asai, T., Hiromitsu, K., & Imamizu, H. (2018). Voluntarily controlled but not merely observed visual feedback affects postural sway. *PeerJ, 6*, e4643. doi:10.7717/peerj.4643

Irving, E., L., Sivak, A., M., & Spafford, M., M. (2018). "I can see fine": patient knowledge of eye care. *Ophthalmic Physiol Opt, 38*, 422-431.

Jackson, S., Harrad, R. A., Morris, M., & Rumsey, N. (2006). The psychosocial benefits of corrective surgery for adults with strabismus. *Br J Ophthalmol, 90*(7), 883-888. doi:10.1136/bjo.2005.089516

Jackson, S., Morris, M., & Gleeson, K. (2013). The long-term psychosocial impact of corrective surgery for adults with strabismus. *Br J Ophthalmol, 97*(4), 419-422. doi:10.1136/bjophthalmol-2012-302983

Ji, N., Xu, M., Yu, H., Xu, J., & Yu, X. (2020). Factors influencing the self-reported sense of deviation in adults with successful surgical outcomes for strabismus. *BMC Ophthalmol, 20*(1), 9. doi:10.1186/s12886-019-1299-3

Johnson, R., Wynn, S., & Coffey, B. (2004). Influences of examiner position and effective prism power on nearpoint alternate cover test. *Optometry, 75*(8), 496-502. doi:10.1016/s1529-1839(04)70174-8

Jones, R. K., & Lee, D. N. (1981). Why two eyes are better than one: the two views of binocular vision. *J Exp Psychol Hum Percept Perform, 7*(1), 30-40.

Joosse, M. V., Simonsz, H. J., van Minderhout, E. M., Mulder, P. G., & de Jong, P. T. (1999). Quantitative visual fields under binocular viewing conditions in primary and consecutive divergent strabismus. *Graefes Arch Clin Exp Ophthalmol, 237*(7), 535-545. doi:10.1007/s004170050276

Joosse, M. V., Simonsz, H. J., Van Minderhout, H. M., De Jong, P. T. V. M., Noordzij, B., & Mulder, P. G. H. (1997). Quantitative perimetry under binocular viewing conditions in microstrabismus. *Vision Res, 37*(19), 2801-2812. doi:10.1016/S0042-6989(97)00067-9

Ju, Y.-Y., Liu, Y.-H., Cheng, C.-H., Lee, Y.-H., Chang, S.-T., Sun, C.-C., & Cheng, H.-Y. K. (2018). Effects of combat training on visuomotor perfromance in children aged 9-12 years - an eye-tracking study. *BMC Pediatrics, 18*(39), 1-9.

Jung, E. H., & Kim, S. J. (2018). Surgical results and factors affecting outcome in adult patients with sensory exotropia. *Eye, 32*(12), 1851-1857. doi:10.1038/s41433-018-0189-x

Kaban, T., Smith, K., & Orton, R., B. (1997). The Normal Field of Binocular Vision as Tested on the Goldmann Perimeter. *Am Orthopt J, 47*(1), 164-171.

Karimi, M., & Brazier, J. (2016). Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *PharmacoEconomics, 34*(7), 645-649. doi:10.1007/s40273-016-0389-9

Kattan, J. M., Velez, F. G., Demer, J. L., & Pineles, S. L. (2016). Relationship Between Binocular Summation and Stereoacuity After Strabismus Surgery. *Am J Ophthalmol, 165*, 29-32. doi:10.1016/j.ajo.2016.02.020

Kelly, K. R., Jost, R. M., De La Cruz, A., Dao, L., Beauchamp, C. L., Stager Jr, D., & Birch, E. E. (2017). Slow reading in children with anisometropic amblyopia is associated with fixation instability and increased saccades. *J AAPOS, 21*, 447-451.

Keskinbora, K. H., Gonen, T., & Horozoglu, F. (2011). Outcome of surgery in long-standing infantile esotropia with cross fixation. *J Pediatr Ophthalmol Strabismus, 48*(2), 77-83. doi:10.3928/01913913-20110303-03

Kim, M. J., Khwarg, S. I., Kim, S. J., & Chang, B. L. (2008). Results of re-operation on the deviated eye in patients with sensory heterotropia. *Korean J Ophthalmol, 22*(1), 32-36. doi:10.3341/kjo.2008.22.1.32

Kim, Y., Kim, C., Kim, S., Han, Y., Chung, I., Seo, S., . . . Yoo, J. (2016). Difference of Self-identity Levels between Strabismus Patients and Normal Controls. *Korean J Ophthalmol, 30*(6), 410-415. doi:10.3341/kjo.2016.30.6.410

Kishimoto, F., & Ohtsuki, H. (2012). Comparison of VF-14 scores among different ophthalmic surgical interventions. *Acta Med Okayama, 66*(2), 101-110. doi:10.18926/amo/48260

Kitaoji, H., & Toyama, K. (1987). Preservation of position and motion stereopsis in strabismic subjects. *Invest Ophthalmol Vis Sci, 28*(8), 1260-1267.

Koc, F., Erten, Y., & Yurdakul, N. S. (2013). Does restoration of binocular vision make any difference in the quality of life in adult strabismus. *Br J Ophthalmol, 97*(11), 1425-1430. doi:10.1136/bjophthalmol-2013-303704

Kothari, M., & Joshi, V. (2014). The perceived personality traits of adults with digitally induced large angle strabismus and the impact of its correction. *Indian J Ophthalmol, 62*(7), 773-776. doi:10.4103/0301-4738.138617

Kumaran, S. E., Khadka, J., Baker, R., & Pesudovs, K. (2018). Patient-reported outcome measures in amblyopia and strabismus: a systematic review. *Clin Exp Optom, 101*(4), 460-484. doi:10.1111/cxo.12553

Kurz, J., Hegele, M., Reiser, M., & Munzert, J. (2017). Impact of task difficulty on gaze behavior in a sequential object manipulation task. *Exp Brain Res, 235*(11), 3479-3486. doi:10.1007/s00221-017-5062-9

Kushner, B. J. (1994). Binocular field expansion in adults after surgery for esotropia. *Arch Ophthalmol, 112*(5), 639-643. doi:10.1001/archopht.1994.01090170083027

Kushner, B. J., & Morton, G. V. (1992). Postoperative Binocularlity in Adults with Longstanding Strabismus. *Ophthalmol, 99*, 316-319.

Kutschke, P. J., & Scott, W. E. (2004). Prism adaptation in visually mature patients with esotropia of childhood onset. *Ophthalmol, 111*(1), 177-179. doi:10.1016/j.ophtha.2003.04.010

Lafayette Instrument Company, I. (2014-2015). Grooved Pegboard User's Manual MAN192 Model 32025. In L. Instrument (Ed.), (Rel. 1.30.15 ed.).

Lafayette Instrument Company, I. (2015). Purdue Pegboard Test User Instructions. Model 32020A. In (Rel. 10.21.15 ed., pp. 1-28): Lafayette Instrument. P.O. Box 5729, Lafayette, IN 47903, USA.

LaRoche, R., & Von Noorden, G. K. (1982). Theoretical and practical evaluation of a simple stereotest. *Invest Ophthalmol Vis Sci, 22*, 137.

Lee, K., Junghans, B. M., Ryan, M., Khuu, S., & Suttle, C. M. (2014). Development of a novel appraoch to the assessment of eye-hand coordination. *J Neurosci Methods, 228*, 50-56.

Lee, S. S.-Y., Wood, J. M., & Black, A. A. (2020). Impact of glaucoma on executive function and visual search. *Ophthalmic Physiol Opt, 40*, 333– 342. doi:10.1111/opo.12679

Lee, T. E., Kim, S. H., & Cho, Y. A. (2013). Postoperative changes in spatial localization following exotropia surgery. *Curr Eye Res, 38*(1), 210-214. doi:10.3109/02713683.2012.713151

Lenoir, M., Musch, E., & La Grange, N. (1999). Ecological relevance of stereopsis in one-handed ball-catching. *Percept Mot Skills, 89*, 495-508.

Leske, D. A., Hatt, S. R., & Holmes, J. M. (2010). Test-retest reliability of health-related quality-of-life questionnaires in adults with strabismus. *Am J Ophthalmol, 149*(4), 672-676. doi:10.1016/j.ajo.2009.11.004

Leske, D. A., Hatt, S. R., Liebermann, L., & Holmes, J. M. (2012). Evaluation of the Adult Strabismus-20 (AS-20) questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci, 53*(6), 2630-2639. doi:10.1167/iovs.11-8308

Levi, D. M., Knill, D. C., & Bavelier, D. (2015). Stereopsis and amblyopia: A mini-review. *Vision Res, 114*, 17-30. doi:10.1016/j.visres.2015.01.002

Liebermann, L., Hatt, S. R., Leske, D. A., & Holmes, J. M. (2013). Adjustment versus no adjustment when using adjustable sutures in strabismus surgery. *J AAPOS, 17*(1), 38-42. doi:10.1016/j.jaapos.2012.10.017

Liebermann, L., Hatt, S. R., Leske, D. A., & Holmes, J. M. (2014). Improvement in specific function-related quality-of-life concerns after strabismus surgery in nondiplopic adults. *J AAPOS, 18*(2), 105-109. doi:10.1016/j.jaapos.2013.11.018

Lin, J., He, Q., & Xiao, H. (2015). Effect of Diplopia on Daily Life and Evaluation of Nursing Measures Following Strabismus Surgery. *Eye Sci, 30*(2), 77-80.

Lin, S., Congdon, N., Yam, J. C., Huang, Y., Qiu, K., Ma, D., . . . Zhang, M. (2014). Alcohol use and positive screening results for depression and anxiety are highly prevalent among Chinese children with strabismus. *Am J Ophthalmol, 157*(4), 894-900. doi:10.1016/j.ajo.2014.01.012

Lipton, J. R., & Willshaw, H. E. (1995). Prospective multicentre study of the accuracy of surgery for horizontal strabismus. *Br J Ophthalmol, 79*(1), 10-11. doi:10.1136/bjo.79.1.10

Llorens, R., Latorre, J., Noé, E., & Keshner, E. A. (2016). Posturography using the Wii Balance Board™: A feasibility study with healthy adults and adults post-stroke. *Gait & Posture, 43*, 228-232. doi:10.1016/j.gaitpost.2015.10.002

Loftus, A., Servos, P., Goodale, M. A., Mendarozqueta, N., & Mon-Williams, M. (2004). When two eyes are better than one in prehnsion: monocular viewing and end-point variance. *Exp Brain Res, 158*, 317-327.

Lu, T. C., Semsarian, C. R., White, A., Dunn, H., Angell, B., Ford, B., & Keay, L. (2019). Journey to glaucoma care - trusting but uncertain and uninformed: a qualitative study. *Clin Exp Optom*. doi:10.1111/cxo.12955

MacKenzie, K., Hancox, J., McBain, H., Ezra, D. G., Adams, G., & Newman, S. (2016). Psychosocial interventions for improving quality of life outcomes in adults undergoing strabismus surgery. *Cochrane Database Syst Rev*(5), CD010092. doi:10.1002/14651858.CD010092.pub4

Mangione, C. M., Lee, P. P., Gutierrez, P. R., & et al. (2001). Development of the 25-list-item national eye institute visual function questionnaire. *Arch Ophthalmol, 119*(7), 1050-1058. doi:10.1001/archopht.119.7.1050

Mangione, C. M., Lee, P. P., Pitts, J., & et al. (1998). Psychometric properties of the national eye institute visual function questionnaire (nei-vfq). *Arch Ophthalmol, 116*(11), 1496-1504. doi:10.1001/archopht.116.11.1496

Mann, V. A., Hein, A., & Diamond, R. (1979). Localization of targets by strabismic subjects: Contrasting patterns in constant and alternating suppressors. *Percept Psychophys, 25*(1), 29-34. doi:10.3758/BF03206106

Marsh, I. B. (2015). We need to pay heed to the psychosocial aspects of strabismus. *Eye, 29*, 238-240.

Mason, J. (2002). *Qualitative Researching*. London, UK.: SAGE Publications Ltd.

Matsuo, T., Narita, A., Senda, M., Hasebe, S., & Ohtsuki, H. (2006). Body sway increases immediately after strabismus surgery. *Acta Med Okayama, 60*(1), 13-24. doi:10.18926/amo/30754

Matsuo, T., Negayama, R., Sakata, H., & Hasebe, K. (2014). Correlation between Depth Perception by Three-Rods Test and Stereoacuity by Distance Randot Stereotest. *Strabismus, 22*(3), 133-137. doi:10.3109/09273972.2014.939766

Mazyn, L. I. N., Lenoir, M., Montagne, G., Delaey, C., & Savelsbergh, G. J. P. (2007). Stereo vision enhances the learning of a catching skill. *Exp Brain Res, 179*(4), 723-726. doi:10.1007/s00221-007-0957-5

Mazyn, L. I. N., Lenoir, M., Montagne, G., & Savelsbergh, G. J. P. (2004). The contribution of stereo vision to one-handed catching. *Exp Brain Res, 157*(3), 383-390. doi:10.1007/s00221-004-1926-x

McBain, H. B., Au, C. K., Hancox, J., MacKenzie, K. A., Ezra, D. G., Adams, G. G. W., & Newman, S. P. (2014a). The impact of strabismus on quality of life in adults with and without diplopia: a systematic review. *Surv Ophthalmol, 59*, 185-191.

McBain, H. B., MacKenzie, K., Hancox, J., Ezra, D. G., Adams, G. G., & Newman, S. P. (2016a). What do patients with strabismus expect post surgery? The development and validation of a questionnaire. *Br J Ophthalmol, 100*(3), 415-419. doi:10.1136/bjophthalmol-2015-307027

McBain, H. B., MacKenzie, K. A., Au, C., Hancox, J., Ezra, D. G., Adams, G. G., & Newman, S. P. (2014b). Factors associated with quality of life and mood in adults with strabismus. *Br J Ophthalmol, 98*(4), 550-555. doi:10.1136/bjophthalmol-2013-304220

McBain, H. B., MacKenzie, K. A., Hancox, J., Ezra, D. G., Adams, G. G., & Newman, S. P. (2016b). Does strabismus surgery improve quality and mood, and what factors influence this? *Eye, 30*(5), 656-667. doi:10.1038/eye.2016.70

McKeon, C., Wick, B., Aday, L. A., & Begley, C. (1997). A Case-Comparison of Intermittent Exotropia and Quality of Life Measurements. *Optom Vis Sci, 74*(2), 105-110.

Melmoth, D. R., Finlay, A. L., Morgan, M. J., & Grant, S. (2009). Grasping Deficits and Adaptations in Adults with Stereo Vision Losses. *Invest Ophthalmol Vis Sci, 50*(8), 3711-3720.

Melmoth, D. R., & Grant, S. (2006). Advantages of binocular vision for the control of reaching and grasping. *Exp Brain Res, 171*(3), 371-388. doi:10.1007/s00221-005-0273-x

Melmoth, D. R., Storoni, M., Todd, G., Finlay, A. L., & Grant, S. (2007). Dissociation between vergence and binocular disparity cues in the control of prehension. *Exp Brain Res, 183*, 283-298. doi:10.1007/s00221-007-1041-x

Menon, V., Saha, J., Tandon, R., Mehta, M., & Khokhar, S. (2002). Study of the psychosocial aspects of strabismus. *J Pediatr Ophthalmol Strabismus, 39*(4), 203-208.

Merrill, K., Satterfield, D., & O'Hara, M. (2010). Strabismus surgery on the elderly and the effects on disability. *J AAPOS, 14*(2), 196-198. doi:10.1016/j.jaapos.2009.12.282

Mets, M. B., Beauchamp, C., & Haldi, B. A. (2004). Binocularity following surgical correction of strabismus in adults. *J AAPOS, 8*(5), 435-438. doi:10.1016/j.jaapos.2004.07.003

Mohney, B. G., McKenzie, J. A., Capo, J. A., Nusz, K. J., Mrazek, D., & Diehl, N. N. (2008). Mental Illness in Young Adults Who Had Strabimus as Children. *Pediatrics, 122*(5), 1033-1038.

Mojon-Azzi, S. M., Kunz, A., & Mojon, D. S. (2011). The perception of strabismus by children and adults. *Graefes Arch Clin Exp Ophthalmol, 249*(5), 753-757. doi:10.1007/s00417-010-1555-y

Mojon-Azzi, S. M., & Mojon, D. S. (2009). Strabismus and employment: the opinion of headhunters. *Acta Ophthalmol, 87*(7), 784-788. doi:10.1111/j.1755-3768.2008.01352.x

Mojon-Azzi, S. M., Potnik, W., & Mojon, D. S. (2008). Opinions of dating agents about strabismic subjects' ability to find a partner. *Br J Ophthalmol, 92*(6), 765-769. doi:10.1136/bjo.2007.128884

Mojon, D. S. (2007). Comparison of a new, minimally invasive strabismus surgery technique with the usual limbal approach for rectus muscle recession and plication. *Br J Ophthalmol, 91*(1), 76-82. doi:10.1136/bjo.2006.105353

Mojon, D. S. (2008). Minimally invasive strabismus surgery for horizontal rectus muscle reoperations. *Br J Ophthalmol, 92*(12), 1648-1652. doi:10.1136/bjo.2008.145110

Mojon, D. S. (2009). A new stereotest for strabismologists: the covered two pencil test. *Eur J Ophthalmol, 19*(4), 530-534.

Mon-Williams, M., & Dijkerman, H. C. (1999). The use of vergence information in the programming of prehension. *Exp Brain Res, 128*(4), 578-582.

Moores, E., Frisby, J. P., Buckley, D., Reynolds, E., & Fawcett, A. (1998). Vergence control across saccades in dyslexic adults. *Ophthal Physiol Opt, 18*(5), 452-462.

Morris, R. J., Scott, W. E., & Dickey, C. F. (1993). Fusion after Surgical Alignment of Longstanding Strabismus in Adults. *Ophthalmol, 100*(1), 135-138. doi:10.1016/S0161-6420(93)31703-3

Murdoch, J., R., McGhee, C., N, J., & Glover, V. (1991). The Relationship Between Stereopsis and Fine Manual Dexterity: Pilot Study of a New Instrument. *Eye, 5*(5), 642-643.

Murray, A. D., Orpen, J., & Calcutt, C. (2007). Changes in the functional binocular status of older children and adults with previously untreated infantile esotropia following late surgical realignment. *J AAPOS, 11*(2), 125-130. doi:10.1016/j.jaapos.2006.10.016

Nelson, B. A., Gunton, K. B., Lasker, J. N., Nelson, L. B., & Drohan, L. A. (2008). The psychosocial aspects of strabismus in teenagers and adults and the impact of surgical correction. *J AAPOS, 12*(1), 72-76. doi:10.1016/j.jaapos.2007.08.006

NHS Improvement. (2019-2020). National tariff payment system. Retrieved from https://improvement.nhs.uk/resources/national-tariff/ (Accessed 7/11/20)

Niechwiej-Szwedo, E., Goltz, H. C., Chandrakumar, M., Hirji, Z. A., & Wong, A. M. (2010). Effects of anisometropic amblyopia on visuomotor behavior, I: saccadic eye movements. *Invest Ophthalmol Vis Sci, 51*(12), 6348-6354. doi:10.1167/iovs.10-5882

Niechwiej-Szwedo, E., Goltz, H. C., Colpa, L., Chandrakumar, M., & Wong, A. M. F. (2017). Effects of Reduced Acuity and Stereo Acuity on Saccades and Reaching Movements in Adults With Amblyopia and Strabismus. *Invest Ophthalmol Vis Sci, 58*, 914-921. doi:10.1167/iovs.16-20727

Niechwiej-Szwedo, E., Kennedy, S. A., Colpa, L., Chandrakumar, M., Goltz, H. C., & Wong, A. M. F. (2012). Effects of Induced Monocular Blur versus Anisometropia Amblyopia on Saccades, Reaching and Eye-Hand Coordination. *Invest Ophthalmol Vis Sci, 53*, 4354-4362. doi:10.1167/iovs.12-9855

O'Cathain, A., & Thomas, K. (2006). Combining Qualitative and Quantitative Methods. In C. Pope & N. Mays (Eds.), *Qualitative Research in Health Care* (Third Edition ed., pp. 102-111): Blackwell Publishing Ltd, Oxford, UK.

O'Connor, A. R., Birch, E. E., Anderson, S., & Draper, H. (2010a). The functional significance of stereopsis. *Invest Ophthalmol Vis Sci, 51*(4), 2019-2023. doi:10.1167/iovs.09-4434

O'Connor, A. R., Birch, E. E., Anderson, S., & Draper, H. (2010b). Relationship between binocular vision, visual acuity, and fine motor skills. *Optom Vis Sci, 87*(12), 942-947. doi:10.1097/OPX.0b013e3181fd132e

O'Connor, A., R., & Tidbury, L. P. (2018). Stereopsis: are we assessing it in enough depth? *Clin Exp Optom, 101*(4), 485-494.

Olitsky, S. E., Sudesh, S., Graziano, A., Hamblen, J., Brooks, S. E., & Shaha, S. H. (1999). The negative psychosocial impact of strabismus in adults. *J AAPOS, 3*(4), 209-211.

Otto, J. M., Bach, M., & Kommerell, G. (2010). Advantage of binocularity in the presence of external visual noise. *Graefe's Arch Clin Exp Ophthalmol, 248*(4), 535-541. doi:10.1007/s00417-010-1304-2

Owsley, C., McGwin, G., Antin, J. F., Wood, J. M., & Elgin, J. (2018). The Alabama VIP older driver study rationale and design: examining the relationship between vision impairment and driving using naturalistic driving techniques. *BMC Ophthalmol, 18*(1), 32. doi:10.1186/s12886-018-0686-5

Ozates, S., Ezerbolat Ozates, M., Can, C. U., Polat, S., Yasar, H. H., Taskale, B., & Gogus, A. K. (2019). Improvement in psychiatric symptoms after strabismus surgery in adolescent patients in long-term follow-up. *Br J Ophthalmol, 103*(7), 966-970. doi:10.1136/bjophthalmol-2018-312510

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffman, T. C., Mulrow, C. D., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*, n71. doi:doi: 10.1136/bmj.n71

Paysse, E. A., Steele, E. A., McCreery, K. M., Wilhelmus, K. R., & Coats, D. K. (2001). Age of the emergence of negative attitudes toward strabismus. *J AAPOS, 5*(6), 361-366. doi:10.1067/mpa.2001.119243

Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clin Vis Sci, 2*(3), 187-199.

Pesudovs, K. (2010). Item banking: a generational change in patient reported outcome measurement. *Optom Vis Sci, 87*(4), 285-293. doi:10.1097/OPX.0b013e3181d408d7.

Piano, M., E, F., & O'Connor, A., R. (2013). The effect of degrading binocular single vision on fine visuomotor skill task performance. *Invest Ophthalmol Vis Sci., 54*, 8204–8213. doi:10.1167/iovs.12-10934

Pineles, S. L., Demer, J. L., Isenberg, S. J., Birch, E. E., & Velez, F. G. (2015). Improvement in binocular summation after strabismus surgery. *JAMA Ophthalmol, 133*(3), 326-332. doi:10.1001/jamaophthalmol.2014.5265

Pineles, S. L., Lee, P. J., Velez, F., & Demer, J. (2014). Effects of visual noise on binocular summation in patients with strabismus without amblyopia. *J Pediatr Ophthalmol Strabismus, 51*(2), 100-104. doi:10.3928/01913913-20140205-02

Pineles, S. L., Repka, M. X., Yu, F., Lum, F., & Coleman, A. L. (2017). Evaluating the impact of strabismus surgery on the association of musculoskeletal injuries, fractures, and falls with disorders of binocular vision in medicare beneficiaries. *J AAPOS, 21*(3), 239-241. doi:10.1016/j.jaapos.2016.11.023

Pineles, S. L., Velez, F. G., Isenberg, S. J., Fenoglio, Z., Birch, E., Nusinowitz, S., & Demer, J. L. (2013). Functional burden of strabismus: decreased binocular summation and binocular inhibition. *JAMA Ophthalmol, 131*(11), 1413-1419. doi:10.1001/jamaophthalmol.2013.4484

Plainis, S., Petratou, D., Giannakopoulou, T., Atchison, D. A., & Tsilimbaris, M. K. (2011). Binocular Summation Improves Performance to Defocus-Induced Blur. *Invest Ophthalmol Vis Sci, 52*(5), 2784-2789. doi:10.1167/

psteph. (2017). Image 'London shopping England 2268971'. Retrieved from https://pixabay.com/photos/london-shopping-england-2268971/ (Accessed 15/10/18)

Rajavi, Z., Feizi, M., Mughadasifar, H., Yaseri, M., Haftabadi, N., & Sheibani, K. (2013). Surgical results of consecutive exotropia. *J Pediatr Ophthalmol Strabismus, 50*(5), 274-281. doi:10.3928/01913913-20130730-03

Raw, R. K., Wilkie, R. M., Culmer, P. R., & Mon-Williams, M. (2012). Reduced motor asymmetry in older adults when manually tracing paths. *Exp Brain Res, 217*(1), 35-41. doi:10.1007/s00221-011-2971-x

Raw, R. K., Wilkie, R. M., Mon-Williams, M., Ross, S. A., Deniz, K., Goddard, T., & Patankar, T. (2017). Kinematic measures provide useful information after intracranial aneurysm treatment. *J Rehabil Assist Technol Eng, 4*, 1-11. doi:10.1177/2055668317744999

Read, J. C. A., Begum, S. F., McDonald, A., & Trowbridge, J. (2013). The binocular advantage in visuomotor tasks involving tools. *i-Perception, 4*, 101-110.

Read, J. C. A., Godfrey, A., Bohr, I., Simonotto, J., Galna, B., & Smulders, T. V. (2016). Viewing 3D TV over two months produces no discernible effects on balance, coordination or eyesight,. *Ergonomics*. doi:10.1080/00140139.2015.1114682

Read, J. C. A., Simonotto, J., Bohr, I., Godfrey, A., Galna, B., Rochester, L., & Smulders , T. V. (2015). Balance and coordination after viewing stereoscopic 3D television. *R Soc Open Sci, 2*, 140522. doi:10.1098/rsos.140522

Repka, M. X. (1997). Strabismus Surgery Among Aged Medicare Beneficiaries. *J AAPOS, 1*, 231-234.

Repka, M. X., Lum, F., & Burugapalli, B. (2018). Strabismus, Strabismus Surgery, and Reoperation Rate in the United States: Analysis from the IRIS Registry. *Ophthalmol, 125*(10), 1646-1653. doi:10.1016/j.ophtha.2018.04.024

Repka, M. X., Yu, F., & Coleman, A. (2012). Strabismus among aged fee-for-service Medicare beneficiaries. *J AAPOS, 16*(6), 495-500. doi:10.1016/j.jaapos.2012.07.010

Repka, M. X., Yu, F., Lum, F., & Coleman, A. L. (2013). Strabismus among aged Medicare beneficiaries: impact of health status and region. *J AAPOS, 17*(4), 391-394. doi:10.1016/j.jaapos.2013.05.005

Ribeiro G de, B., Bach, A. G., Faria, C. M., Anastásia, S., & Almeida, H. C. (2014). Quality of life of patients with strabismus. *Arq Bras Oftalmol, 77*(2), 110-113. doi:10.5935/0004-2749.20140027

Ribeiro, G. d. B., Bach, A. G. Z., Faria, C. M., Anastásia, S., & Almeida, H. C. d. (2014). Quality of life of patients with strabismus. *Arq Bras Oftalmol, 77*, 110-113. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492014000200110&nrm=iso

Ritchie, A., Colapinto, P., & Jain, S. (2013). The psychological impact of strabismus: does the angle really matter? *Strabismus, 21*(4), 203-208. doi:10.3109/09273972.2013.833952

Ritchie, A. E., & Ali, N. (2019). The incidence and clinical outcome of complications in 4,000 consecutive strabismus operations. *J AAPOS, 23*(3), 140.e141-140.e146. doi:10.1016/j.jaapos.2018.08.015

Royal College of Ophthalmologists. (2016). *Strabismus (squint) surgical intervention is not a cosmetic procedure*. Retrieved from https://www.rcophth.ac.uk/2016/09/squint-surgical-intervention-is-not-a-cosmetic-procedure/ (Accessed 7/11/20)

Ryu, W. Y., Sohn, E. J., & Ahn, H. B. (2015). Differences between patient expectations and actual strabismus surgery experience in adults. *Can J Ophthalmol, 50*(1), 61-67. doi:https://doi.org/10.1016/j.jcjo.2014.09.009

Sabri, K., Knapp, C. M., Thompson, J. R., & Gottlob, I. (2006). The VF-14 and Psychological Impact of Amblyopia and Strabismus. *Invest Ophthalmol Vis Sci, 47*(10), 4386-4392. doi:10.1167/iovs.05-1365

Sah, S. P., Sharma, I. P., Chaudhry, M., & Saikia, M. (2017). Health-Related Quality of Life (HRQoL) in Young Adults with Strabismus in India. *J Clin Diagn Res, 11*(2), NC01-NC04. doi:10.7860/JCDR/2017/24541.9389

Saladin, J. J. (2005). Stereopsis From a Performance Perspective. *Optom Vis Sci, 82*(3), 186-205.

Sandercoe, T., Beukes, S., & Martin, F. (2014). Adults with strabismus seek surgery for pyschosocial benefits. *Taiwan J Ophthalmol, 4*(1), 17-20. doi:10.1016/j.tjo.2013.10.004

Satterfield, D., Keltner, J. L., & Morrison, T. L. (1993). Psychosocial Aspects of Strabismus Study. *Arch Ophthalmol, 111*, 1100-1105.

Schiller, P., H., Kendall, G., L. , Kwak, M., C. , & Slocum, W., M. (2012). Depth Perception, Binocular Integration and Hand-Eye Coordination in Intact and Stereo Impaired Human Subjects. *J Clin Exp Ophthalmol, 3*, 210. doi:10.4172/2155-9570.1000210

Scott, I. U., Feuer, W. J., & Jacko, J. A. (2002). Impact of visual function on computer task accuracy and reaction time in a cohort of patients with age-related macular degeneration. *Am J Ophthalmol, 133*(3), 350-357. doi:10.1016/S0002-9394(01)01406-4

Şener, E. C., Mocan, M. C., Gedik, Ş., Ergin, A., & Sanaç, A. Ş. (2002). The reliability of grading the fixation preference test for the assessment of interocular visual acuity differences in patients with strabismus. *J AAPOS, 6*(3), 191-194. doi:10.1067/mpa.2002.122364

Senthil, M., P., Khadka, J., Gilhotra, J., S., Simon, S., Fenwick, E., K., Lamouteux, E., & Pesudovs, K. (2019). Understanding quality of life impact in people with retinal vein occlusion: a qualitative inquiry. *Clin Exp Optom, 102*, 406-411.

Serafino, M., Granet, D. B., Kushner, B. J., Dagi, L. R., Kekunnaya, R., & Nucci, P. (2019). Use of the Delphi process for defining successful outcomes for strabismus surgery. *J AAPOS, 23*(6), 309-312. doi:10.1016/j.jaapos.2019.07.006

Servos, P., Goodale, M. A., & Jakobson, L. S. (1992). The role of binocular vision in prehension: a kinematic analysis. *Vision Res, 32*(8), 1513-1521.

Severini, G., Straudi, S., Pavarelli, C., Da Roit, M., Martinuzzi, C., Di Marco Pizzongolo, L., & Basaglia, N. (2017). Use of Nintendo Wii Balance Board for posturographic analysis of Multiple Sclerosis patients with minimal balance impairment. *J Neuroeng Rehabil, 14*(1), 19. doi:10.1186/s12984-017-0230-5

Shaheen, A. F., Sourlas, A., Horton, K., McLean, C., Ewins, D., Gould, D., & Ghoussaynia, S. (2018). Effects of lighting illuminance levels on stair negotiation performance in individuals with visual impairment. *J Electromyogr Kinesiol, 39*, 8-15. doi:10.1016/j.jelekin.2018.01.004

Sheedy, J. E., Bailey, I. L., Buri, M., & Bass, E. (1986). Binocular vs. monocular task performance. *Am J Optom Physiol Opt, 63*(10), 839-846.

Sheedy, J. E., Harris, M. G., Bronge, M. R., Joe, S. M., & Mook, M. A. (1991). Task and visual performance with concentric bifocal contact lenses. *Optom Vis Sci, 68*(7), 537-541.

Sheedy, J. E., Harris, M. G., Poon, L., & Sakuda, T. (1992). Task and visual performance with contact lenses and spectacles. *Optom Vis Sci, 69*(5), 337-341.

Shire, K., A. , Hill, L., J, B. , Snapp-Childs, W., Bingham, G., P. , Kountouriotis, G., K. , Barber, S., & Mon-Williams, M. (2016). Robot Guided 'Pen Skill' Training in Children with Motor Difficulties. *PLoS One, 11*(3), e0151354. doi:10.1371/journal.pone.0151354

Sight Loss and Vision Priority Setting Partnership. (2013). Setting Priorities for Eye Research. Final Report. In (pp. 1-47). jla.nihr.ac.uk/priority-setting-partnerships/sight-loss-and-vision/ (accessed 14/9/21).

Silverman, D. (2013). *Doing Qualitative Research* (K. Metzler Ed. Fourth ed.). London: SAGE Publications Ltd.

Sim, J., & Lewis, M. (2012). Sample size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. *J Clin Epidemiol, 65*, 301-308.

Sim, P. Y., Cleland, C., Dominic, J., & Jain, S. (2018). Investigation of factors associated with the success of adult strabismus surgery from the patient's perspective. *J AAPOS, 22*(4), 266-271. doi:10.1016/j.jaapos.2018.03.006

Snapp-Childs, W., Fath, A. J., Watson, C. A., Flatters, I., Mon-Williams, M., & Bingham, G. P. (2015). Training to improve manual control in 7-8 and 10-12 year old children: Training

eliminates performance differences between ages. *Hum Mov Sci, 43*, 90-99. doi:10.1016/j.humov.2015.07.006

SPSS Inc., C., Ill. USA. (2017). SPSS Statistics for Windows, version 25.0 (Version Version 25).

SR Research. (2020). Eyelink 1000+ Technical Specifications. In. https://www.sr-research.com/eyelink-1000-plus/ (accessed 03/08/20).

Stayte, M., Reeves, B., & Wortham, C. (1993). Ocular and vision defects in preschool children. *Br J Ophthalmol, 77*(4), 228-232.

Steinberg, E. P., Tielsch, J. M., Schein, O. D., Javitt, J. C., Sharkey, P., Cassard, S. D., . . . Sommer, A. (1994). The VF-14: An Index of Functional Impairment in Patients With Cataract. *Arch Ophthalmol, 112*(5), 630-638. doi:10.1001/archopht.1994.01090170074026

Suttle, C. M., Melmoth, D. R., Finlay, A. L., Sloper, J. J., & Grant, S. (2011). Eye-hand coordination skills in children with and without amblyopia. *Invest Ophthalmol Vis Sci, 52*(3), 1851-1864. doi:10.1167/iovs.10-6341

Swain, T. A., McGwin, G., Wood, J. M., & Owsley, C. (2021). Motion perception as a risk factor for motor vehicle collision involvement in drivers ≥ 70 years. *Accid Anal Prev, 151*, 105956. doi:10.1016/j.aap.2020.105956

Tandon, A. K., Velez, F. G., Isenberg, S. J., Demer, J. L., & Pineles, S. L. (2014). Binocular inhibition in strabismic patients is associated with diminished quality of life. *J AAPOS, 18*(5), 423-426. doi:10.1016/j.jaapos.2014.05.009

Taroyan, N. A., Buckley, D., Porrill, J., & Frisby, J. P. (2000). Exploring sequential stereopsis for co-planarity tasks. *Vision Res, 40*(24), 3373-3390. doi:10.1016/S0042-6989(00)00174-7

Taylor, D. J., Smith, N. D., & Crabb, D. P. (2017). Searching for Objects in Everyday Scenes: Measuring Performance in People With Dry Age-Related Macular Degeneration. *Invest Ophthalmol Vis Sci, 58*(3), 1887-1892. doi:10.1167/iovs.16-21122

Thayaparan, K., Crossland, M. D., & Rubin, G. S. (2007). Clinical assessment of two new contrast sensitivity charts. *Br J Ophthalmol, 91*(6), 749-752. doi:10.1136/bjo.2006.109280

Thibaut, M., Boucart, M., & Tran, T. H. C. (2019). Object search in neovascular age-related macular degeneration: the crowding effect. *Clin Exp Optom*. doi:10.1111/cxo.12982

Thompson, B., Richard, A., Churan, J., Hess, R. F., Aaen-Stockdale, C., & Pack, C. C. (2011). Impaired spatial and binocular summation for motion direction discrimination in strabismic amblyopia. *Vision Res, 51*(6), 577-584. doi:10.1016/j.visres.2011.02.001

Tidbury, L. P., Black, R. H., & O'Connor, A. R. (2014). Perceiving 3D in the absence of measureable stereo-acuity. *Brit Ir Orthopt J, 11*, 34-38.

Turan, E. K., Şekeroğlu, T. H., Şener, E. C., & Sanaç, A. (2015). Effect of Visual Acuity on the Surgical Outcomes of Secondary Sensory Strabismus. *Turk J Ophthalmol, 45*(6), 254-258. doi:10.4274/tjo.67878

Umazume, F., Ohtsuki, H., & Hasebe, S. (1997). Predictors of postoperative binocularity in adult strabismus. *Jpn J Ophthalmol, 41*(6), 414-421.

Unknown photographer. (2017). London street scene 773176. Retrieved from https://pxhere.com/en/photo/773176 (Accessed 15/10/18)

van de Graaf, E. S., Borsboom, G., van der Sterre, G. W., Felius, J., Simonsz, H. J., & Kelderman, H. (2017). Differences in quality-of-life dimensions of Adult Strabismus Quality of Life and Amblyopia & Strabismus Questionnaires. *Graefes Arch Clin Exp Ophthalmol*. doi:10.1007/s00417-017-3694-x

van de Graaf, E. S., van der Sterre, G. W., Polling, J. R., van Kempen, H., Simonsz, B., & Simonsz, H. J. (2004). Amblyopia & Strabismus Questionnaire: design and initial validation. *Strabismus, 12*(3), 181-193. doi:10.1080/09273970490491196

van de Graaf, E. S., van Kempen-du Saar, H., Looman, C. W., & Simonsz, H. J. (2010). Utility analysis of disability caused by amblyopia and/or strabismus in a population-based, historic cohort. *Graefes Arch Clin Exp Ophthalmol, 248*(12), 1803-1807. doi:10.1007/s00417-010-1457-z

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behav Res Methods, 49*(2), 653-673. doi:10.3758/s13428-016-0721-5

Vedamurthy, I., Suttle, C. M., Alexander, J., & Asper, L. J. (2007). Interocular interactions during acuity measurement in children and adults, and in adults with amblyopia. *Vision Res, 47*(2), 179-188. doi:10.1016/j.visres.2006.08.017

Wakayama, A., Matsumoto, C., Ohmure, K., Inase, M., & Shimomura, Y. (2012). Influence of Background Complexity on Visual Sensitivity and Binocular Summation Using Patterns with and without Noise. *Invest Ophthalmol Vis Sci, 53*(1), 387-393. doi:10.1167/iovs.11-8022

Wang, J. Y., Leske, D. A., Hatt, S. R., & Holmes, J. M. (2019). Diplopia after strabismus surgery for adults with nondiplopic childhood-onset strabismus. *J AAPOS, 23*(6), 313.e311-313.e315. doi:10.1016/j.jaapos.2019.07.005

Wang, L., & Nelson, L. B. (2011). Outcome study of graded unilateral medial rectus recession for small to moderate angle esotropia. *J Pediatr Ophthalmol Strabismus, 48*(1), 20-24. doi:10.3928/01913913-20100818-05

Wang, Z., Ren, H., Frey, R., Liu, Y., Raphael, D., Bian, W., & Wang, X. (2014). Comparison of the Adult Strabismus Quality of Life Questionnaire (AS-20) with the Amblyopia and Strabismus Questionnaire (ASQE) among adults with strabismus who seek medical care in China. *BMC Ophthalmol, 14*, 139. doi:10.1186/1471-2415-14-139

Wang, Z., Zhou, J., Xu, Y., Yin, H., She, X., Bian, W., & Wang, X. (2018). Development of a conceptual model regarding quality of life in Chinese adult patients with strabismus: a mixed method. *Health Qual Life Outcomes, 16*(1), 171. doi:10.1186/s12955-018-0991-y

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Med Care, 30*(6), 473-483. Retrieved from www.jstor.org/stable/3765916

Watt, S. J., & Bradshaw, M. F. (2000). Binocular cues are important in controlling the grasp but not the reach in natural prehension movements. *Neuropsychologia, 38*, 1473-1481.

West, S. K., Rubin, G. S., Broman, A. T., Munoz, B., Bandeen-Roche, K., & Turano, K. (2002). How does visual impairment affect performance on tasks of everyday life? The see project. *Arch Ophthalmol, 120*(6), 774-780. doi:10.1001/archopht.120.6.774

Whoqol Group, T. (1998). The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties. *Social Science & Medicine, 46*(12), 1569-1585. doi:10.1016/S0277-9536(98)00009-4

Wortham, E. V., & Greenwald, M. J. (1989). Expanded Binocular Peripheral Visual Fields Following Surgery for Esotropia. *J Pediatr Ophthalmol Strabismus, 26*(3), 109-112.

Wright, K. W., Edelman, P. M., Walonker, F., & Yiu, S. (1986). Reliability of Fixation Preference Testing in Diagnosing Amblyopia. *Arch Ophthalmol, 104*(4), 549-553. doi:10.1001/archopht.1986.01050160105023

Xu, J., Yu, X., Huang, Y., Chen, J., Yu, H., Wang, Y., & Zhang, F. (2012). The psychosocial effects of strabismus before and after surgical correction in chinese adolescents and adults. *J Pediatr Ophthalmol Strabismus, 49*(3), 170-175. doi:10.3928/01913913-20110920-02

Xu, M., Yu, H., Chen, Y., Xu, J., Zheng, J., & Yu, X. (2016). Long-Term Quality of Life in Adult Patients with Strabismus after Corrective Surgery Compared to the General Population. *PLoS One, 11*(11), e0166418. doi:10.1371/journal.pone.0166418

Ziebland, S., & McPherson, A. (2006). Making sense of qualitative data analysis: an introduction will illustrations from DIPEx (personal experiences of health and illness). *Med Educ, 40*, 405-414.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370. doi:10.1111/j.1600-0447.1983.tb09716.x

Zipori, A. B., Colpa, L., Wong, A. M. F., Cushing, S. L., & Gordon, K. A. (2018). Postural stability and visual impairment: Assessing balance in children with strabismus and amblyopia. *PLoS One, 13*(10), e0205857. doi:10.1371/journal.pone.0205857

Zlatkova, M. B., Anderson, R. S., & Ennis, F. A. (2001). Binocular summation for grating detection and resolution in foveal and peripheral vision. *Vision Res, 41*(24), 3093-3100. doi:10.1016/S0042-6989(01)00191-2

Appendix A. Strabismus surgery procedures and costs

Table A-1. Number of surgical procedures on the extraocular muscles in the NHS in England

1 year period (April-March)	Number of procedures performed on the extraocular muscles	Procedures in children 0-17 years	Procedures in adults 18-90+ (% of all cases)
2019-2020	11,214	6,281	4,933 (44%)
2018-2019	11,784	6,605	5,179 (44%)
2017-2018	11,810	6,641	5,169 (44%)
2016-2017	11,987	6,666	5,321 (44%)
2015-2016	11,954	6,725	5,229 (44%)
2014-2015	11,904	6,614	5,290 (44%)

(Health and Social Care Information Centre) Information accessed 07/11/2020

Table A-2. NHS tariff costs relating to strabismus surgery

Description of procedure	£
Complex Ocular Motility Procedures	1,420
Very Major Ocular Motility Procedures, 19 years and over	1,287
Very Major Ocular Motility Procedures, 18 years and under	1,414
Major Ocular Motility Procedures, 19 years and over	1,277
Major Ocular Motility Procedures, between 4 and 18 years	1,391
Major Ocular Motility Procedures, 3 years and under	1,342
Intermediate Ocular Motility Procedures, 19 years and over	1,106
Intermediate Ocular Motility Procedures, 18 years and under	1,270
Minor Ocular Motility Procedures, 19 years and over	680
Minor Ocular Motility Procedures, 18 years and under	911

Adult strabismus surgery and ocular motility procedures highlighted in yellow
(NHS Improvement, 2019-2020) Information accessed 07/11/2020

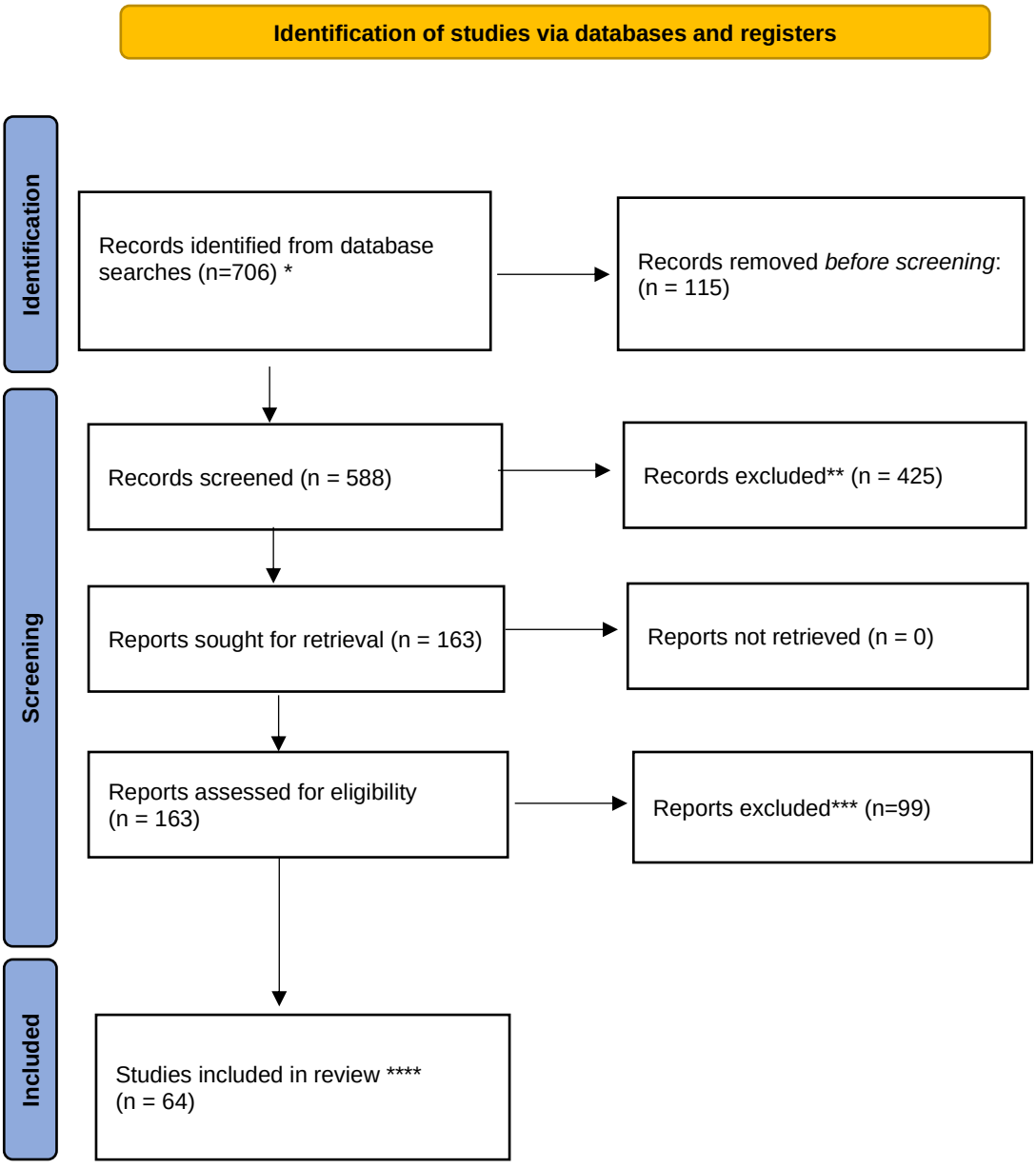
Appendix B. Systematic literature search

Appendix B-1. Literature search terms

Table B-1. Table of literature search terms used		
Terms	Boolean operator	Filters used
Strabismus Adult Surgery Outcomes	AND	English Humans All adult age categories
Thyroid Graves Myasthenia Nerve palsy Myopia Fracture Intermittent Duane syndrome	NOT	
Additional search performed using the MeSH terms: Strabismus AND Surgery including the term AND psychosocial (all fields)		
Additional search performed using the terms: outcome AND functional AND eye alignment AND squint		

Appendix B-2. PRISMA flow diagram of literature search outputs.

Flow chart template from Page et al. (2021)



*Databases searched: PubMed, Scopus, Cochrane Library, NICE, PsycINFO, Web of Science, Google Scholar, the British and Irish Orthoptic Journal online and an EndNote database of non- or pre-Medline indexed sources (American Orthoptic Journal, Australian Orthoptic Journal, British Orthoptic Journal, Strabismus, Binocular Vision, Journal of AAPOS, and the Transactions of the International Orthoptic Congress, the International Strabismological Association, and the European Strabismological Association).

** No automated tools were used, all records were excluded by GA.

** Exclusions due to:

- strabismus surgery for visual benefit only (to gain BSV or to eliminate diplopia), investigative outcomes in patients with potential BSV (for example prism adaptation to restore BSV prior to strabismus surgery)
- strabismus secondary to or associated with other aetiologies such as neurogenic palsy, mechanical condition (for example Duane syndrome), high myopia, retinal detachment, orbital fractures, congenital fibrosis of the extraocular muscles, age related distance ET (with diplopia)
- other strabismus diagnoses reported only (for example acute acquired concomitant esotropia, DVD, double elevator palsy)
- strabismus surgery outcomes in co-existing ocular pathology (for example glaucoma)
- strabismus surgery anaesthetic techniques or surgical techniques (without strabismus outcome data)
- strabismus surgery techniques and outcomes following specific vertical muscle procedures for a vertical or torsional deviation (for example Harada-Ito procedure)
- intermittent strabismus or heterophoria only
- paediatric patients only (with the following exceptions: childhood strabismus that had recurred in adulthood and childhood onset strabismus that had received the primary surgical treatment in adulthood)
- other surgical outcomes (for example refractive surgery outcomes performed in patients with strabismus)
- treatments for diplopia (with the exception of diplopia resulting from psychosocial strabismus surgery, which was included)
- slipped extraocular muscles during surgery (for example, description of surgical technique but no reported strabismus outcome)
- outcomes from Botulinum Toxin (BT) injections
- Poster abstracts
- Review papers reporting no original data
- Editorial articles

*** Exclusions due to:

- strabismus surgery outcomes reported in a heterogeneous cohort and not possible to extract outcomes in those undergoing strabismus surgeries for psychosocial reasons only

- insufficient evidence reported to be able to determine postoperative outcomes of strabismus surgery in those undergoing strabismus surgery for psychosocial reasons

Appendix B-3. Table showing the studies included in the literature review

Author	Study purpose	Patients	Outcome criteria	Time postoperative outcome judged	Study design
(Adams et al., 2016)	Investigating psychological issues in patients before and after strabismus surgery	All strabismus patients (n=220) Age 17-88 No diplopia (n=96)	Clinical assessment of success, partial success or failure using criteria 1=largest angle of deviation <12PD (for ET, XT and HT), <20PD HoT; 2=no (or rare) diplopia or visual confusion in primary and reading position; and 3=no prisms or Bangerter foil occlusion Success = 3/3 criteria met Partial success = 1 or 2/3 criteria met Failure = 0/3 criteria met Psychological questionnaires (QoL: Adult Strabismus quality of life questionnaire AS-20, Mood: Hospital Anxiety & Depression Scale (HADS), Appearance related social anxiety and social avoidance: The Derriford Appearance Scale (DAS24), Beliefs about strabismus: Revised Illness Perception Questionnaire (IPQ-R), Beliefs about strabismus surgery: Treatment Representations Inventory (TRI), Fear of negative evaluation: Fear of Negative Evaluation (FNE), Perceived visibility: 7-point Likert scale from 1 (not at all visible) to 7 (extremely visible), Importance of appearance: The Centre of Appearance Research Salience Scale (CARSAL), Perception of their appearance: The Centre of Appearance Research Valence Scale (CARVAL), Satisfaction with social support: Multidimensional Scale of Perceived Social Support (MSPSS), Expectations about the outcome of surgery: designed by psychology team (ESSQ) Reasons for having surgery: designed by psychology team (RSSQ), Satisfaction with surgery: designed by psychology team	3 months clinical 3 and 6 months psychological	Prospective
(Akbari et al., 2015)	Persian version of AS-20 pre and postoperatively	All types of strabismus N=112 Age 15-43 years	AS-20 (Persian version) VFQ-25 (Persian version) Diplopia (yes / no) PCT <10PD and ≥10PD	3 months	Prospective

(Alam et al., 2014)	Investigating AS-20 outcomes in those considered surgical success	Concomitant manifest strabismus >15PD (preop) successfully aligned within 10PD orthotropia N=30 Age 11-34 years	AS-20	6 weeks 3 months	Prospective
(Aletaha et al., 2016)	Comparison of surgical techniques	Horizontal strabismus N=54 Age 2-50 years	PCT Number of reoperations	3 months	Prospective
(Alkharashi & Hunter, 2017)	Comparison of surgical techniques	All rectus strengthening procedures (resection or plication) N=72 Age 1-86 years	Success = distance PCT ≤10 PD horizontal deviation and ≤6 PD vertical deviation Reoperation rate Postoperative alignment drift (change from immediate postoperative measurement to final visit measurement)	6-12 week	Retrospective
(Al-Wadaani, 2017)	Retrospective review of all strabismus surgery	All non-adjustable strabismus surgery N=96 Age 16-61 years	Improvement in deviation postoperatively	6-47 months	Retrospective
(Ball et al., 1993)	Case series of unexpected stereopsis postoperatively	N=8	BSV tests		Retrospective
(Bayramlar & Gunduz, 2006)	Review of long term outcome of strabismus surgery in dense	N=33 Age 8-61 years	Krimsky measurement of deviation Success ±12 PD deviation	2 months and 24-108 months	Retrospective

	amblyopes (6/60 or worse)				
(Beauchamp et al., 2003)	Review of strabismus outcomes (all patients combined)	All patients who had strabismus surgery (6 centres) N=299 Age 16 years +	Success alignment = ≤8 PD horizontal deviation and ≤2 PD vertical deviation Success motor = ≤ +1 o/a Success sensory = no diplopia	1 day – 19 months	Multicentre retrospective
(Berland et al., 1998)	Patients undergoing 8-9mm bilateral LR recession for XT	N=30	Abduction limitation Reoperation rate	3-30 months	Retrospective
(Biglan et al., 1994)	Comparison of surgical procedures	All strabismus patients (all aetiologies) N=24 adjustable N=113 nonadjustable Mean age 43 and 42	Success = ±8 PD horizontal deviation and ±4 PD vertical deviation % success BSV Correction of diplopia	1 week and 6 weeks	Retrospective
(Bucci et al., 2009)	Horizontal saccades and vergence pre and postoperatively	With and without BSV N=9 Age 8-20 years	PCT BSV Saccades (measured onset: time to reach 5% of peak velocity, offset: time when velocity reduced to <10 degrees/sec, gain, mean velocity) Vergence (convergence and divergence) (measured onset: time when velocity reached > 5 degrees/sec, offset: time when velocity reduced to 5 degrees/sec, gain, mean velocity). Saccades combined with vergence	2 weeks – 2 months and 3 -10 months	Prospective
(Burke et al., 1997)	Psychosocial implications of strabismus and surgery	All had surgery for alignment N=31 Age 18-68 years	Self-reporting repertory grid – self rating psychosocial issues (pre op and post op) PCT	3 months	Prospective

(Chang et al., 2017)	Binocular summation in strabismic amblyopia and effect of surgery	N=15 strabismic amblyopia & Sx N=30 normal N=30 strabismus but no amblyopia	VA at 100%, 2.5% and 1.25% contrast (BEO & monocularly) Calculation of BiS Stereopsis PCT	6-10 weeks	Prospective
(Cifuentes et al., 2018)	Outcomes after 3 muscle surgery for large angle horizontal deviations	Consecutive patients having 3 muscle surgery for large angle horizontal strabismus patients N=28 Age 1 – 79 years	Motor alignment success criteria: Dist = Primary position 10PD residual deviation – 4PD consecutive deviation and no induced lateral incomitance 5PD between lateral gazes Nr = Primary position 10PD residual deviation – 4PD consecutive deviation Sensory success: improvement in stereopsis of 2 octaves Overcorrection >4PD consecutive deviation Dist & Near (primary position) Undercorrection >10PD deviation Dist & Nr (primary position)	6 weeks – 57 months	Retrospective
(Currie et al., 2003)	Outcomes after surgery for large angle XT	Consecutive patients having surgery for large angle XT N=26 Age 14-68 years	PCT Success criteria Dist = ≤10 PD heterotropia or phoria BSV Subjective question – Happy? Yes / No	8-12 months 18-36 months	Retrospective
(Dadeya et al., 2002)	Use of a drug during surgery to reduce restrictions postoperatively	Strabismus patients having a second surgery, +ve FDT but ≤25 PD N=20 Age 6-25 years	PCT FDT score Success criteria Satisfactory = ± 5 PD of orthophoria Undercorrection Overcorrection	1, 4 and 8 weeks then monthly for 12 months Outcome at 12 months	Prospective RCT
(Dawson et al., 2013)	Outcomes of strabismus treatment with poor VA (6/24 – PL)	Strabismus treatment outcomes in patients with reduced VA	PCT Comments documented in clinical notes about patient satisfaction postoperatively	2 weeks	Retrospective

		BT n=11 (n=2 then Sx) Sx (n=8 total) N=17 Age 19-74 years			
(Dotan et al., 2014)	Strabismus surgery in patients with unilateral vision loss and horizontal strabismus	Horizontal strabismus and unilateral VA in worst eye 1.0 or worse, VA in better seeing eye 0.3 or better N=21 Age 3 – 64 years	PCT Success ≤10 PD horizontal deviation and 1 surgical procedure was required Not success if >10PD or if >1 surgical procedure required	6-60 months	Retrospective
(Eino & Kraft, 1997)	Adjustable surgery for horizontal deviation	Compared predetermined target angle (after adjustment) to deviation at 6-8-months N=109 Age 15-72 years	PCT Drift from final alignment to 6-8 month measurement (in PCT and direction) Success if <10PD	Final alignment after adjustment 1-2 weeks 6-8 weeks 6-8 months	Retrospective
(Elkamshoushy & Langue, 2019)	biLR recession for recurrent XT (prev biMR resect)	Previous biMR resection for XT, but recurrent XT N=15 Age 20-31 years	PCT OM limitation of ABDuction Success 8PD ET – 10PD XT	6 months	Retrospective
(Estes et al., 2020)	Strabismus surgery, social anxiety and self consciousness	N=95 >18 years old	Questionnaire to evaluate self-consciousness (private and public) and social anxiety (self-consciousness survey instrument) Pre op and post op	6 months	Prospective

(Faridi et al., 2007)	All surgery for primary XT, no previous surgery	Intermittent or constant XT N=124 Mdn age at surgery 13 years (IQR 6-34 years)	Good motor outcome = ± 10PD orthotropia (SPCT) BSV	1-79 months	Retrospective
(Fatima et al., 2009)	Report postoperative BSV when none predicted preoperatively	Constant strabismus with no predicted BSV (free space with prisms) N=15 Age 12-40 years	BSV Success = ≤10 PD horizontal deviation and ≤4 PD vertical deviation	6 weeks	Retrospective
(Felius et al., 2001)	Re-recession of MR for recurrent ET	N=115 Age 11 months – 77 years	PCT Success ET ≤10 PD or XT ≤8 PD OM on versions (underaction of MR)	4 weeks – 8 months Long term follow up 8-120 months	Retrospective
(Frangouli & Adams, 2013)	Amniotic membrane in complex repeat strabismus surgery	Strabismus surgery complicated by fibrosis, range of aetiology N=8 Age 10-70 years	PCT Objective improvement Subjective improvement in patient symptoms (mainly relating to diplopia, but also includes report of binocular field of vision) Need for further interventions	9-24 months	Retrospective
(Ganesh et al., 2011)	Long term follow up of patients who had surgery for childhood ET	Surgery for ET until aligned to 0-10PD ET. Review 32-44 years later N=85	Initial surgery success = 0-10PD ET Incidence of consecutive XT = ≥10PD XT Near and Dist Reoperations OM restriction of ADDuction BSV	32-44 years	Prospective long term follow up study

		Age 2-24 at surgery			
(Ghiasi et al., 2013)	Psychosocial improvement after strabismus surgery	N=124 Age 15 years+ (71% no diplopia)	Used questionnaires from (Nelson et al, 2008) translated (Iranian population)	3 months	Prospective
(Gigante et al., 2018)	10 year follow up after monocular surgery for large angle ET	Range of aetiologies of large angle ET N=36 Age at surgery 4-58 years	PCT Good ≤15PD Fair 16-20PD Poor >20PD Rate of consecutive XT	6 months 10 years	Prospective long term follow up
(Glasman et al., 2013)	QoL following strabismus surgery – all patients with complete data	Horizontal and vertical deviations N=86 Age 17-76 years	PCT AS-20 (total, function subscale and psychosocial subscale)	12 days – 1 year	Prospective
(Gusek-Schneider & Boss, 2010)	Secondary sensory strabismus surgery outcomes	All patients having surgery for secondary sensory strabismus N=26 Age 3-45 years	PCT Dist VA BSV Diplopia yes / no Patient satisfaction with surgery yes / no	3 months Last follow up (1 year 8 m – 13 years 3 m)	Retrospective
(Hatt et al., 2010a)	HRQoL questionnaires in strabismus surgery	All strabismus, with diplopia (n=80) and without diplopia (n=26) N=106 Age 18-84 years	AS-20 VFQ-25 PCT (SPCT) Success criteria 1. no diplopia / visual confusion in primary position or for reading 2. <10PD heterotropia primary position Near or Dist 3. No prism / Bangerter foil / occlusion 4. No symptoms relating to misalignment or strabismus surgery Partial success	4-13 weeks	Prospective

			1. No diplopia / visual confusion in primary position or reading 2. <20PD heterotropia in primary position at Dist and Near 3. No prism / Bangerter foil / occlusion 4. Mild / intermittent symptoms relating misalignment or strabismus surgery (eyestrain / blur / photophobia / suture reaction) Failure 1. Diplopia / visual confusion in primary position and reading 2. ≥20PD heterotropia in primary position at Dist or Near 3. Using prism / Bangerter foil / occlusion 4. Moderate / severe symptoms related to misalignment or strabismus surgery		
(Hatt et al., 2018)	Identify factors associated with failure of AS-20 scores to improve following strabismus surgery	All strabismus patients – looked at failure to improve on each of the 4 AS-20 domains N=276 Age 18-91 years	PCT (SPCT) Near 1/3m and Dist 3m AS-20 (4 domains) Diplopia questionnaire Center for Epidemiologic Studies Depression Scale–Revised (CESD-R) (depressive symptoms) Type-D Scale 14 questionnaire (type-Distressed [type-D] personality)	6 weeks	Prospective
(Hatt et al., 2012a)	Changes in HRQoL 1 year after successful strabismus Sx	All strabismus patients included, all aetiologies N=73 Age 18-88 years	PCT (SPCT & PACT, but SPCT used in criteria) AS-20 Change in AS-20 psychosocial score Change in AS-20 function score Revised diplopia questionnaire Success: no/rare diplopia / visual confusion straight ahead at distance and for reading, <10PD heterotropia in primary position at distance and near Partial success: diplopia / visual confusion “sometimes” or less straight ahead distance and for reading (with or without prism), and <15PD heterotropia Failure: either diplopia / visual confusion was “often” or “always” straight ahead distance or for reading, >15PD heterotropia at distance or near, or the patient was using a Bangerter foil / occlusion	6 weeks (but between 4-14 weeks) 1 year (but between 5-22 months)	Retrospective

(Hatt et al., 2016)	Incorporating HRQoL into the assessment of outcome after strabismus surgery	Assess 'failures' by motor and diplopia criteria and evaluate change in HRQoL. Any strabismus type with and without diplopia. All aetiologies. N=227 Failures (n=40) Age 18-88 years	PCT (SPCT) Dist 3m and Near 1/3m Diplopia questionnaire AS-20 Motor criteria Diplopia criteria Failure: if 1 of the following criteria was met: (1) SPCT ≥15 PD (horizontal or vertical) at distance or near; (2) diplopia or visual confusion was present more than ''sometimes'' straight ahead at distance or for reading (unless atypical diplopia due to decompensated childhood strabismus was present preoperatively, in which case diplopia was allowed postoperatively); (3) occlusive patch / Bangerter foil needed. Partial success: SPCT ≤15 PD (horizontal and vertical) at distance and near, and diplopia / visual confusion was present never / rarely / sometimes. Correction of diplopia with prism was allowed. Success: if SPCT <10 PD (horizontal and vertical) at distance and near, and diplopia / visual confusion was present never or only rarely.	1 year (but between 5 months – 2 years)	Prospective
(Hertle, 1998)	Compare clinical characteristics of strabismus surgery with different onset	Compared strabismus onset before visual maturation (BVM) and after visual maturation (AVM). All surgery and all patients reported. N=255 Age 14-72 years	PCT BSV Subjective report Success – sensory: restoration of function field of BSV (>20◦), regaining central or peripheral fusion, orthotropia or heterophoria in primary position and at near Success – motor: absence of binocular function without diplopia, horizontal alignment <12PD and vertical alignment <5PD in primary position and near Success – subjective: subjective interpretation on improved eye position, binocular function and appearance (including happy / unhappy with eye position, tolerant / intolerant of residual diplopia, happy / unhappy with eye movement) Incomitance = difference ≥8PD	6 months – 5 years	Retrospective
(Jackson et al., 2006)	What are the psychosocial benefits of strabismus surgery	All strabismus patients. N=46 Age 16-61 years	PCT 1/3m Visual Analogue Scales (VAS) (0-10) for 5 questions on coping, lifestyle, worry, noticeable strabismus, strabismus severity Derriford Appearance Scale (DAS-24) Hospital Anxiety and Depression Scale (HADS)	3 months (but between 1-6 months)	Prospective

		(40% diplopia 60% no diplopia)	WHOQoLBref (four quality of life domains: physical, psychological, social, and environmental) BSV		
(Ji et al., 2020)	Self-reported sense of deviation in adults successfully aligned with surgery	All deviations N=91	PCT EOM BSV AS-20 (Chinese version) Self-report of deviation: no deviation / still have some deviation / still have obvious deviation (some and obvious were classed as self-reported sense of deviation) Success: (>1 year of follow up) no / rare diplopia / visual confusion in primary position and for reading, <10PD horizontal deviation, <5PD vertical deviation at near or dist	Follow up >1 year Last postoperative visit (12 – 42 months)	Retrospective
(Jung & Kim, 2018)	Surgery outcomes in sensory XT	Unilateral visual loss and constant horizontal strabismus VA <6/30 (0.7) N=64 Age 18-71 years	Success = <10PD dist Failure = recurrence or overcorrection Recurrence ≥10PD XT Overcorrection ≥10PD ET	1 year	Retrospective
(Kattan et al., 2016)	Binocular summation and stereoacuity after strabismus surgery	All types of strabismus and surgery N=130 Age 20 – 60 years	VA 100% contrast VA reduced contrast 2.5%, 1.25% in dimly lit room Binocular summation Stereoacuity near and dist Diplopia Measures only taken postoperatively	2 months	Prospective case series
(Keskinbora et al., 2011)	Long standing infantile ET – outcomes in late surgery	Alignment and BSV despite late surgery and early onset ET N=21 Age 8-26 years	PCT BSV <5PD heterotropia = orthotropia Residual ET ≥5PD ET Exotropia ≥5PD XT	3 – 9 years	Retrospective

(Kim et al., 2008)	Reoperation in sensory strabismus	N=11 Age 4-33 years	PCT Success = 0-10PD	1 month Last visit 1-48 months	Retrospective
(Kim et al., 2016)	Self-identity in strabismus and after surgery	N=351 Age 19 years + 3 groups Strabismus (n=96) Surgery age 4-15 years (n=108) No strabismus (n=147)	Korean self-identity scale (subscales: subjectivity, self-acceptance, future confidence, goal orientation, initiative, and familiarity)	3 independent groups – not before and after surgery	Retrospective
(Kishimoto & Ohtsuki, 2012)	VF14 in different ophthalmic conditions	Concomitant and incomitant strabismus N=625 Age 40-85 years	VF-14 questionnaire PCT BSV (Concomitant group)	3 months	Prospective
(Koc et al., 2013)	Strabismus surgery outcomes – does binocular vision make a difference to QoL	N=61 Age ≥18 years	AS-20 A&SQ (Amblyopia and strabismus questionnaire) BSV Diplopia score (from A&SQ) Motor success <10PD horizontal deviation and <5PD vertical deviation Sensory results BVP (binocular vision positive) and BVN (binocular vision negative)	3 months	Prospective
(Kushner, 1994)	Visual field (binocular or BEO) after surgery for ET	ET Sx N=37 Age 16-62 years	PCT Binocular VF (BEO) BSV (BG)	6 weeks	Prospective
(Kutschke & Scott, 2004)	PAT in ET (childhood onset, but Sx when visually mature)	All types of ET N=85 Age 9-70 years	Success 0-8PD SPCT at Near and Dist + peripheral fusion Those with no BSV postoperatively are reported	6 weeks to 13.7 years	Retrospective

(Lee et al., 2013)	Postoperative change in spatial localisation after XT surgery	XT N=60 Age 4-43 years	PCT Computer touch screen – spatial localisation (pointing errors)	1 day 1 month	Prospective
(Liebermann et al., 2013)	Compare long term outcomes in reoperation of horizontal strabismus-adjustment Vs no adjustment following surgery	ET and XT With and without potential BSV N=89 Age 12-83 years	Success: <10PD dist deviation (primary and near), no / rare diplopia (primary and reading), no prism or occlusion Partial success: ≤15PD dist deviation (primary and near) without prism, diplopia none / rare / sometimes in primary and reading, prism allowed, no occlusion Failure: if any of these are met >15PD dist deviation in primary or reading, diplopia always / often in primary and reading, needs occlusion	6 weeks (3-21 weeks) 1 year (23 weeks-2 years)	Retrospective
(Liebermann et al., 2014)	Improvement in specific function HRQoL concerns after strabismus surgery in nondiplopic adults	N=20 Age 22-79 years	Same success criteria as (Liebermann et al., 2013) AS-20 PCT BSV	1 year (but 6-19 months)	Retrospective
(Lipton & Willshaw, 1995)	Comparison of surgery accuracy – specialist centre compared to general	N=205 Age ?	PCT Success: Grade 1 within 0-5 PD of surgical goal Grade 2 within 6-10 PD of surgical goal Grade 3 >10PD of surgical goal	6 months	Prospective multicentre study
(McBain et al., 2014b)	QoL and mood postoperatively	Range of aetiologies N=210 Age 17-88 years	PCT (APCT 6m) Self-reports of pain, swelling, scarring, redness 0-10 scale At 3 months: Success: 3 out of 3 criteria met: <12PD ET/ XT / HT <20PD HoT, no / rare diplopia / visual confusion in primary position and reading, no prism / occlusion needed Partial success: 1 of the 3 criteria met Failure: 0 out of 3 criteria met AS-20	3 months 6 months	Prospective

			Success AS-20: >17.7 point increase in psychosocial subscale and >19.5 point increase in function subscale (>95% LOA) Psychosocial measures: Revised Illness Perception Questionnaire (IPQ-R) Treatment Representations Inventory (TRI) Fear of Negative Evaluation (FNE) scale The Derriford Appearance Scale (DAS24) Perceived Visibility of Strabismus Salience of Appearance scale (CARSAL) Valence of Appearance scale (CARVAL) Multidimensional Scale of Perceived Social Support (MSPSS) Hospital Anxiety and Depression Scale (HADS) Questionnaires: Reasons for strabismus surgery (RSSQ) Expectations of strabismus surgery (ESSQ) Additional questions: Do you regret having strabismus surgery: Yes definitely 1 – Not at all 4 Would you go through the surgery again: No hesitation at all 1 – Certainly not 4		
(Menon et al., 2002)	Psychosocial aspects of strabismus	All having surgery for alignment N=40 Age 15-25 years	Semi-structured interview to complete questionnaire and score questionnaire items (pre op and post op) Neuroticism questionnaire	3 months	Prospective
(Murray et al., 2007)	Changes in binocular status after late surgery for infantile ET	N=17 (if aligned 0-8PD at 1 day post op)	BSV (Worth 4 dot test, BG, Titmus, fusion on Synoptophore) Visual field BEO	Last follow up N=6 <1 month N=5 < 3 months N=6 >1 year	Retrospective
(Nelson et al., 2008)	Psychosocial impact of strabismus and surgery	N=128 Age ≥15 years N=20 teenagers N=108 adults	Postoperative telephone interviews to complete questionnaire about psychosocial issues (1-10) and postoperative outcome (1-7)	Unclear	Retrospective

(Ozates et al., 2019)	Psychological impact of strabismus surgery	N=83 Age 14-21 years XT & X(T)	Grouped by constant / manifest deviation XT or X(T) Turkish versions of: Social Appearance Anxiety Scale (SAAS) Depression subscale of the HADS (HAD-D Brief Fear of Negative Evaluation Scale (BFNE) state anxiety subscale of State-Trait Anxiety Inventory (STAI-S) trait anxiety subscale of State-Trait Anxiety Inventory (STAI-T)	1 year	Prospective
(Pineles et al., 2015)	Binocular summation after strabismus surgery	All strabismus types N=97 Age 2.5-90 years	VA high contrast (100%) VA low contrast (2.5% and 1.25%) Binocular summation calculation PCT Diplopia Success= 0-10PD horizontal strabismus and 0-4PD vertical strabismus	6-10 weeks	Prospective
(Ribeiro et al., 2014)	QoL in strabismus	N=101 Age 7-67 years 75% no surgery 25% had surgery	Semi-structured interviews to complete questionnaire (own modified version of AS-20)	?	Prospective
(Sandercoe et al., 2014)	Retrospective review of strabismus surgery	Categorised reasons for surgery (78% for psychosocial reasons) N=83 Mean age 37 years	PCT BSV Diplopia Objective criteria for success <10PD and acceptable 10-20PD results Subjective criteria = satisfaction with surgical outcome (very satisfied / satisfied / neutral / unsatisfied / very dissatisfied)	Mean 16 weeks	Retrospective
(Sim et al., 2018)	Factors associated with patient perception of success	N=87 Age 16-83 years 35% had no diplopia	AS-20 (used >95% limits of agreement as evidence of change) Diplopia PCT	24-126 days	unclear
(Wang & Nelson, 2011)	Sm-mod ET surgery outcomes	N=123 Age 11 months – 48 years	Success 0-5PD (PCT Near and Dist, primary position and lateral gaze)	6 months Last follow up (6 months-8 years)	Retrospective

(Wortham & Greenwald, 1989)	Binocular visual field in ET	N=10 Age 22-49 years	PCT BEO VF BSV	1-2 months	Retrospective
(Xu et al., 2012)	Psychosocial effect of strabismus surgery	N=56 Age 16-49 years No diplopia pre-op 64% surgery for BSV 36% had surgery for alignment	Own questionnaire (social function and psychological function scores) CT = fair alignment (small manifest deviation) or excellent alignment (no manifest deviation)	2-3 months	Prospective
(Xu et al., 2016)	Long term follow up and HRQoL following strabismus surgery	N=122 Compared AS-20 results to control group without strabismus N=89	AS-20 (Chinese version) PCT OM BSV Sense of deviation (no deviation / still have some deviation / still have obvious deviation) Diplopia	Last follow up 12-24 months)	Prospective

Appendix C. Ethical milestones throughout the study

Table C-1. Ethical approvals gained during the study

Phase	Approval	Date	Reference	Appendix
Phase one	REC North West – Liverpool Central REC	21/9/17	REC reference: 17/NW/0561 IRAS: 231502 Protocol: STH19274 version 1.0	C.1
	HRA	19/10/17	REC reference: 17/NW/0561 IRAS: 231502 Protocol: STH19274 version 1.0	C.2
	Non-substantial / Minor amendment	11/12/17	IRAS: 231502 Sponsor amendment notification number: NSA01 Protocol: STH19274 version 2.0	C.3
Phase two	REC London – Queen Square REC	22/11/18	REC reference: 18/LO/2013 IRAS: 256407 Protocol: STH20677 version 1.0	C.4
	HRA	22/11/18	REC reference: 18/LO/2013 IRAS: 256407 STH:20677 version 1.0	C.5
	Non-substantial / Minor amendment	18/3/20	IRAS: 256407 Protocol: STH20677 / 256407 version 2.0TC	C.6

Appendix C.1 REC approval – phase one

NHS
Health Research Authority
North West - Liverpool Central Research Ethics Committee
3rd Floor
Barlow House
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0207 104 8019

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

21 September 2017

Mrs Gemma Arblaster
NIHR Clinical Research fellow
University of Sheffield
E100f, Academic Unit of Ophthalmology and Orthoptics
School of Medicine and Biomedical Sciences, University of Sheffield
The Medical School, Beech Hill Road
S10 2RX

Dear Mrs Arblaster

Study title: **Measuring the effects of eye alignment surgery - a feasibility study. Phase one - what do patients report following eye alignment surgery?**
REC reference: **17/NW/0561**
Protocol number: **STH19274**
IRAS project ID: **231502**

Thank you for your responding to the Proportionate Review Sub-Committee's request for changes to the documentation for the above study.

The revised documentation has been reviewed and approved by the sub-committee.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact please contact hra.studyregistration@nhs.net outlining the reasons for your request.

Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above

research on the basis described in the application form, protocol and supporting documentation as revised.

Conditions of the favourable opinion

You should notify the REC once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. Revised documents should be submitted to the REC electronically from IRAS. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which you can make available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA Approval (England)/ NHS permission for research is available in the Integrated Research Application System, www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" above).

Approved documents

The documents reviewed and approved by the Committee are:

Document	*Version*	*Date*
Covering letter on headed paper [231502 Cover letter]	1.0	13 September 2017
Interview schedules or topic guides for participants [231502 Topic Guide for Interviews v1.0]	1.0	25 August 2017
IRAS Application Form [IRAS_Form_05092017]		05 September 2017
IRAS Application Form XML file [IRAS_Form_05092017]		05 September 2017
IRAS Checklist XML [Checklist_05092017]		05 September 2017
IRAS Checklist XML [Checklist_13092017]		13 September 2017
Participant consent form [231502 Consent form v2.0]	2.0	13 September 2017
Participant information sheet (PIS) [231502 Information sheet v2.0]	2.0	13 September 2017
Referee's report or other scientific critique report [Summary Reviews ICA-CDRF-2016-02-063]	1.0	13 April 2017
Research protocol or project proposal [231502 Protocol v1.0]	1.0	22 August 2017
Summary CV for Chief Investigator (CI) [231502 CV Gemma Arblaster v1.0]	1	25 August 2017
Summary CV for supervisor (student research) [David Buckley (lead supervisor) CV]	1.0	25 August 2017
Summary CV for supervisor (student research) [Prof Helen Davis]		
Summary CV for supervisor (student research) [Sarah Barnes]		
Summary, synopsis or diagram (flowchart) of protocol in non technical language [231502 Flow Chart v1.0]	1.0	22 August 2017

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance

We are pleased to welcome researchers and R & D staff at our RES Committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

17/NW/0561 **Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely

Mrs Julie Brake
Chair

Email: nrescommittee.northwest-liverpoolcentral@nhs.net

Enclosures: *"After ethical review – guidance for researchers" [SL-AR2]*

Appendix C.2 HRA approval – phase one

NHS
Health Research Authority

Mrs Gemma Arblaster
NIHR Clinical Research fellow
University of Sheffield
E100f, Academic Unit of Ophthalmology and Orthoptics
School of Medicine and Biomedical Sciences, University of Sheffield
The Medical School, Beech Hill Road
S10 2RX

Email: hra.approval@nhs.net

19 October 2017

Dear Mrs Arblaster

Letter of HRA Approval

Study title:	**Measuring the effects of eye alignment surgery - a feasibility study. Phase one - what do patients report following eye alignment surgery?**
IRAS project ID:	**231502**
Protocol number:	**STH19274**
REC reference:	**17/NW/0561**
Sponsor	**Sheffield Teaching Hospitals NHS Foundation Trust**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

It is critical that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details and further information about working with the research management function for each organisation can be accessed from www.hra.nhs.uk/hra-approval.

Appendices

The HRA Approval letter contains the following appendices:

- A – List of documents reviewed during HRA assessment
- B – Summary of HRA assessment

After HRA Approval

The document *"After Ethical Review – guidance for sponsors and investigators"*, issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

In addition to the guidance in the above, please note the following:

- HRA Approval applies for the duration of your REC favourable opinion, unless otherwise notified in writing by the HRA.
- Substantial amendments should be submitted directly to the Research Ethics Committee, as detailed in the *After Ethical Review* document. Non-substantial amendments should be submitted for review by the HRA using the form provided on the HRA website, and emailed to hra.amendments@nhs.net.
- The HRA will categorise amendments (substantial and non-substantial) and issue confirmation of continued HRA Approval. Further details can be found on the HRA website.

Scope

HRA Approval provides an approval for research involving patients or staff in NHS organisations in England.

If your study involves NHS organisations in other countries in the UK, please contact the relevant national coordinating functions for support and advice. Further information can be found at http://www.hra.nhs.uk/resources/applying-for-reviews/nhs-hsc-rd-review/.

If there are participating non-NHS organisations, local agreement should be obtained in accordance with the procedures of the local participating non-NHS organisation.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/.

HRA Training

We are pleased to welcome researchers and research management staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

Your IRAS project ID is **231502**. Please quote this on all correspondence.

Yours sincerely

Andrea Bell
Assessor

Email: hra.approval@nhs.net

Copy to: *Mrs Samantha Walmsley, Sheffield Teaching Hospitals NHS Foundation Trust – Sponsor's Representative*

Appendix A - List of Documents

The final document set assessed and approved by HRA Approval is listed below.

Document	*Version*	*Date*
Interview schedules or topic guides for participants [231502 Topic Guide for Interviews v1.0]	1.0	25 August 2017
IRAS Application Form [IRAS_Form_05092017]		05 September 2017
Letter from funder [Funding schedule]	1.0	
Letter from funder [NIHR confirmation]	1.0	19 January 2017
Participant consent form [231502 Consent form v2.0]	2.0	13 September 2017
Participant consent form [231502 Consent Form v1.0]	1.0	22 August 2017
Participant information sheet (PIS) [231502 Participant Information Sheet v1.0]	1.0	22 August 2017
Participant information sheet (PIS) [231502 Information sheet v2.0]	2.0	13 September 2017
Participant information sheet (PIS) [PIS]	3	05 October 2017
Referee's report or other scientific critique report [Summary Reviews ICA-CDRF-2016-02-063]	1.0	13 April 2017
Research protocol or project proposal [231502 Protocol v1.0]	1.0	22 August 2017
Summary CV for Chief Investigator (CI) [231502 CV Gemma Arblaster v1.0]	1	25 August 2017
Summary CV for supervisor (student research) [David Buckley (lead supervisor) CV]	1.0	25 August 2017
Summary CV for supervisor (student research) [Prof Helen Davis]		
Summary CV for supervisor (student research) [Sarah Barnes]		
Summary, synopsis or diagram (flowchart) of protocol in non technical language [231502 Flow Chart v1.0]	1.0	22 August 2017

IRAS project ID	231502

Section	HRA Assessment Criteria	Compliant with Standards	Comments
4.2	Insurance/indemnity arrangements assessed	Yes	The management, conduct and design of the study is covered by NHS indemnity. Where applicable, independent contractors (e.g. General Practitioners) should ensure that the professional indemnity provided by their medical defence organisation covers the activities expected of them for this research study
4.3	Financial arrangements assessed	Yes	No comments
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	REC favourable opinion issued on 21/09/2017
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments

IRAS project ID 231502

Appendix B - Summary of HRA Assessment

This appendix provides assurance to you, the sponsor and the NHS in England that the study, as reviewed for HRA Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England to assist in assessing and arranging capacity and capability.

For information on how the sponsor should be working with participating NHS organisations in England, please refer to the *participating NHS organisations, capacity and capability* and *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* sections in this appendix.

The following person is the sponsor contact for the purpose of addressing participating organisation questions relating to the study:

Name: Mrs Samantha Walmsley (Sheffield Teaching Hospitals NHS Foundation Trust)
Tel: 01142265932
Email: samantha.walmsley@sth.nhs.uk

HRA assessment criteria

Section	HRA Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	Application applies to Phase one only.
2.1	Participant information/consent documents and consent process	Yes	The applicant has clarified what will happen to personal data should the participant stop taking part. The applicant submitted an updated PIS version 3.0 dated 05/10/2017, to the HRA on 05/10/2017
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	This is a single site, same sponsor study; therefore an agreement is not required. The sponsor is not requesting and does not expect any other site agreement.

IRAS project ID 231502

Participating NHS Organisations in England

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.
There is only site type and the participating organisation will undertake the activities as described in the protocol. If this study is subsequently extended to other NHS organisation(s) in England, an amendment should be submitted to the HRA, with a Statement of Activities and Schedule of Events for the newly participating NHS organisation(s) in England. The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. For NIHR CRN Portfolio studies, the Local LCRN contact should also be copied into this correspondence. For further guidance on working with participating NHS organisations please see the HRA website. If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England which are not provided in IRAS or on the HRA website, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net. The HRA will work with these organisations to achieve a consistent approach to information provision.

Confirmation of Capacity and Capability

This describes whether formal confirmation of capacity and capability is expected from participating NHS organisations in England.
This is a single site study sponsored by the site. The R&D office will confirm to the CI when the study can start.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and the minimum expectations for education, training and experience that PIs should meet (where applicable).
The Clinical Research Fellow, who is a member of the direct care team, is the CI for the study. No PI or Local Collaborator is required. GCP training is not a generic training expectation, in line with the HRA statement on training expectations.

IRAS project ID 231502

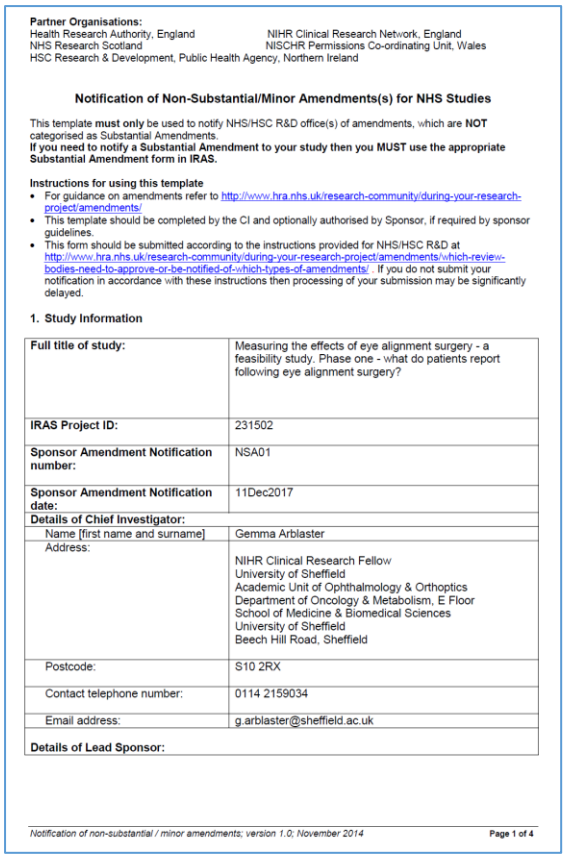
HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken
No Honorary Research Contracts, Letters of Access or pre-engagement checks are expected for local staff employed by the participating NHS organisations. Where arrangements are not already in place, research staff not employed by the NHS host organisation undertaking any of the research activities listed in the research application would be expected to obtain an honorary research contract. This would be on the basis of a Research Passport (if university employed) or an NHS to NHS confirmation of pre-engagement checks letter (if NHS employed). These should confirm enhanced DBS checks, including appropriate barred list checks, and occupational health clearance.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England to aid study set-up.
The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix C.3 Non-substantial amendment – phase one



Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

Notification of Non-Substantial/Minor Amendments(s) for NHS Studies

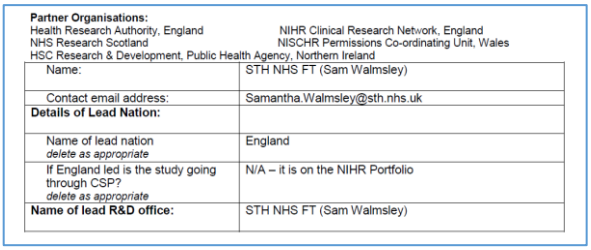
This template **must only** be used to notify NHS/HSC R&D office(s) of amendments, which are **NOT** categorised as Substantial Amendments.
If you need to notify a Substantial Amendment to your study then you MUST use the appropriate Substantial Amendment form in IRAS.

Instructions for using this template

- For guidance on amendments refer to http://www.hra.nhs.uk/research-community/during-your-research-project/amendments/
- This template should be completed by the CI and optionally authorised by Sponsor, if required by sponsor guidelines.
- This form should be submitted according to the instructions provided for NHS/HSC R&D at http://www.hra.nhs.uk/research-community/during-your-research-project/amendments/which-review-bodies-need-to-approve-or-be-notified-of-which-types-of-amendments/ . If you do not submit your notification in accordance with these instructions then processing of your submission may be significantly delayed.

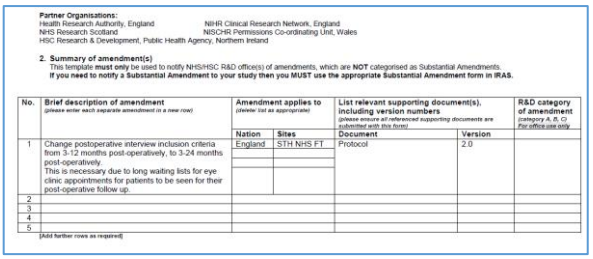
1. Study Information

Full title of study:	Measuring the effects of eye alignment surgery - a feasibility study. Phase one - what do patients report following eye alignment surgery?
IRAS Project ID:	231502
Sponsor Amendment Notification number:	NSA01
Sponsor Amendment Notification date:	11Dec2017
Details of Chief Investigator:	
Name [first name and surname]	Gemma Arblaster
Address:	NIHR Clinical Research Fellow University of Sheffield Academic Unit of Ophthalmology & Orthoptics Department of Oncology & Metabolism, E Floor School of Medicine & Biomedical Sciences University of Sheffield Beech Hill Road, Sheffield
Postcode:	S10 2RX
Contact telephone number:	0114 2159034
Email address:	g.arblaster@sheffield.ac.uk
Details of Lead Sponsor:	



Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

Name:	STH NHS FT (Sam Walmsley)
Contact email address:	Samantha.Walmsley@sth.nhs.uk
Details of Lead Nation:	
Name of lead nation *delete as appropriate*	England
If England led is the study going through CSP? *delete as appropriate*	N/A – it is on the NIHR Portfolio
Name of lead R&D office:	STH NHS FT (Sam Walmsley)

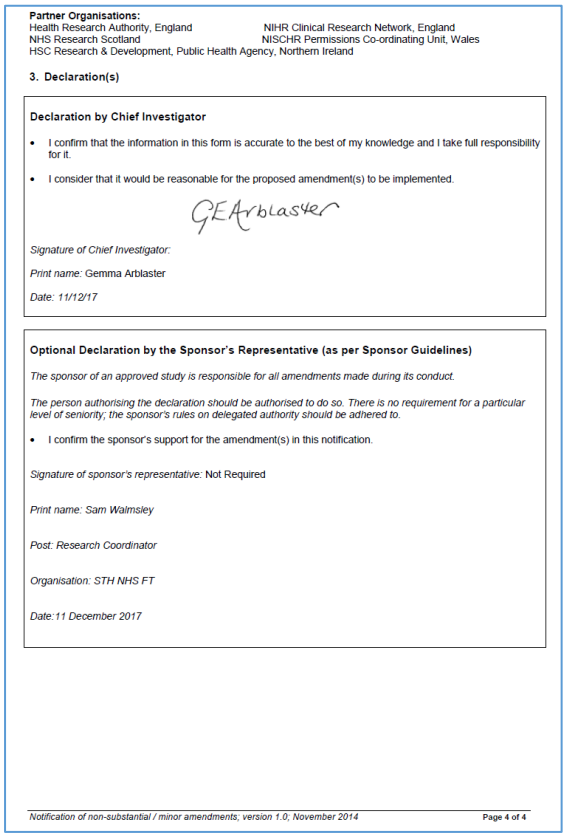


Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

2. Summary of amendment(s)
This template **must only** be used to notify NHS/HSC R&D office(s) of amendments, which are **NOT** categorised as Substantial Amendments.
If you need to notify a Substantial Amendment to your study then you MUST use the appropriate Substantial Amendment form in IRAS.

No.	Brief description of amendment *(please enter each separate amendment in a new row)*	Amendment applies to *(delete 'list as appropriate)*		List relevant supporting document(s), including version numbers *(please ensure all referenced supporting documents are submitted with this form)*		R&D category of amendment *(category A, B, C) For office use only*
		Nation	Sites	Document	Version	
1	Change postoperative interview inclusion criteria from 3-12 months post-operatively, to 3-24 months post-operatively. This is necessary due to long waiting lists for eye clinic appointments for patients to be seen for their post-operative follow up.	England	STH NHS FT	Protocol	2.0	
2						
3						
4						
5						

[Add further rows as required]



Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

3. Declaration(s)

Declaration by Chief Investigator

- I confirm that the information in this form is accurate to the best of my knowledge and I take full responsibility for it.
- I consider that it would be reasonable for the proposed amendment(s) to be implemented.

GE Arblaster

Signature of Chief Investigator:

Print name: Gemma Arblaster

Date: 11/12/17

Optional Declaration by the Sponsor's Representative (as per Sponsor Guidelines)

The sponsor of an approved study is responsible for all amendments made during its conduct.

The person authorising the declaration should be authorised to do so. There is no requirement for a particular level of seniority; the sponsor's rules on delegated authority should be adhered to.

- I confirm the sponsor's support for the amendment(s) in this notification.

Signature of sponsor's representative: Not Required

Print name: Sam Walmsley

Post: Research Coordinator

Organisation: STH NHS FT

Date:11 December 2017

Appendix C.4 REC approval – phase two

NHS
Health Research Authority

London - Queen Square Research Ethics Committee
HRA NRES Centre Manchester
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0207 104 8019

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

22 November 2018

Mrs Gemma Arblaster
NIHR Clinical Research Fellow
University of Sheffield
Academic Unit of Ophthalmology and Orthoptics, Department of Oncology and Metabolism,
Medical School (Room E104c), Faculty of Medicine, Dentistry and Health,
University of Sheffield, Beech Hill Road
S10 2RX

Dear Mrs Arblaster

Study title:	**Measuring the effects of eye alignment surgery - a feasibility study.Phase two - Measuring vision and task performance before and after eye alignment surgery.**
REC reference:	**18/LO/2013**
IRAS project ID:	**256407**

Thank you responding to the Proportionate Review Sub-Committee's request for changes to the documentation for the above study.

The revised documentation has been reviewed and approved by the sub-committee.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact please contact hra.studyregistration@nhs.net outlining the reasons for your request.

Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

A Research Ethics Committee established by the Health Research Authority

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

You should notify the REC once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. Revised documents should be submitted to the REC electronically from IRAS. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which you can make available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System, at www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

A Research Ethics Committee established by the Health Research Authority

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" above).

Approved documents

The documents reviewed and approved by the Committee are:

Document	*Version*	*Date*
IRAS Application Form [IRAS_Form_07112018]		07 November 2018
IRAS Application Form XML file [IRAS_Form_29102018]		29 October 2018
IRAS Application Form XML file [IRAS_Form_07112018]		07 November 2018
IRAS Checklist XML [Checklist_29102018]		29 October 2018
IRAS Checklist XML [Checklist_07112018]		07 November 2018
Letter from funder [Letter of Intent ICA-CDRF-2016-02-063]		19 January 2018
Letters of invitation to participant [256407 Invitation letter v1.0]	1.0	30 October 2018
Letters of invitation to participant [256407 Invitation email v1.0]	1.0	30 October 2018
Non-validated questionnaire [256407 study questionnaire v1.0]	1.0	24 October 2018
Other [ICA-CDRF-2016-02-063-Award Holder Acceptance-04-01-2017 16-22-03]		
Other [18 LO 2013 G Arblaster response to REC 061118]	1.0	06 November 2018
Participant consent form [256407 Consent Form v1.0]	1.0	24 October 2018
Participant information sheet (PIS) [256407 Information sheet v3.0]	3.0	05 November 2018
Participant information sheet (PIS) [256407 Information sheet v4.0]	4.0	05 November 2018
Referee's report or other scientific critique report [Summary Reviews ICA-CDRF-2016-02-063 20170413-142822]	1.0	19 May 2016
Research protocol or project proposal [256407 Protocol v1.0]	1.0	24 October 2018
Summary CV for Chief Investigator (CI) [256407 Arblaster CV v1.0]	1.0	24 October 2018
Summary CV for student [256407 Arblaster CV v1.0]	1.0	24 October 2018
Summary CV for supervisor (student research) [David Buckley CV]	1.0	24 October 2018
Summary CV for supervisor (student research) [Sarah Barnes CV signed]	1.0	05 July 2017
Summary CV for supervisor (student research) [Prof Helen Davis CV]	1.0	24 October 2018
Summary, synopsis or diagram (flowchart) of protocol in non technical language [256407 Flow chart v1.0]	1.0	24 October 2018
Validated questionnaire [256407 AS_20_Questionnaire v1.0]	1.0	24 October 2018
Validated questionnaire [256407 VFQ_25 Questionnaire v1.0]	1.0 (Version 2000 written in header by authors)	01 January 2000

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

A Research Ethics Committee established by the Health Research Authority

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance

We are pleased to welcome researchers and R & D staff at our RES Committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

18/LO/2013	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely

PP: Dr Eamonn Walsh
Chair

Email: nrescommittee.london-queensquare@nhs.net

Enclosures: *"After ethical review – guidance for researchers" [SL-AR2]*

Copy to: *Ms Jemima Clarke, Sheffield Teaching Hospitals NHS Foundation Trust*

Appendix C.5 HRA approval – phase two

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Mrs Gemma Arblaster
NIHR Clinical Research Fellow
University of Sheffield
Academic Unit of Ophthalmology and Orthoptics, Department of Oncology and Metabolism,
Medical School (Room E104c), Faculty of Medicine, Dentistry and Health,
University of Sheffield, Beech Hill Road
S10 2RX

Email: hra.approval@nhs.net
Research-permissions@wales.nhs.uk

22 November 2018

Dear Mrs Arblaster

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Measuring the effects of eye alignment surgery - a feasibility study.Phase two - Measuring vision and task performance before and after eye alignment surgery.**
IRAS project ID:	**256407**
REC reference:	**18/LO/2013**
Sponsor	**Sheffield Teaching Hospitals NHS Foundation Trust**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

How should I continue to work with participating NHS organisations in England and Wales?
You should now provide a copy of this letter to all participating NHS organisations in England and Wales, as well as any documentation that has been updated as a result of the assessment.

This is a single site study sponsored by the site. The sponsor R&D office will confirm to you when the study can start following issue of HRA and HCRW Approval.

It is important that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details of the research management function for each organisation can be accessed here.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?

HRA and HCRW Approval does not apply to NHS/HSC organisations within the devolved administrations of Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) has been sent to the coordinating centre of each participating nation. You should work with the relevant national coordinating functions to ensure any nation specific checks are complete, and with each site so that they are able to give management permission for the study to begin.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?
The document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

I am a participating NHS organisation in England or Wales. What should I do once I receive this letter?
You should work with the applicant and sponsor to complete any outstanding arrangements so you are able to confirm capacity and capability in line with the information provided in this letter.

The sponsor contact for this application is as follows:

Name: Ms Jemima Clarke
Tel: 0114226 5943
Email: Jemima.Clarke@sth.nhs.uk

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **256407**. Please quote this on all correspondence.

Yours sincerely

Lucy Roberts
Assessor

Email: hra.approval@nhs.net

Copy to: *Ms Jemima Clarke, Sponsor Contact, Sheffield Teaching Hospitals NHS Foundation Trust*
Dr David Buckley, Academic Supervisor, University of Sheffield
Dr Sarah Barnes , Academic Supervisor, University of Sheffield
Professor Helen Davis , Academic Supervisor, University of Sheffield

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
IRAS Application Form [IRAS_Form_07112018]		07 November 2018
Letter from funder [Letter of Intent ICA-CDRF-2016-02-063]		19 January 2018
Letters of invitation to participant [256407 Invitation letter v1.0]	1.0	30 October 2018
Letters of invitation to participant [256407 Invitation email v1.0]	1.0	30 October 2018
Non-validated questionnaire [256407 study questionnaire v1.0]	1.0	24 October 2018
Other [ICA-CDRF-2016-02-063-Award Holder Acceptance-04-01-2017 16-22-03]		
Other [18 LO 2013 G Arblaster response to REC 061118]	1.0	06 November 2018
Participant consent form [256407 Consent Form v1.0]	1.0	24 October 2018
Participant information sheet (PIS) [256407 Information sheet v3.0]	3.0	05 November 2018
Participant information sheet (PIS) [256407 Information sheet v4.0]	4.0	05 November 2018
Referee's report or other scientific critique report [Summary Reviews ICA-CDRF-2016-02-063 20170413-142822]	1.0	19 May 2016
Research protocol or project proposal [256407 Protocol v1.0]	1.0	24 October 2018
Summary CV for Chief Investigator (CI) [256407 Arblaster CV v1.0]	1.0	24 October 2018
Summary CV for student [256407 Arblaster CV v1.0]	1.0	24 October 2018
Summary CV for supervisor (student research) [David Buckley CV]	1.0	24 October 2018
Summary CV for supervisor (student research) [Sarah Barnes CV signed]	1.0	05 July 2017
Summary CV for supervisor (student research) [Prof Helen Davis CV]	1.0	24 October 2018
Summary, synopsis or diagram (flowchart) of protocol in non technical language [256407 Flow chart v1.0]	1.0	24 October 2018
Validated questionnaire [256407 AS_20_Questionnaire v1.0]	1.0	24 October 2018
Validated questionnaire [256407 VFQ_25 Questionnaire v1.0]	1.0 (Version 2000 written in header by authors)	01 January 2000

Summary of assessment
The following information provides assurance to you, the sponsor and the NHS in England and Wales that the study, as assessed for HRA and HCRW Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England and Wales to assist in assessing, arranging and confirming capacity and capability.

Assessment criteria

Section	Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	No comments
2.1	Participant information/consent documents and consent process	Yes	No comments
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	An agreement will not be required as this is a single site study, where the single site is also the study sponsor.
4.2	Insurance/indemnity arrangements assessed	Yes	NHS indemnity applies.
4.3	Financial arrangements assessed	Yes	Funding has been secured from NIHR.
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	No comments

IRAS project ID	256407

Section	Assessment Criteria	Compliant with Standards	Comments
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments

Participating NHS Organisations in England and Wales

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.

There is one site-type. This is a single site study, where the single NHS site is also the study Sponsor.

If this study is subsequently extended to other NHS organisation(s) in England or Wales, an amendment should be submitted, with a Statement of Activities and Schedule of Events for the newly participating NHS organisation(s) in England or Wales.

The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England and Wales in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. Where applicable, the local LCRN contact should also be copied into this correspondence.

If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England and Wales which are not provided in IRAS, the HRA or HCRW websites, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net or HCRW at Research-permissions@wales.nhs.uk. We will work with these organisations to achieve a consistent approach to information provision.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and Wales, and the minimum expectations for education, training and experience that PIs should meet (where applicable).

A Principal Investigator is expected at the site.

If this study is extended to other sites in England or Wales, a further assessment of the need for a PI or LC at the additional sites will be made.

GCP training is not a generic training expectation, in line with the HRA/HCRW/MHRA statement on training expectations.

IRAS project ID	256407

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken

Access arrangements will not be applicable, as all study activity is being conducted by staff employed by the site.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England and Wales to aid study set-up.

The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix C.6 Non-substantial amendment – phase two

Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

Notification of Non-Substantial/Minor Amendments(s) for NHS Studies

This template **must only** be used to notify NHS/HSC R&D office(s) of amendments, which are **NOT** categorised as Substantial Amendments.
If you need to notify a Substantial Amendment to your study then you MUST use the appropriate Substantial Amendment form in IRAS.

Instructions for using this template

- For guidance on amendments refer to http://www.hra.nhs.uk/research-community/during-your-research-project/amendments/
- This template should be completed by the CI and optionally authorised by Sponsor, if required by sponsor guidelines.
- This form should be submitted according to the instructions provided for NHS/HSC R&D at http://www.hra.nhs.uk/research-community/during-your-research-project/amendments/which-review-bodies-need-to-approve-or-be-notified-of-which-types-of-amendments/ . If you do not submit your notification in accordance with these instructions then processing of your submission may be significantly delayed.

1. Study Information

Full title of study:	Measuring the Effects of Eye Alignment Surgery Phase two – Measuring vision and task performance before and after eye alignment surgery
IRAS Project ID:	IRAS 256407
Sponsor Amendment Notification number:	
Sponsor Amendment Notification date:	18/3/20
Details of Chief Investigator:	
Name [first name and surname]	Gemma Arblaster
Address:	E104c, Division of Ophthalmology & Orthoptics, Health Sciences School, University of Sheffield, Beech Hill Road,
Postcode:	S10 2RX
Contact telephone number:	0114 2159055
Email address:	g.arblaster@sheffield.ac.uk
Details of Lead Sponsor:	

Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

Name:	Sheffield Teaching Hospitals NHS Foundation Trust
Contact email address:	Jemima.clarke2@nhs.net
Details of Lead Nation:	
Name of lead nation *delete as appropriate*	England
If England led is the study going through CSP? *delete as appropriate*	Yes / No
Name of lead R&D office:	Sheffield Teaching Hospitals NHS Foundation Trust – Research and Innovation Office

Partner Organisations:
Health Research Authority, England — NIHR Clinical Research Network, England
NHS Research Scotland — NISCHR Permissions Co-ordinating Unit, Wales
HSC Research & Development, Public Health Agency, Northern Ireland

2. Summary of amendment(s)
This template **must only** be used to notify NHS/HSC R&D office(s) of amendments, which are **NOT** categorised as Substantial Amendments.
If you need to notify a Substantial Amendment to your study then you MUST use the appropriate Substantial Amendment form in IRAS.

No.	Brief description of amendment *(please enter each separate amendment in a new row)*	Amendment applies to *(delete/ list as appropriate)*		List relevant supporting document(s), including version numbers *(please ensure all referenced supporting documents are submitted with this form)*		R&D category of amendment *(category A, B, C) For office use only*
		Nation	Sites	Document	Version	
1	Due to COVID-19. At the typical post-operative visit time (minimum 3 months after surgery) patients will be invited to complete the questionnaires only (AS-20, VFQ-25 and the additional study questions). Patients will be offered the options of completing and returning these questionnaires by email or by post, or completing the questionnaires during a telephone interview with the researcher. If a telephone interview is preferred, the patient will be given their choice of a preferred time to be telephoned by the researcher.	England	Single site – Sheffield Teaching Hospitals NHS Foundation Trust	Protocol (256407 Protocol v2.0 TC)	2	
2	Due to COVID-19. The final study visit (post-operative visit) will be offered to the patient when face-to-face consultations are allowed. The timing of this visit will remain as a minimum of 3 months post-operatively.	England. Single site – Sheffield Teaching Hospitals NHS Foundation Trust		Protocol (256407 Protocol v2.0 TC)	2	

	When face-to-face consultations are allowed, patients will be offered a study visit at the University of Sheffield on a separate day (as per the original protocol). As this time may be some time before a clinical appointment can be offered (due to expected clinical appointment back logs), patients will be offered the option of having their study visit results copied into their clinical notes. If patients decline a study visit at the University of Sheffield and would rather wait until clinical outpatient appointments resume at Sheffield Teaching Hospitals NHS Foundation Trust, patients will be offered a study visit on the same day as their planned clinical appointment (as per the original protocol).				
3	Due to COVID-19. The study end date will be extended to allow all patients who have been recruited to the study and have already taken part in their pre-operative visit to be seen for their post-operative visit and complete their involvement in the study.	England. Single site – Sheffield Teaching Hospitals NHS Foundation Trust	Protocol (256407 Protocol v2.0 TC)	2	
4	For the grading of the cover test result, a 7th additional option will be included '7 – heterophoria' to allow the inclusion of any unexpected binocular results post-operatively to be graded	England. Single site – Sheffield Teaching Hospitals NHS Foundation Trust	Protocol (256407 Protocol v2.0 TC)	2	
5					

[Add further rows as required]

3. Declaration(s)

Declaration by Chief Investigator

- I confirm that the information in this form is accurate to the best of my knowledge and I take full responsibility for it.
- I consider that it would be reasonable for the proposed amendment(s) to be implemented.

GEArblaster

Signature of Chief Investigator:

Print name: Gemma Arblaster

Date: 18/3/20

Optional Declaration by the Sponsor's Representative (as per Sponsor Guidelines)

The sponsor of an approved study is responsible for all amendments made during its conduct.

The person authorising the declaration should be authorised to do so. There is no requirement for a particular level of seniority; the sponsor's rules on delegated authority should be adhered to.

- I confirm the sponsor's support for the amendment(s) in this notification.

Signature of sponsor's representative: .....

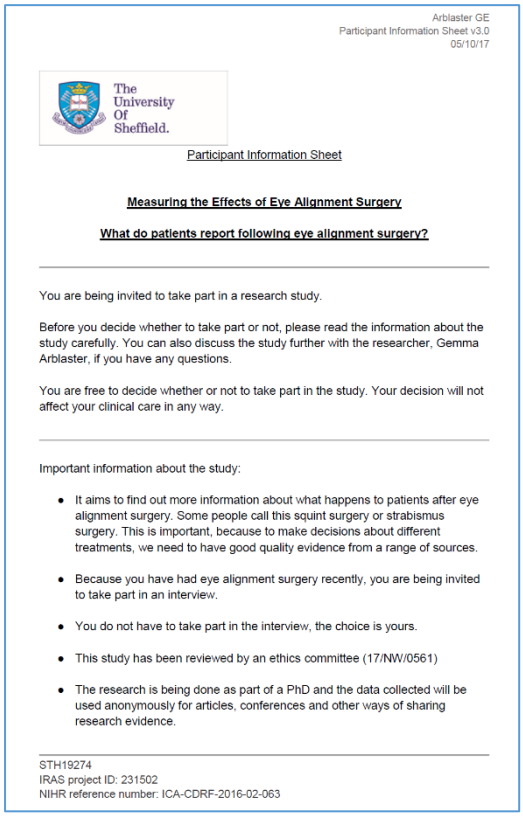
Print name:.....

Post: .....

Organisation:.....

Date:.....

Appendix D. Participant information sheet - phase one



The University Of Sheffield.

Participant Information Sheet

Measuring the Effects of Eye Alignment Surgery

What do patients report following eye alignment surgery?

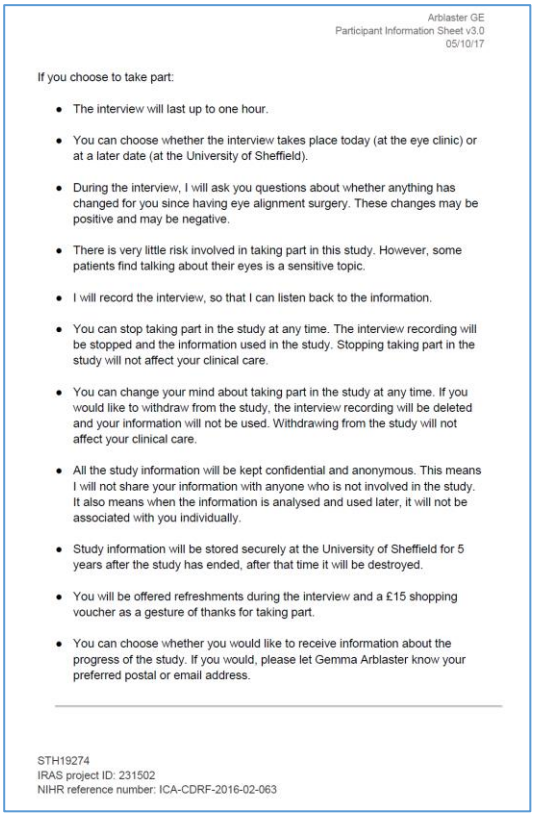
You are being invited to take part in a research study.

Before you decide whether to take part or not, please read the information about the study carefully. You can also discuss the study further with the researcher, Gemma Arblaster, if you have any questions.

You are free to decide whether or not to take part in the study. Your decision will not affect your clinical care in any way.

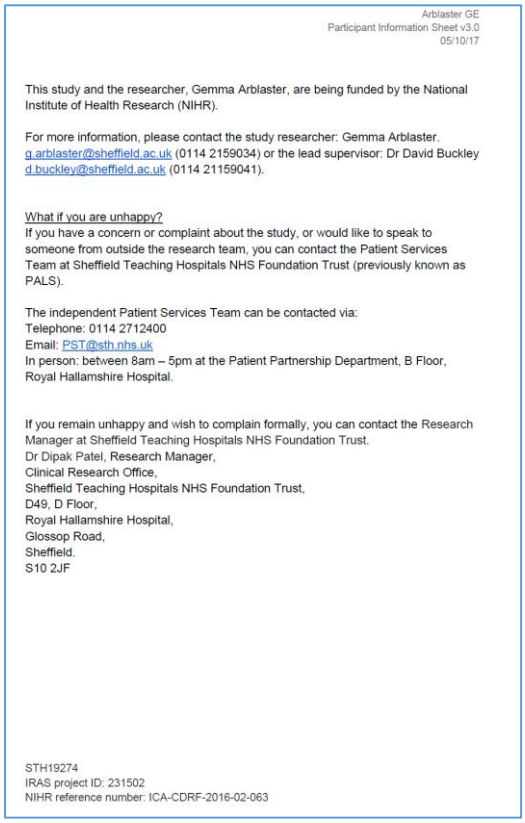
Important information about the study:

- It aims to find out more information about what happens to patients after eye alignment surgery. Some people call this squint surgery or strabismus surgery. This is important, because to make decisions about different treatments, we need to have good quality evidence from a range of sources.
- Because you have had eye alignment surgery recently, you are being invited to take part in an interview.
- You do not have to take part in the interview, the choice is yours.
- This study has been reviewed by an ethics committee (17/NW/0561)
- The research is being done as part of a PhD and the data collected will be used anonymously for articles, conferences and other ways of sharing research evidence.



If you choose to take part:

- The interview will last up to one hour.
- You can choose whether the interview takes place today (at the eye clinic) or at a later date (at the University of Sheffield).
- During the interview, I will ask you questions about whether anything has changed for you since having eye alignment surgery. These changes may be positive and may be negative.
- There is very little risk involved in taking part in this study. However, some patients find talking about their eyes is a sensitive topic.
- I will record the interview, so that I can listen back to the information.
- You can stop taking part in the study at any time. The interview recording will be stopped and the information used in the study. Stopping taking part in the study will not affect your clinical care.
- You can change your mind about taking part in the study at any time. If you would like to withdraw from the study, the interview recording will be deleted and your information will not be used. Withdrawing from the study will not affect your clinical care.
- All the study information will be kept confidential and anonymous. This means I will not share your information with anyone who is not involved in the study. It also means when the information is analysed and used later, it will not be associated with you individually.
- Study information will be stored securely at the University of Sheffield for 5 years after the study has ended, after that time it will be destroyed.
- You will be offered refreshments during the interview and a £15 shopping voucher as a gesture of thanks for taking part.
- You can choose whether you would like to receive information about the progress of the study. If you would, please let Gemma Arblaster know your preferred postal or email address.



This study and the researcher, Gemma Arblaster, are being funded by the National Institute of Health Research (NIHR).

For more information, please contact the study researcher: Gemma Arblaster, g.arblaster@sheffield.ac.uk (0114 2159034) or the lead supervisor: Dr David Buckley d.buckley@sheffield.ac.uk (0114 21159041).

What if you are unhappy?
If you have a concern or complaint about the study, or would like to speak to someone from outside the research team, you can contact the Patient Services Team at Sheffield Teaching Hospitals NHS Foundation Trust (previously known as PALS).

The independent Patient Services Team can be contacted via:
Telephone: 0114 2712400
Email: PST@sth.nhs.uk
In person: between 8am – 5pm at the Patient Partnership Department, B Floor, Royal Hallamshire Hospital.

If you remain unhappy and wish to complain formally, you can contact the Research Manager at Sheffield Teaching Hospitals NHS Foundation Trust.
Dr Dipak Patel, Research Manager,
Clinical Research Office,
Sheffield Teaching Hospitals NHS Foundation Trust,
D49, D Floor,
Royal Hallamshire Hospital,
Glossop Road,
Sheffield.
S10 2JF

Appendix E. Consent form – phase one

Arblaster GE
Consent form v2.0
13/9/17

The University Of Sheffield.

Participant Identification Number:

Consent Form

Measuring the Effects of Eye Alignment Surgery.

What do patients report following eye alignment surgery?

Name of Researcher: Gemma Arblaster

Please initial each box:

1. I confirm that I have read the information sheet dated __/__/20__ (version__) and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my medical notes may be looked at by individuals from the University of Sheffield or from Sheffield Teaching Hospitals NHS Foundation Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I give permission for my interview to be recorded. ☐
5. I understand the information I provide, as part of this study, will only be accessed by members of the research team and this will be stored securely at the University of Sheffield. ☐
6. I understand that my information, as part of this study, will be used to support other research in the future, and may be shared anonymously with other researchers. ☐
7. I understand that I can receive information about the progress of the study if I wish, and this will require me to provide the researcher with my preferred postal or email address. ☐
8. I agree to take part in the above study. ☐

Name of Participant Date Signature

Name of Person taking consent Date Signature

IRAS ID: 231502
Study Number: STH19274
NIHR reference number: ICA-CDRF-2016-02-063

Copies to: patient
investigator site file
medical notes

Appendix F. Topic guide – semi structured interviews

Amendments to the topic guide during the research shown in red

Introduction

- Greet and thank participant for taking part in study, introduce self (Gemma Arblaster) and explain the purpose of the research.
- Remind participant interview will be recorded and explain I may also take notes during the interview.
- Remind patients interview is confidential.
- Explain that during the interview we might discuss things that you find difficult. You don't have to answer any questions that make you feel uncomfortable, but at the same time, I'm very interested in your experiences. If there are any questions that you are uncomfortable with, or you find difficult to answer, we can move on so please don't worry.
- I am interested in your views and they are personal to you. There is no right or wrong answer to the questions, it is your experiences that are important. Positive and negative experiences are both important, as are the experiences where nothing has changed, so don't feel as though you only have to say good things during the interview.
- Explain interview format.

Topics

1. Would you like to tell me a bit about your eye position and the operation you had?

2. How do you feel after having eye alignment surgery?
[prompts: What happened for you after the surgery? How does that make you feel? What effect has the surgery had on you/your life? Can you give me any examples of how you feel after having surgery?]

3. Do you think anything has changed for you after having eye alignment surgery?
[prompts: Is anything different after the surgery? Is anything better? Is anything worse? Have some things changed? Have some things not changed? How have things changed/not changed your life? Do you think surgery has made any difference to you?]

4. Has anything about your vision changed after having eye alignment surgery?
[prompts: Vision / how you use your eyes / how your eyes work. Is your vision the same/different after the surgery? How has your vision changed? Is your vision better or worse now? Can you give an example? Do you think your vision at the side (peripheral vision) has changed? When do you notice that change in your vision? Do you think the change in your vision makes a difference to you?]

5. Has anything about your ability to perform tasks changed after having the eye alignment surgery?
[prompts: Perform tasks / do things / do the things you enjoy / do the things you need to do. Is your ability to perform tasks different/the same after the surgery? How has your ability to perform tasks changed? Is your ability to perform tasks better or worse now? Can you give an example? Do you think that change in task performance makes a difference to you?]

6. Has anything about your daily life changed after having the eye alignment surgery?
[prompts: Daily life / activities / work. Is anything about your daily life different/the same after the surgery? How has your daily life changed? Is anything about your daily life better or worse now? Can you give an example?]

7. Did you need to take breaks because of your eyes? Do you think surgery has made any difference to taking breaks?

8. Has anything about your depth perception changed after having the eye alignment surgery?
[prompts: Depth perception / 3D vision / ability to see depth / ability to judge distances. Is anything about your depth perception different/the same after the surgery? How has your depth perception changed? Is anything about your depth perception better or worse now? Can you give an example?]

9. Do you think you see in 3D now? Do you think you used to see in 3D before the surgery? Have you ever tried a 3D film or game? *[can you give an example of that?]*

10. Do you think surgery has made a difference to your confidence? *[can you give an example of that?]*

11. Do you think surgery has made a difference to busy environments or being in busy places? *[can you give an example of that?]*

12. When your eye was turning did you ever feel like it was getting in the way? Does it still feel like that now after your surgery? Did you ever need to close one eye? Do you still do that now after the surgery? *[can you give an example of that?]*

13. Did you ever feel like you could control the eye position? How does that compare to now after the surgery? *[can you give an example of that?]*

14. Was the eye ever uncomfortable? How does the eye feel after the surgery? Did you ever need to close the eye at all? Do you still need to do that now? *[can you give an example of that?]*

15. Do you feel like your balance was affected? Do you think there were any changes in your balance after the surgery? *[can you give an example of that?]*

16. Do you drive? Has anything about driving changed after the surgery? *[can you give an example of that?]*

17. Do you think you use your eyes together as a pair? Do you think you used to use your eyes together before the surgery? *[can you give an example of that?]*

18. Do you think there is anything different about your eye movements / the way your eyes move since having surgery? *[can you give an example of that?]*

19. Do you think anything else has changed for you after having the eye alignment surgery?
[prompts: Anything else that you feel is important or have noticed. This could be anything – small or large. Is anything else different/the same after the surgery? Has anything else changed since the surgery? Is anything else better or worse now? Can you give an

example? Work? Social activities? New activities? Old activities? Do you think you try have tried new things after the surgery? Do you think your ability to do things has changed after the surgery? Do you think you do things faster or slower after the surgery?]

20. Has anything happened to you after eye alignment surgery that you didn't expect?
[prompts: Has anything unexpected happened after the surgery? Is this a good or bad thing? Did anything happen/not happen that you were told would/wouldn't happen after the surgery? How has this affected you/your life? can you give an example of that?]

21. Is there anything that you wanted to change following eye alignment surgery that didn't change?
[prompts: This could be something positive or negative can you give an example of that?]

22. Do you think surgery made anything better for you (that we haven't already talked about)?

23. Do you think surgery made anything worse for you (that we haven't already talked about)?

24. Do you think anything stayed the same after surgery (that we haven't already talked about)?

25. Was there anything that happened to you after the surgery that you didn't expect? *[can you give an example of that?]*

26. Is there anything you wish you had known before the surgery that you know now?
[prompts: This could be something positive or negative can you give an example of that?]]

27. Are you satisfied with the result of your eye alignment surgery or not?
[prompts: Are you happy or unhappy with the result of your surgery? Do you feel the surgery has made a difference to your eye position? can you give an example of that?]

28. Is there anything else you would like to share about your experiences since having eye alignment surgery?

29. Do you feel any important issues have been left out of the discussion so far?

Close
Thank the participant and ensure they have the Participant Information Sheet, which gives researcher contact details if they need further support or information.

Appendix G. Coding framework

<table>
<tr><td colspan="3">Table G-1. A summary of the coding framework, including the initial codes and categories, the development of the themes and the final themes from the qualitative analysis</td></tr>
<tr><td>Initial categories and codes in coding framework</td><td>Development of themes</td><td>Final themes</td></tr>
<tr><td>Activities of daily life
Driving
Social
Work
Other</td><td rowspan="2">Reviewed codes:
Activities of daily life
Adaptations
Changes postoperatively
Describing examples of visual or task based activities. Examples given of scenarios where vision or task performance, or both, are described. Codes merged and separated into ‘vision’ or ‘task performance’ themes.</td><td>Vision</td></tr>
<tr><td>Adaptations

Changes post-operatively</td><td>Task performance</td></tr>
<tr><td>Emotions
Anxiety
Embarrassment
Feeling hurt
Fitting in or belonging
Frustration
Nervous
Relief
Stress
Upset
Worry</td><td rowspan="2">Reviewed codes:
Emotions
Self-image
Describing emotions, issues around confidence and self-perception. Codes merged into ‘confidence and emotions’.</td><td rowspan="2">Confidence and emotions</td></tr>
<tr><td>Self-image
Confidence and perception of self
Views of others</td></tr>
<tr><td>Signs and symptoms
Pain and discomfort
Visual</td><td>Reviewed Codes:
Signs and symptoms</td><td>Physical symptoms</td></tr>
</table>

	Describing physical aspects of strabismus not covered by other themes. Keep ‘physical symptoms’ as separate theme. Visual signs or symptoms moved to ‘vision’ theme.	
Previous and other treatments	Merged into final themes: ‘vision’, ‘task performance’, ‘physical symptoms’ and ‘confidence and emotions’.	
Other	Merged into final themes: ‘vision’, ‘task performance’, ‘physical symptoms’ and ‘confidence and emotions’.	

Appendix H. Clinical characteristics of the participants – phase one

Table H-1. The clinical characteristics of the phase one participants (n=13).

Participant number	Gender	Age group	Pre-op diagnosis (strabismus)	Pre-op diagnosis (other)	Surgery	Post-op diagnosis (strabismus)	VA better eye	VA worse eye	Postoperative time (months)
001	F	18-35	Longstanding L XT & L HT ? congenital R superior division 3rd NP (fixes RE so L XT & HT)	anisometropic amblyopia	L SR recession	residual L HT & XT	-0.1	0.98 (PH 0.86)	11
002	F	18-35	secondary ET	strabismic amblyopia	L MR recession	residual ET (Nr) & consecutive XT (Dist)	0	CF (PH no improvement)	20
003	F	36+	longstanding ET with accommodative element		R LR advancement (hangback sutures) removal of adhesions (around IO) & granuloma	residual ET & HoT with accommodative element	0	0.02	17
004	F	18-35	residual ET with accommodative element	strabismic amblyopia	L MR recession & BT	residual ET with accommodative element	-0.2	1.06 (PH no improvement)	15

005	F	36+	infantile ET & DVD	strabismic amblyopia	R MR recession	residual ET & DVD	-0.04	1.04 (PH 0.98)	14
006	F	36+	infantile ET & DVD, then consecutive XT	strabismic & anisometropic amblyopia	L MR resection & advancement	residual ET & HoT & pseudoptosis	0.16 (PH 0.10)	0.82 (PH 0.54)	15
007	M	36+	secondary ET	traumatic LR partial avulsion	L SR & IR transposed to borders of LR (Foster sutures) & L MR BT	residual ET & HoT	0.12 (PH 0.06)	0.52 (PH 0.06)	9
008	M	36+	secondary XT	keratoconus (unilat CL & gls)	L MR resection & LR recession	Residual XT	0.06	0.66 (CL)	10
009	M	36+	constant XT & ADDuction limitation	strabismic amblyopia	R MR advancement	orthotropia	0.14 (PH 0.04)	CF (PH no improvement)	17
010	M	18-35	infantile ET, then consecutive XT		L MR resection & advancement	Residual consecutive XT & HoT	0.04 (PH 0.00)	0.14 (PH 0.04)	9
011	F	18-35	longstanding residual ET & HT, A pattern	strabismic amblyopia	R SO tendon spacer (7mm) & L SR recession	residual ET	0.08 (PH 0.00)	0.26 (PH no improvement)	4.5

012	M	18-35	infantile ET, then consecutive XT	strabismic & anisometropic amblyopia	L MR advancement	consecutive ET	-0.16	0.98 (PH no improvement)	5
013	M	36+	secondary XT	traumatic aphakia	L LR recession & L MR resection	consecutive ET	0	1.80 (PH 0.50) (CL 0.20)	12

F = female, M = male, L = left, R = right

SR = superior rectus, IR = inferior rectus, MR = medial rectus, LR = lateral rectus, IO = inferior oblique, SO = superior oblique

BT = Botulinum toxin, CL = contact lens, gls = glasses, CF = count fingers, PH = pinhole

ET = esotropia, XT = exotropia, HT = hypertropia, HoT = hypotropia, DVD = dissociated vertical deviation, 3rd NP = 3rd nerve palsy

Appendix I. Selecting and refining the quantitative methods

Appendix I.1 Stage one – combining the findings from the qualitative interviews and the results from the literature review

Within each theme (vision, task performance, physical symptoms and confidence and emotions) participants described a number of outcomes following strabismus surgery (Chapter 6). These are displayed in Table I-1 alongside the outcomes (from the interviews) and outcome measures reported in the literature (following strabismus surgery undertaken for psychosocial reasons) (Chapter 3). Possible measures or clinical tests to measure each of these outcomes are presented, in addition to other factors considered for each of the possible outcome measures.

Table I-1. Strabismus surgery outcomes described in the qualitative interviews and in the literature evidence, and how these may be measured.

Theme	Outcome described postoperatively - interview finding	Outcome measurement from literature review	Possible outcome measure	Other associated factors considered
Vision	Improved focussing		Accommodation	Monocular and binocular measurements
	Clearer vision	Binocular summation	Visual acuity Other visual function measures Binocular summation Low contrast visual acuity Contrast sensitivity	Viewing conditions: natural viewing during a specific task in 'busy' environments Monocular and binocular measurements Concentration Work performance
	Central, straighter vision		Visual field both eyes open Vernier acuity	

			Judgements about object position or alignment	
	Improved ability to judge object position or alignment	Unexpected BSV	Binocular single vision tests Judgement of object depth or position Sequential stereopsis Judgement of alignment Vernier acuity	
	Improved ability to take up fixation with the strabismic eye		Cover test Objective measurement of fixation ability	
	Increased peripheral vision	Increased peripheral field of vision	Visual field both eyes open (Goldmann perimeter or other perimeter) Navigation Mobility	Making head movements to see to the side
	Improved eye movements	Eye movements - saccade & vergence accuracy & velocity	Visual field both eyes open Uniocular field of fixation Ocular motility Sequential stereopsis Measurement of ductions and versions Objective measurement of eye movements	Eye movements: excursions, smooth pursuit, saccades, fixation, optokinetic response, or other eye movements
	Using the eyes together as a pair	Unexpected BSV	Binocular single vision tests Alternative measures of gross binocular single vision Binocular summation Accommodation Sequential stereopsis Contribution of the apparently suppressed eye	
	Using the strabismic eye more - contribution of the apparently suppressed eye	Contribution of the suppressed eye to both eyes open viewing • planning of saccades • visual field both eyes open Binocular summation	Binocular single vision tests Other test to specifically measure the contribution of the apparently suppressed eye • saccade task using distractors • visual field both eyes open using coloured filters & stimuli • other	

	Improved coordination of the eyes moving together		Binocular single vision tests Binocular summation Ocular motility Objective measurement of eye movements	
	Improved control of strabismus		Cover test Binocular single vision tests Objective scoring of control Subjective reporting of control	
	Strabismus causing less confusion, less distraction or getting in the way less	Binocular summation	Subjective perception of confusion, distraction, or strabismus getting in the way Visual acuity Binocular single vision tests Binocular summation	Other visual measures Leading to blurred vision & eye strain
	Improved vision in strabismic eye		Uniocular measures Visual acuity Other visual function measures	
	Needing to close the strabismic eye less	Binocular summation	Binocular summation Subjective reporting of eye closure Objective scoring of eye closure during a task	
	Perception that one eye no longer works much harder than the other eye	Binocular summation	Binocular summation Subjective reporting of perception of one eye working harder than the other eye	
	Change in colour vision		Colour vision test Subjective reporting of colour vision	
	Improved vision or task performance in poorer lighting conditions		Performing a visual or task-based measure under different lighting conditions	Task performance
Task performance	Improved driving		On-the-road driving Driving simulator task Driving proxy task	Vision measures: Visual acuity Other visual task Accommodation Binocular summation Visual field both eyes open
	Improved ability at work	Task performance	Eye hand coordination task Physical task or skill	Self-reports of: Concentration

		• AS-20 functional subscale • Ability to perform daily activities – no test or measure specified (reported during telephone interviews)	Visual acuity Other visual task Accommodation Computer and / or screen-based task	Work ability The need for rest breaks Confidence
	Improved near task performance	Task performance • AS-20 functional subscale • Ability to perform daily activities – no test or measure specified (reported during telephone interviews)	Near task performance Reading Eye hand coordination task Physical task or skill Scoring or measure of mistakes / accuracy	Vision measures: Accommodation Binocular single vision Binocular summation Judgement of the position of objects Sequential stereopsis Measurement of the contribution of the strabismic eye Visual field both eyes open Self-reports of: Concentration
	Improved balance		Balance Mobility	Self-reports of: Balance Mobility Clumsiness Vision measures: Accommodation Binocular single vision Binocular summation Cover test
	Improved judgements about depth and object position	Unexpected BSV Task performance • AS-20 functional subscale • Ability to perform daily activities – no test or measure specified (reported	Binocular single vision Binocular summation Sequential stereopsis Judgements about depth Judgements about object position	

		during telephone interviews)		
Physical symptoms	Improved (or worsened) physical symptoms		Self-reports of: Pain Discomfort Tiredness – eyes feeling tired Eye strain Eye feeling tight Pulling sensation Feeling of the eyes Feeling the strabismus Headaches Eye watering Eye sensitivity Schirmer's test	Eye alignment: Cover test Prism and cover test People noticing strabismus Vision measures Work ability: Time off work (sick leave) Rest breaks at work Task performance: Near task performance Reading Eye hand coordination task Physical task or skill Concentration
	Improved eye closure signs (closing the strabismic eye less)		Eye closure frequency or duration during a specific activity or task Self-report of: Eye closure	
Confidence and emotions	Improvements in self-perception (described as experiencing more positive emotions or less negative emotions)	HRQoL – AS-20 psychosocial subscale	Self-reports of: Self-confidence Feeling better about themselves Happiness Stress Anxiety Worry Fitting in Depression Being treated differently Feeling frustrated Having photographs taken	Overall improvements in their lives Feeling able to do things Socialising

	Improvements in ability to interact with others	HRQoL – AS-20 psychosocial subscale	Self-reports of: Confidence in social situations Confidence at work Interactions and face-to-face communication with people Being able to make eye contact with people Going out and socialising Stress Receiving negative comments and bullying	Work opportunities Work ability
	Improvements in confidence in own abilities	HRQoL – AS-20 psychosocial subscale	Self-reports of: Confidence in ability to perform tasks Confidence in work ability Confidence to try new activities Confidence to put themselves forward for opportunities	Task performance Driving Work ability
	Improved confidence in eyes and vision		Self-reports of: Confidence in vision Confidence in their eyes	Task performance Driving Trying new activities Confidence in their abilities Physical symptoms

Appendix I.2 Stage one – prioritising the quantitative measures

A significant number of measures or clinical tests arose as possible quantitative measures for phase two of the study (Table I-1). Due to the time required to perform all of these measures, not all could be included in the quantitative phase two of the study. The list of possible measures was then combined into a single list (Table I-2).

Some of the different aspects described during the interviews were considered to be best measured by a subjective report or scoring system. Within the literature, the AS-20 questionnaire had already been used as a measure of functional aspects of HRQoL (Alam et al, 2014, Hatt et al, 2010; Hatt et al, 2012; Koc et al, 2013; Liebermann et al, 2014). Subjectively reporting whether a particular symptom was occurring or not, or a grading of how frequently a symptom was occurring was considered as a more practical way to gather quantitative information about that symptom. The single list of possible quantitative measures was then reviewed to consider whether each measure could be measured by an objective test, a subjective report, or both. The ultimate aim was to avoid duplication of measures and excessive testing burden for participants in the quantitative phase.

To prioritise the list of possible quantitative measures further, all the original interview transcripts and the notes made during the interviews were reread thoroughly. The number of times each clinical test or measure could have been used to capture the information from each of the interview transcripts was counted. The results of the combined list of possible quantitative measures are shown in Table I-2. Included in the table are the themes relating to each measure, whether each measure could be captured by an objective measure, a subjective measure, or both, and the number of times each measure was mentioned during the qualitative interviews.

Table I-2. Summary of the measures considered for the quantitative phase during stage one

Objective clinical tests or measure considered	Theme that would be assessed	Objective measure	Subjective self-report	Number of times each 'measure' could have

		Yes (*) No (blank)	Yes (*) No (blank)	been used – referring to the interviews
Binocular summation of contrast	Vision	*		57
Visual field both eyes open (peripheral vision)	Vision	*		40
Visual acuity: Near and distance (100% contrast visual acuity chart) Low contrast visual acuity / contrast sensitivity	Vision	*		39
Focussing / accommodation: objective / subjective	Vision	*		30
Task performance – physical task / skill / activity of daily living / near task / eye hand coordination / under different lighting conditions / making mistakes	Task performance	*		18
Computer / screen based activity / task	Vision Task performance	*		18
Detection and measurement of strabismus: Cover Test, Prism cover test / Synoptophore (measurements in primary, secondary and tertiary positions of gaze)	Vision	*		15
Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)	Vision Task performance	*		15
Driving – on the road / simulator / proxy task	Vision Task performance	*		10
Balance / dizziness	Task performance	*	*	10
Control of strabismus – objective score	Vision	*	*	8

Concentration	Vision Task performance	*	*	8
Eye closure – closing strabismic eye	Vision Task performance	*	*	8
Eye movement recordings (excursions / smooth pursuit / saccades / fixation / optokinetic response / other eye movements)	Vision	*		7
Judgement of object depth / sequential stereopsis	Vision	*		6
Ocular motility	Vision	*		6
Reading – rate / speed / accuracy	Vision Task performance	*		6
Mobility / clumsiness	Task performance	*	*	6
Binocular single vision: Sensory fusion / motor fusion / stereopsis	Vision	*		5
Objective score of fixation: Ability to take up fixation with strabismic eye / swapping fixation / alternation	Vision	*		4
Other visual function measure: Under natural viewing conditions / in a busy environment / during a task / under different lighting conditions	Vision	*	*	3
Judgement of object position	Vision	*		3
Judgement of object alignment / straightness / vernier acuity	Vision	*		3
Eyes watering – Schirmer’s test	Physical symptoms	*	*	3
Uniocular field of fixation	Vision	*		2

Colour vision	Vision	*	*	2
Measurement of ductions and versions	Vision	*		1
Subjective self-report considered	Theme that would be assessed	Objective measure Yes (*) No (blank)	Subjective self-report Yes (*) No (blank)	Number of times each 'measure' could have been used – referring to the interviews
Interactions / communication with people / talking to people / looking at people / face-to-face communication / making eye contact	Confidence		*	29
Self-confidence / self-conscious / feeling better about self / self-perception	Confidence		*	28
Headaches	Physical symptoms		*	21
Socialising / going out / social situations / fitting in socially / avoiding social situations	Confidence		*	18
People noticing strabismus / hiding strabismus from people / embarrassed about appearance of strabismus	Confidence		*	17
Pain	Physical symptoms		*	15
Confidence (general)	Confidence		*	12
Balance / dizziness	Task performance	*	*	10
Discomfort / comfort	Physical symptoms		*	9
Eye strain	Physical symptoms		*	9
Pulling sensation around the eyes	Physical symptoms		*	9
Stress	Confidence		*	9
Control of strabismus – objective score	Vision	*	*	8

Concentration	Vision Task performance	*	*	8
Eye closure – closing strabismic eye	Vision Task performance	*	*	8
Rest breaks: Need for rest breaks / number of breaks	Task performance Physical symptoms		*	7
Eyes feeling tired / feeling worse when tired	Physical symptoms		*	7
Feeling of the eyes / feeling of the strabismus	Physical symptoms		*	7
Anxiety	Confidence		*	7
Worry	Confidence		*	7
Feeling more able to do things / doing more things / avoid things / trying new things / trying new activities	Confidence		*	7
Mobility / clumsiness	Task performance	*	*	6
Confidence in vision / eyes	Confidence Vision		*	6
Work ability / feeling better able to do job / working harder / confidence in work ability	Confidence Task performance Vision		*	5
Being able to have a photograph taken	Confidence		*	5
Strabismus causing confusion	Vision		*	4
Strabismus being a distraction	Vision		*	4

Opportunities / confidence to put oneself forward for opportunities / work opportunities / being held back by strabismus	Confidence Task performance		*	4
Feeling frustrated	Confidence		*	4
Happiness	Confidence		*	4
Receiving negative comments / bullying	Confidence		*	4
Other visual function measure: Under natural viewing conditions / in a busy environment / during a task / under different lighting conditions	Vision	*	*	3
Eyes watering – Schirmer's test	Physical symptoms	*	*	3
Time off work sick	Task performance Physical symptoms		*	3
Eye feeling tight / tight feeling around the eyes	Physical symptoms		*	3
Confidence: Social situations / talking to people	Confidence		*	3
Colour vision	Vision	*	*	2
Strabismus getting in the way	Vision		*	2
Tiredness	Physical symptoms		*	2
Confidence in ability to perform tasks / do things	Confidence Task performance Vision		*	2
Being treated differently	Confidence		*	2
Depression	Confidence		*	2
Eye sensitivity	Physical symptoms		*	1

Appendix I.3 Stage two – core measures to be extracted from the clinical orthoptic report

Following discussion with supervisors and the advisory group, it was decided that 'standard' clinical tests that would be performed at a clinical orthoptic appointment would be extracted from the clinical notes, rather than repeated during a study visit. These clinical tests included: visual acuity (near and distance); cover test (near and distance), from which fixation could be scored; binocular single vision tests using standard clinical tests; ocular motility, with grading of ductions and versions in each of the nine positions of gaze using a standard clinical 9 point scale (-4 to 0 to +4) graded with respect to the midline; prism cover test, performed at near and distance in primary position and performed in secondary positions of gaze for distance fixation; see Table I-3.

Table I-3. Core measures – to be extracted from the clinical orthoptic report	
Visual acuity	Near and distance
Cover test	Near and Distance Scoring of fixation from cover test observation
Investigation of Binocular Single Vision	Near and Distance Sensory fusion (Bagolini glasses) Motor fusion (Prism fusion range, with Bagolini glasses control if required) Stereopsis (Near: Frisby near sterotest, TNO stereotest, Wirt stereotest, and Distance: FD2)
Ocular motility	Grading of ductions and versions 9 point scale, with respect to the midline (-4 to 0 to +4) 9 positions of gaze
Prism cover test	Near and Distance (primary position) Secondary positions of gaze (distance fixation)

Appendix I.4 Stage two – refining the objective measures

The list of objective measures considered for the quantitative phase (first part of Table I-2) were considered and discussed with supervisors and the advisory group. The measures considered 'core measures' that were to be extracted from the clinical orthoptic report (Table I-3) were removed (highlighted in orange in Table I-4). Measures considered more appropriate for measurement by subjective self-report were also removed (highlighted in blue in Table I-4). These included colour vision, eye watering, mobility or clumsiness, closure of the strabismic eye, concentration, and scoring or grading of perceived control of the strabismus.

Table I-4. Objective measures considered for the quantitative phase – stage two

Objective clinical tests or measures	Theme	Objective measure	Subjective self-report	Count
Binocular summation of contrast	Vision	*		57
Visual field both eyes open (peripheral vision)	Vision	*		40
Visual acuity: Near and distance (100% contrast visual acuity chart)	Vision	*		39
Low contrast visual acuity / contrast sensitivity				
Focussing / accommodation: objective / subjective	Vision	*		30
Task performance – physical task / skill / activity of daily living / near task / eye hand coordination / under different lighting conditions / making mistakes	Task performance	*		18
Computer / screen based activity / task	Vision Task performance	*		18
Detection and documentation of strabismus: Cover Test Prism cover test / Synoptophore (measurements in primary, secondary and tertiary positions of gaze)	Vision	*		15
Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)	Vision Task performance	*		15
Driving – on the road / simulator / proxy task	Vision Task performance	*		10
Balance / dizziness	Task performance	*	*	10
Control of strabismus – objective score	Vision	*	*	8

Concentration	Vision Task performance	*	*	8
Eye closure – closing strabismic eye	Vision Task performance	*	*	8
Eye movement recordings (excursions / smooth pursuit / saccades / fixation / optokinetic response / other eye movements)	Vision	*		7
Judgement of object depth / sequential stereopsis	Vision	*		6
Ocular motility	Vision	*		6
Reading – rate / speed / accuracy	Vision Task performance	*		6
Mobility / clumsiness	Task performance	*	*	6
Binocular single vision: Sensory fusion / motor fusion / stereopsis	Vision	*		5
Objective score of fixation: Ability to take up fixation with strabismic eye / swapping fixation / alternation	Vision	*		4
Other visual function measure: Under natural viewing conditions / in a busy environment / during a task / under different lighting conditions	Vision	*	*	3
Judgement of object position	Vision	*		3
Judgement of object alignment / straightness / vernier acuity	Vision	*		3
Eyes watering – Schirmer's test	Physical symptoms	*	*	3
Uniocular field of fixation	Vision	*		2
Colour vision	Vision	*	*	2
Measurement of ductions and versions	Vision	*		1

Orange – removed from consideration of objective measures, included in the standard clinical Orthoptic report.

Blue – removed from consideration of objective measures, included in subjective self-report

The remaining measures suitable for inclusion in the quantitative phase (those not highlighted in orange or blue in Table I-4) were discussed, with supervisors and the advisory group, with particular emphasis on avoiding duplication of measures and avoiding testing burden for participants. A refined selection of possible objective tests or measures was made and are shown in Table I-5. A summary of the discussion and decision making for each test is shown at the bottom of the table.

Table I-5. The refined list of objective measures for the quantitative phase – stage two		
Objective clinical tests or measures considered	Objective clinical tests or measures selected	Theme
• Binocular summation of contrast • Low contrast visual acuity • Contrast sensitivity	• Binocular summation of contrast • Contrast sensitivity – both eyes open	Vision (a)
• Visual field both eyes open • Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)	• Visual field both - eyes open and monocularly (with strabismic eye covered)	Vision (b)
• Focussing / accommodation	• Objective measure of accommodation – both eyes open	Vision (c)
• Judgement of object depth / sequential stereopsis • Judgement of object position • Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)	• Sequential stereopsis: - Both eyes open and monocularly (with strabismic eye covered) - Different stimuli • Gross binocular single vision test / judgement of object depth • Prehension test	Vision Task performance (d)
• Task performance – physical task / skill / activity of daily living / near task / eye hand coordination / under different lighting conditions / making mistakes • Computer / screen-based activity / task • Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or	• Task performance: - Both eyes open - Performed under different lighting conditions - Using a screen based test - Using a physical task or skill - Using a real world task	Task performance Vision (e)

awareness / perception that this doesn't happen)		
• Balance / dizziness • Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)	• Balance measure - Both eyes open and monocularly (with strabismic eye covered)	Task performance (f)
• Eye movement recordings (excursions / smooth pursuit / saccades / fixation / optokinetic response / other eye movements) • Other visual function measure: Under natural viewing conditions / in a busy environment / during a task / under different lighting conditions • Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen) • Computer / screen based activity / task	• Objective eye movement recordings - Both eyes open - Fixation - Smooth pursuit - Saccades - Horizontal and vertical eye movments - Backgrounds of different complexity - Screen presentation of stimuli and backgrounds	Vision (g)
Tests eliminated		
Uniocular field of fixation		Vision (h)
Judgement of object alignment / straightness / vernier acuity	Consider again as a subjective self-report	Vision (i)
Reading – rate / speed / accuracy	Consider whether could be covered by questionnaire or other measure of visual function	Vision (j)
Driving – on the road / simulator / proxy task	Consider whether could be covered by questionnaire or other measure of visual function	Vision Task performance (k)

(a) Binocular summation of contrast, low contrast visual acuity and contrast sensitivity testing considered to have overlap. Binocular summation of contrast and contrast sensitivity selected. Contrast sensitivity was selected to be performed both eyes open.
(b) Visual field both eyes open was selected and was not considered to have significant overlap with any of the other possible measures. Additionally, performing a visual field test with the strabismic eye covered would allow quantification of the contribution of the strabismic eye to peripheral vision and the visual field.
(c) Measures of focussing and accommodation, including objective and subjective measures, were refined to an objective measure of accommodation both eyes open.
(d) Judgements about object depth and object position, as well as the contributions of the suppressed eye to both eyes open viewing were selected to be measured by testing sequential stereopsis, using different stimuli. Additionally, measures of gross binocular vision involving judgements of the depth and position of objects, and measures of prehension, involving reaching to grasp real world objects were selected.
(e) Task performance measures were selected with the need to incorporate tasks that were screen based, involved a physical skills or task and used a real-world task or skill. These were selected to be performed both eyes open, with a preference for a task that could be performed under different lighting conditions
(f) Measures of balance and dizziness were refined to an objective measure of balance that could incorporate a measurement with both eyes open, as well as monocularly to investigate the contribution of the suppressed eye to balance.
(g) Eye movement recordings were selected with a preference for detailed high-quality recordings that could be performed both eyes open. Recordings and measurements of fixation, smooth pursuit and saccades performed with different backgrounds presented on a screen were suggested to be the preferred method for measuring both eye movements and visual function in different viewing environments.
(h) It was decided to remove a measurement of uniocular field of fixation with each eye. This was mentioned less frequently and there was a potential overlap with ocular motility testing, including the grading of ductions and versions in all 9 positions of gaze (core measurements extracted from the clinical orthoptic report).
(i) The judgement of object alignment, straightness or vernier acuity suggested tests were mentioned a small number of times and were moved to the list of possible subjective self-report tests.
(j) A test of reading was considered better captured using a questionnaire-based assessment of function or as a patient reported outcome measure.
(k) A test of driving ability was considered better captured using a questionnaire-based assessment of function or as a patient reported outcome measure. Driving was additionally considered time consuming aspect to investigate and measure, even in a simulation. Whilst a proxy task may have been appropriate to measure driving ability, it was felt a self-report of driving ability may be sufficient considering the scope of the study and considering not all participants may drive

Appendix I.5 Stage three – suitable PROMs

A systematic review of PROMs in amblyopia and strabismus was published in 2018 (Kumaran et al., 2018). Seventy-one PROMs were identified for inclusion in the systematic review. For the purposes of this study, only the PROMs suitable for use in adults with strabismus were considered. Eliminated from the list of PROMs identified (Kumaran et al., 2018) were those that were suitable for children only, specific only to amblyopia, psychological measures, behavioural inventories, and instruments used only to measure either beliefs and cognition, social support, appearance related distress and perceptions, functional measures of activities of daily living and utilities to measure the value of health to the patient. Eliminated from the list was also those that were strabismus-specific but were not relevant to the patient group in this study as they were specific to intermittent exotropia (Hatt et al., 2010b; McKeon et al., 1997), exotropia only (Ha & Kim, 2016), the effect of diplopia following strabismus surgery (Lin et al., 2015) or the effect of post strabismus surgery symptoms (Ryu et al., 2015). The remaining PROMs (n=17) were either generic (n=2), vision specific (n=2), strabismus specific (n=10), adapted from general instruments to include strabismus (n=1) or specific to amblyopia and strabismus (n=2).

The generic instruments included questions about general well-being and quality of life. These included the medical outcomes Short Form (SF) health surveys, such as the SF-36, SF-20, SF-12 and SF-8 (Ware & Sherbourne, 1992), and the WHOQOL-BREF (Whoqol Group, 1998). The vision specific instruments included questions to measure the functional impact of vision or more specifically reduced vision, including the widely used National Eye Institute Visual Function Questionnaire (VFQ-25) (Mangione et al., 1998) and the Visual Function 14 (VF-14) (Steinberg et al., 1994).

The strabismus specific instruments included a survey developed to capture the psychosocial aspects of living with strabismus (Satterfield et al., 1993). The Adult Strabismus-20 (AS-20) (Hatt et al., 2009b) was the most commonly used PROM to evaluate adult strabismus (Kumaran et al., 2018) and was typically completed pre-operatively and post-operatively to measure change in HRQoL following strabismus surgery. In a study of Chinese adolescents and adults with strabismus, a questionnaire with different pre and post-operative components was developed specifically to include aspects considered culturally important to Chinese individuals (Xu et al., 2012). Some strabismus specific instruments were used or developed specifically for pre or post-operative assessment of the strabismus. The Expectations of Strabismus Surgery Questionnaire (ESSQ) (McBain et al., 2016a) and

Likert scales to report the visibility of the strabismus perceived by the patient (McBain et al., 2014b) were used pre-operatively. An instrument used to evaluate satisfaction with the outcome of strabismus surgery was used post-operatively only (Satterfield et al., 1993). Other strabismus specific instruments were developed for post-operative use but included questions about the effect of strabismus both pre and post-operatively. Pre-operative questions therefore required the patient to recall their pre-operative state retrospectively and after receiving strabismus surgery. These include a Disability Questionnaire (Beauchamp et al., 2005a), a self-report grid of different personality traits (Burke et al., 1997), a psychosocial difficulties questionnaire (Nelson et al., 2008), and a Perspectives Questionnaire completed by both patient and Ophthalmologist (Beauchamp et al., 2005b).

Other instruments to measure the impact of strabismus on QoL included asking questions adapted from another study with a Visual Analogue Scale (VAS) for patients to report their response to questions pre and post-operatively (Jackson et al., 2013). The Psychological Impact Questionnaire (PIQ) was developed for use in teenagers with amblyopia, strabismus, or glasses, or who had previously received amblyopia treatment (patching) (Sabri et al., 2006), yet it has been used in a study of older teenagers and adults with strabismus (Ritchie et al., 2013). The Amblyopia and Strabismus Questionnaire (A&SQ) was developed to measure the impact of amblyopia and/or strabismus on QoL (van de Graaf et al., 2004) and it is a commonly used instrument to measure the impact of amblyopia and strabismus on QoL (Kumaran et al., 2018).

An additional search of the literature to identify any sources published since (Kumaran et al., 2018). Using search terms ‘questionnaire’ OR ‘instrument’ OR ‘scale’ OR ‘checklist’ OR ‘patient reported outcome measure’ AND ‘strabismus’ AND ‘adult’. Using filters ‘human’, ‘published in English’ and ‘published in the last 5 years’. One hundred and six sources were identified (search date 25/2/20). No additional PROMs or QoL instruments were identified that had not been included in the systematic review (Kumaran et al., 2018).

The AS-20 and VFQ-25 were selected as the most appropriate measures for the quantitative phase of the study.

Appendix I.6 Stage four – coverage of the AS-20 and VFQ-25

The coverage of both the AS-20 and VFQ-25 questionnaires was reviewed. The content of each instrument was compared to the findings from the qualitative interviews (chapter 6) and the measures planned for the quantitative phase and the findings are shown in Table I-6.

Table I-6 Coverage of the AS-20 and VFQ-25 questionnaires compared to the findings from the qualitative interviews in phase one

Core measures – to be extracted from clinical Orthoptic report		
Clinical test		**Covered by AS-20 or VFQ-25**
Visual acuity	Near and distance	VFQ-25 Q2, Q5, Q6, Q7, Q8, Q11, Q13, Q14
Cover test	Near and Distance Scoring of fixation from cover test observation	
Investigation of Binocular Single Vision	Near and Distance Sensory fusion (Bagolini glasses) Motor fusion (Prism fusion range, with Bagolini glasses control if required) Stereopsis (Near: Frisby near stereotest, TNO stereotest, Wirt stereotest, and Distance: FD2)	AS-20 Q14
Ocular motility	Grading of ductions and versions 9 point scale, with respect to the midline (-4 to 0 to +4) 9 positions of gaze	
Prism cover test	Near and Distance (primary position) Secondary positions of gaze (distance fixation)	
Measures to be included in the study visit		
Study measure		**Covered by AS-20 or VFQ-25**
Binocular summation of contrast		
Contrast sensitivity	Both eyes open	
Visual field both eyes open	Both eyes open and monocularly (with strabismic eye covered)	VFQ-25 Q9, Q10

Contribution of the suppressed eye to both eyes open viewing / feeling of using both eyes / feeling of the eyes working together (or awareness / perception that this doesn't happen)		
Objective measure of accommodation	Both eyes open	
Sequential stereopsis:	Both eyes open and monocularly (with strabismic eye covered) Using different stimuli	
Gross binocular single vision test / judgement of object depth		AS-20 Q14
Prehension test		
Task performance	Both eyes open Performed under different lighting conditions Using a screen based test Using a physical task or skill Using a real world task	VFQ-25 Q6, Q7, Q9
Balance measure	Both eyes open and monocularly (with strabismic eye covered)	
Objective eye movement recordings	Both eyes open Fixation / Smooth pursuit / Saccades Horizontal and vertical eye movements Backgrounds of different complexity Screen presentation of stimuli and backgrounds	
Subjective self-report measures		
Subjective self-report topics		Covered by AS-20 or VFQ-25
Interactions / communication with people / talking to people / looking at people / face-to-face communication / making eye contact		AS-20 Q1, Q2, Q3, Q4, Q7, Q9, Q10 VFQ-25 Q11
Self-confidence / self-conscious / feeling better about self / self-perception		AS-20 Q1, Q2, Q3, Q6, Q8
Headaches		

Socialising / going out / social situations / fitting in socially / avoiding social situations	AS-20 Q1, Q2, Q3, Q4, Q6, Q7, Q9, Q10, Q13, Q14 VFQ-25 Q20
People noticing strabismus / hiding strabismus from people / embarrassed about appearance of strabismus	AS-20 Q1, Q2, Q3, Q4, Q9, Q10
Pain	VFQ-25 Q4, Q19
Confidence (general)	
Balance / dizziness	
Driving	VFQ-25 Q15, Q16
Discomfort / comfort	VFQ-25 Q4, Q19
Eye strain	AS-20 Q15
Pulling sensation around the eyes	
Stress	AS-20 Q17
Control of strabismus – objective score	
Concentration	AS-20 Q13
Eye closure – closing strabismic eye	AS-20 Q11
Rest breaks: Need for rest breaks / number of breaks	AS-20 Q13, Q20 VFQ-25 Q18
Eyes feeling tired / feeling worse when tired	
Feeling of the eyes / feeling of the strabismus	
Anxiety	
Worry	AS-20 Q1, Q18 VFQ-25 Q3
Feeling more able to do things / doing more things / avoid things / trying new things / trying new activities	AS-20 Q19 VFQ-25 Q13, Q14
Mobility / clumsiness	VFQ-25 Q9
Confidence in vision / eyes	
Reading	AS-20 Q12, Q16, Q20 VFQ-25 Q5
Work ability / feeling better able to do job / working harder / confidence in work ability	VFQ-25 Q18
Being able to have a photograph taken	
Strabismus causing confusion	
Strabismus being a distraction	
Opportunities / confidence to put oneself forward for opportunities / work opportunities / being held back by strabismus	AS-20 Q5
Feeling frustrated	VFQ-25 Q21
Happiness	

Receiving negative comments / bullying	
Other visual function measure: Under natural viewing conditions / in a busy environment / during a task / under different lighting conditions	VFQ-25 Q9
Eyes watering	
Time off work sick	
Eye feeling tight / tight feeling around the eyes	
Confidence: Social situations / talking to people	AS-20 Q10
Judgement of object alignment / straightness / vernier acuity	
Colour vision	VFQ-25 Q12
Strabismus getting in the way	AS-20 Q11
Tiredness	
Confidence in ability to perform tasks / do things	
Being treated differently	AS-20 Q5, Q7, Q9
Depression	
Eye sensitivity	

The subjective self-report topics identified (Table I-6) were taken forward and developed into additional study questions. The additional study questions were then compared to the AS-20 and VFQ-25 questions, to identify topics already covered by questions in these PROMs. These findings are shown in Table I-7.

Table I-7 The development of additional study questions for each of the subjective self-report topics and the coverage of these topics by existing questions in the AS-20 and VFQ-25			
Theme	Subjective self-report topic	Possible question wording	Coverage by other questions
Vision	Control of strabismus	I can control my eye turning I can make my eye position straighter I can adjust my head position to control my eye turning My eyes drift / turn	No
	Blurred when manifest / eye drifting and causing blur Experiencing periods of blurred vision Vision blurry / confusion when manifest	I get blurred vision because of my eyes I get blurred vision When my eye turns it causes blurred vision	Partial VFQ-25 difficulty with activities VFQ-25 vision subscales multiple questions related to eyesight, but not related to becoming manifest
	Focussing - distance tasks	I have problems focussing on things far away	Yes VFQ-25 subscale - dist vision A6-A8 A6 Because of your eyesight, how much difficulty do you have recognizing people you know from across a room? A7 Because of your eyesight, how much difficulty do you have taking part in active sports or other outdoor activities that you enjoy (like golf, bowling, jogging, or walking)? A8 Because of your eyesight, how much difficulty do you have seeing and enjoying programs on TV?
	Focussing - near tasks	I have problems focussing on things close to me	Partial VFQ-25 subscale - nr vision A3-A5 not focussing, just nr vision
	Feeling like eyes focus together	My eyes focus together	No
	Confusion	My eyes cause me to get confused I am confused by my vision I get confused by my eyes	No
	Frustration	My eyes cause me to get frustrated	Partial

		I am frustrated by my vision I get frustrated by my eyes	VFQ-25 Q21. I feel frustrated a lot of the time because of my eyesight Question specific to eyesight only
	Distraction	My eyes cause me to get distracted when I am doing tasks / activities I am distracted by my eyes I get distracted by my eyes I have one eye that gets in the way of my vision	Partial VFQ-25 questions about 'difficulty' with vision and activities, not 'distraction' or one eye getting in the way
	See things better Easier to see Vision feels more natural Things looking sharper Things looking clearer Comfortable vision	I have problems seeing things clearly I struggle to see things clearly I have problems with the sharpness of my vision My vision is clear My vision feels natural My vision feels comfortable	Yes VFQ-25 questions about difficulties with eyesight and vision subscales
	Taking up fixation with strabismic eye	I swap which eye I am looking through I swap to using my other eye I have problems / it is difficult to swap and use my other eye I find it easy to swap to using my other eye I have difficulty swapping to look with my other eye	No
	Vision being central	I have problems with my vision not being central	No
	Vision feeling straighter	I have problems with my vision not being straight	No
	Looking around / finding it difficult to look around Easier to look around	I find it difficult to look around I have difficulty looking around	No
	Moving eyes around - feeling easier	I have problems moving my eyes around	No

	Feeling like one eye works harder than the other Using fixing eye a lot more than strabismic eye One eye working harder than the other	I have one eye that works much harder than the other eye One of my eyes works much harder than the other eye	No
	Looking through two eyes at the same time Using both eyes	I am able to look through both my eyes at the same time I look through both my eyes at the same time I can use both my eyes at the same time	No
	Feeling like eyes work together	My eyes work together	No
	Eyes coordinating	My eyes coordinate together	No
	Eye movement	My eyes move together	No
	Time to see	It takes me a long time to see / focus on something	Partial VFQ-25 questions about vision and difficulties with vision, not specifically about time to see
	Colour	I have problems seeing colour	Partial VFQ-25 Q12 Because of your eyesight, how much difficulty do you have picking out and matching your own clothes?
	Contrast	I have problems seeing different shades of black, grey and white	No
	Vision in strabismic eye	I have one eye that sees worse than the other I have one eye that sees much worse than the other	No
	Feeling like can see depth / depth perception	I have problems with depth perception	Yes AS-20 Q14 I have problems with depth perception
	3D	I see things in 3D	Yes AS-20 Q14 I have problems with depth perception
	Having more vision at the side / peripheral vision / Having less	I have difficulty seeing things at the edge of my vision / in my peripheral vision / in the outer part of my vision	Yes VFQ-25 Q10 Because of your eyesight, how much difficulty do you have noticing objects off to the side while you are walking along?

	peripheral vision (XT/ET) Feeling like not fully seeing peripheral vision Being able to see peripheral vision -		
	Having to move head to see something at the side / in periphery -	I have to move my head to see things at the edge of my vision / in my peripheral vision	Yes VFQ-25 Q10 Because of your eyesight, how much difficulty do you have noticing objects off to the side while you are walking along?
	Being cautious when things are at the side / in peripheral vision	I have to be cautious when things are at the side of me / in my peripheral vision	Yes VFQ-25 Q10 Because of your eyesight, how much difficulty do you have noticing objects off to the side while you are walking along?
	Eye closure	I close or cover one eye to see something I close or cover one eye to concentrate on something I need to close or cover one eye I close or cover one eye	Yes AS-20 Q11 I cover or close one eye to see things better
Task performance	Work - longer	My eyes affect / limit how long I work or do things for My eyes cause me problems with working for long periods My eyes cause me to work less than I would like to	Yes AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate VFQ-25 general difficulties with activities questions, not specific to work VFQ-25 Q18 Are you limited in how long you can work or do other activities because of your vision?
	Work - harder	My eyes affect / limit how hard I can work / do things for My eyes cause me problems with working hard	Partial VFQ-25 Q18 Are you limited in how long you can work or do other activities because of your vision? VFQ-25 A11b Are you limited in the kinds of things you can do because of your vision?
	Reading	I have problems reading	Yes AS-20 Q12 I avoid reading because of my eyes AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate

			AS-20 Q16 I have problems reading because of my eye condition AS-20 Q20 I need to take frequent breaks when reading because of my eyes VFQ-25 Q5 How much difficulty do you have reading ordinary print in newspapers? VFQ-25 near vision subscale A3 Wearing glasses, how much difficulty do you have reading the small print in a telephone book, on a medicine bottle, or on legal forms? A4 Because of your eyesight, how much difficulty do you have figuring out whether bills you receive are accurate? A5 Because of your eyesight, how much difficulty do you have doing things like shaving, styling your hair, or putting on makeup?
	Task performance ? split into near and distance	My eyes cause problems with performing tasks I have difficulties performing near tasks and activities I have problems performing distance tasks and activities	Partial AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate VFQ-25 part 2 - difficulty with activities Q6 - 14 VFQ-25 Part 3 - responses to vision problems Q17 - 25 VFQ-25 Subscale near vision A3-A5 VFQ-25 Subscale role limitations A11 a & b all related to ‘eyesight’ rather than task performance
	Task performance - making mistakes	I make mistakes when I’m performing tasks	Partial AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate VFQ-25 part 2 - difficulty with activities Q6 - 14 VFQ-25 Part 3 - responses to vision problems Q17 - 25 VFQ-25 Subscale near vision A3-A5

			VFQ-25 Subscale role limitations A11 a & b all related to ‘eyesight’ rather than task performance
	Task performance - quicker	I have problems doing things / performing tasks quickly	Partial VFQ-25 part 2 - difficulty with activities Q6 - 14 VFQ-25 Part 3 - responses to vision problems Q17 - 25 VFQ-25 Subscale near vision A3-A5 VFQ-25 Subscale role limitations A11 a & b all related to ‘eyesight’ rather than task performance
	Task performance - judgement of depth	When I am performing tasks I have problems judging the position of things	Partial AS-20 Q14 I have problems with depth perception VFQ-25 part 2 - difficulty with activities Q6 - 14 VFQ-25 Part 3 - responses to vision problems Q17 - 25 VFQ-25 Subscale near vision A3-A5 VFQ-25 Subscale role limitations A11 a & b all related to ‘eyesight’ rather than task performance
	Eye hand coordination	I have difficulty with eye hand coordination	No
	Using screens / looking at emails / focussing on computer or screen	I have problems using screens like computers, mobile phones, tablet devices I avoid looking at screens because of my eyes My eyes cause me problems when I’m using screens I have difficulty using screens	Partial VFQ-25 A8 Because of your eyesight, how much difficulty do you have seeing and enjoying programs on TV?
	Adjusting screen brightness Room brightness / dimness	I have problems with the brightness of screens I have problems doing things when the room is very bright I have problems doing thing when the room is very dim I have problems with the brightness of rooms or objects	No

	Concentration	I have problems concentrating on tasks / activities I have problems concentrating on things I have difficulty maintaining concentration on tasks / activities	Partial VFQ-25 part 3 - response to vision problems Q18 Are you limited in how long you can work or do other activities because of your vision? AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate
	Reading - longer	I have problems reading I have problems reading for long periods I need to take breaks when reading I avoid reading because of my eyes	Yes VFQ-25 Q5 How much difficulty do you have reading ordinary print in newspapers? VFQ-25 Q8 How much difficulty do you have reading street signs or the names of stores? VFQ-25 A3 Wearing glasses, how much difficulty do you have reading the small print in a telephone book, on a medicine bottle, or on legal forms? AS-20 Q12 I avoid reading because of my eyes AS-20 Q16 I have problems reading because of my eye condition AS-20 Q20 I need to take frequent breaks when reading because of my eyes
	Judgement of steps / mobility	I have problems getting around I have problems navigating obstacles I have problems judging steps and the position of things when I'm moving around I have problems judging steps and the position of things I have to be careful to judge steps and curbs	AS-20 Q14 I have problems with depth perception VFQ-25 Q9 Because of your eyesight, how much difficulty do you have going down steps, stairs, or curbs in dim light or at night? (but specific to poor light)
	Balance	I have difficulty with balance I have problems with balance I have to put a lot of effort into balancing	No
	Clumsiness	I can be clumsy and bump into things	No
	Dizziness	I can be dizzy I can get dizziness I experience dizziness / dizzy spells I experience problems with dizziness	No
	Experiencing / Having periods of dizziness or lack of balance	I have dizziness or balance problems because of my eyes My eyes cause dizziness or balance problems I experience balance problems	No

	Driving	I have problems driving because of my eyes My eyes cause difficulty with driving May drive / not drive Y/N I don't drive because of my eyes	Yes VFQ-25 driving questions Q15-16a If driving - if no, never drive / gave up If gave up - why If currently driving - asks about driving in daytime, night and in difficult conditions
	Doing things (general)	I avoid doing things because of my eyes My eyes mean that I can't do some things	VFQ-25 Q17 Do you accomplish less than you would like because of your vision? VFQ-25 Q20 I stay home most of the time because of my eyesight VFQ-25 Q22 I have much less control over what I do, because of my eyesight VFQ-25 A11 b Are you limited in the kinds of things you can do because of your vision? but questions are specific to vision / eyesight, rather than eyes
Physical symptoms	Headache	My eyes cause headaches I have headaches because of my eyes I get headaches	No
	Straining to control strabismus	Trying to control my eye turning causes eye strain Trying to control my eye turning causes headache	Partial AS-20 Q15 My eyes feel strained VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)?
	Eye strain	My eyes cause eye strain I have eye strain	Yes AS-20 Q15 My eyes feel strained VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)?
	Discomfort	My eyes cause discomfort I have eye discomfort My eyes are sensitive	Yes VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)?
	Pain	My eyes cause pain I have pain because of my eyes	Yes

			VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)?
	Headache / eye strain leading to taking breaks from work - or just 'eyes'	Eye strain and/or headaches cause me to take breaks from work or the things I enjoy doing My eyes cause me to take breaks from work My eyes cause me to take breaks from the things I enjoy doing I have to take breaks from work I have to take breaks from the things I enjoy doing I have to stop doing things because of my eyes	Partial AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate AS-20 Q20 I need to take frequent breaks when reading because of my eyes VFQ-25 Q17 Do you accomplish less than you would like because of your vision? VFQ-25 Q18 Are you limited in how long you can work or do other activities because of your vision? specific to 'vision', rather than 'eyes' VFQ-25 Q19 How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you'd like to be doing?
	Headache / eye strain leading to time off work - or just 'eyes'	Eye strain and/or headaches cause me to take breaks from work or the things I enjoy doing My eyes cause me to take breaks from work My eyes cause me to take breaks from the things I enjoy doing I have to take breaks from work I have to take breaks from the things I enjoy doing I have to stop doing things because of my eyes	Partial AS-20 Q13 I stop doing things because my eyes make it difficult to concentrate AS-20 Q20 I need to take frequent breaks when reading because of my eyes VFQ-25 Q17 Do you accomplish less than you would like because of your vision? VFQ-25 Q18 Are you limited in how long you can work or do other activities because of your vision? specific to 'vision', rather than 'eyes' VFQ-25 Q19 How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you'd like to be doing?
	Work - taking time off	I take time off work because of my eyes	No
	Taking painkillers	I take painkillers because of eye pain / discomfort I take painkillers because of headaches	No
	Taking rests / breaks	I need to take breaks to rest my eyes I take breaks because of my eyes At work I need to take breaks because of my eyes	Partial AS-20 Q20 I need to take frequent breaks when reading because of my eyes

		I need to take breaks because of my eyes	specific to ‘reading’ VFQ-25 Q18 Are you limited in how long you can work or do other activities because of your vision? specific to ‘vision’, rather than ‘eyes’
	Busy environments	I find it difficult to use my eyes / see things when it’s very busy I avoid busy environments / crowded places I have difficulty seeing / focussing / processing things when it’s very busy	Partial VFQ-25 Q7. Because of your eyesight, how much difficulty do you have finding something on a crowded shelf? VFQ-25 Q13. Because of your eyesight, how much difficulty do you have visiting with people in their homes, at parties, or in restaurants? VFQ-25 Q18. Are you limited in how long you can work or do other activities because of your vision? related to eyesight, rather than eyes not specific to busy environments / difficulty seeing / processing in crowded places
	Pulling sensation	I can feel my eyes pulling I feel a pulling sensation in my eyes My eyes cause an uncomfortable pulling sensation My eyes cause an uncomfortable pulling feeling I feel a pulling sensation in my eyes	Partial VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)? VFQ-25 Q19. How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you’d like to be doing?
	Tight feeling / tightness	I can feel a tightness in my eyes I get / feel a tightness in my eyes My eyes cause an uncomfortable tight feeling	Partial VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)? VFQ-25 Q19. How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you’d like to be doing?
	Stress	My eyes cause me stress	Yes AS-20 Q17) I feel stressed because of my eyes
	Watery eyes	My eyes water	No
	Avoiding looking at things (due to comfort)	I avoid looking at things because my vision isn’t comfortable I avoid looking at things because my eyes are not comfortable I avoid looking at things because of my eyes	Partial AS-20 Q12) I avoid reading because of my eyes AS-20 Q16) I have problems reading because of my eye condition

			VFQ-25 Q4. How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)? VFQ-25 Q19. How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you'd like to be doing?
	Eyes affected by tiredness / tiredness affecting the eyes	My eyes are affected by tiredness My eye position is affected by tiredness My eyes make me tired	No
	Eye feeling like it's in the right place / central / not in straight position	My eye feels like it is not straight My eye feels like it is turning It feels like there is something wrong with my eyes I can feel my eye turning	No
Confidence and emotions	Talking to people	I have problems talking to people I have problems making eye contact with people I avoid making eye contact with people I avoid face-to-face situations I avoid face-to-face situations because of my eyes Face-to-face situations are difficult because of my eyes I have problems communicating with people because of my eyes I have problems meeting new people because of my eyes I avoid looking at people because of my eyes I have problems looking at people because of my eyes	Partial AS-20 Q10) I find it hard to initiate contact with people I don't know because of my eyes AS-20 Q7) People avoid looking at me because of my eyes AS-20 Q2) I feel that people are thinking about my eyes even when they don't say anything AS-20 Q4) I wonder what people are thinking when they are looking at me because of my eyes
	People notice strabismus	Other people notice there is something different about my eyes Other people notice my eye turns Other people notice my eye position is not straight	Partial AS-20 Q3) I feel uncomfortable when people are looking at me because of my eyes AS-20 Q7) People avoid looking at me because of my eyes AS-20 Q9) People react differently to me because of my eyes

	People commenting on strabismus	I get distracted when people comment on my eye position I get distracted when people comment on my eyes Other people comment on my eyes It bothers me when other people comment on my eyes People commenting on my eyes upsets me	No
	People treating me differently	I feel people treat me differently because of my eyes	Partial AS-20 Q5) People don't give me opportunities because of my eyes AS-20 Q9) People react differently to me because of my eyes
	Bullying	I experience bullying because of my eyes I am bullied because of my eyes	No
	Work	I am treated differently at work because of my eyes My eyes cause problems with getting a job My eyes cause problems with performing my job My eyes cause problems with gaining career opportunities My eyes cause problems with gaining promotion at work I miss out on opportunities because of my eyes My eyes cause problems at work	Partial AS-20 Q5) People don't give me opportunities because of my eyes AS-20 Q9) People react differently to me because of my eyes
	Confidence - in vision	I have confidence in my eyes I have confidence in my vision I have confidence that I can see well I can rely on my eyes I can rely on my vision I trust my eyes I trust my vision	No
	Confidence - abilities	I am confident that I can do my work I have confidence in my abilities at work I have confidence in my abilities in my everyday life I am confident to take on roles and responsibilities at work I am confident to take on new things I am confident to take on new personal challenges I avoid doing things because of eyes I avoid doing things at work because of my eyes	No

		I avoid doing things in my everyday life because of my eyes My eyes cause problems with confidence in my abilities	
	Confidence - general	I am confident I am generally confident I have self confidence I have confidence in myself I have problems with my confidence because of my eyes My eyes cause problems with my self confidence My eyes hold me back	Partial AS-20 Q8) I feel inferior to others because of my eyes
	Self-perception	I think about my eyes	Partial AS-20 Q6) I am self-conscious about my eyes AS-20 Q8) I feel inferior to others because of my eyes
	Appearance	I dislike the ways my eyes look The way my eyes look upsets me I dislike the appearance of my eyes I try to hide my eyes I try to cover my eyes	Partial AS-20 Q3) I feel uncomfortable when people are looking at me because of my eyes
	Socialising	I avoid socialising because of my eyes I avoid social situations because of my eyes I avoid having photographs taken because of my eyes I am unconfident in social situations	Partial AS-20 Q19) I can't enjoy my hobbies because of my eyes VFQ-25 Q11. Because of your eyesight, how much difficulty do you have seeing how people react to things you say? VFQ-25 Q13. Because of your eyesight, how much difficulty do you have visiting with people in their homes, at parties, or in restaurants? VFQ-25 Q14. Because of your eyesight, how much difficulty do you have going out to see movies, plays, or sports events? specific to eyesight, rather than eyes VFQ-25 A9. Because of your eyesight, how much difficulty do you have entertaining friends and family in your home? (all VFQ-25 related to eyesight)
	Fitting in	I have problems fitting in because of my eyes My eyes cause problems with fitting in	No

	Happiness	My eyes cause my to be unhappy I am happy because of my eyes I am unhappy because of my eyes I feel unhappy because of my eyes	No
	Tense	My eyes cause me to be tense I feel tense because of my eyes	No
	Bad tempered	My eyes cause me to be bad tempered I am bad tempered because of my eyes	Partial VFQ-25 A12. I am often irritable because of my eyesight. (eyesight rather than eyes)
	Embarrassed	I am embarrassed about my eyes My eyes cause me to feel embarrassed	Partial AS-20 Q8) I feel inferior to others because of my eyes VFQ-25 Q25. I worry about doing things that will embarrass myself or others, because of my eyesight
	Depression	My eyes cause me to feel depressed I feel depressed because of my eyes	No
	Anxiety	My eyes cause me to feel anxious I feel anxious because of my eyes	No
	Stress	My eyes cause me stress My eyes cause me to feel stressed I get stressed because of my eyes	Yes AS-20 Q 17) I feel stressed because of my eyes
	Worry	I worry about my eyes My eyes cause me to feel worried	Yes AS-20 Q1) I worry about what people will think about my eyes AS-20 Q18) I worry about my eyes VFQ-25 Q3. How much of the time do you worry about your eyesight?
	Self-conscious	I am self-conscious about my eyes I am self-conscious about my appearance	Yes AS-20 Q6) I am self-conscious about my eyes

Topics that were adequately covered by existing AS-20 and VFQ-25 questions were removed. The remaining topics and the possible wording of questions to capture these self-reported topics were discussed with study supervisors and the advisory group. Using a similar question and response style to the AS-20 and VFQ-25, questions were worded using the responses: all of the time, most of the time, some of the time, a little

of the time and none of the time. Table I-8 shows the refinement of the question selection and the final wording of the additional study questions.

Table I-8 Refinement of the selection of the additional study questions and the wording of the questions			
Theme	Development of question wording	Refinement of question selection and wording	Final question wording
Vision	My eyes turn	My eyes turn	My eyes have been turning
	I have difficulty controlling my eye position	I can control my eye position	I been able to control my eye position
	I get blurred vision when my eye turns	X	X
	I have difficulty focussing my eyes together	I can focus on things	I been able to focus my eyes
	I get confused by my eyes	I get confused by what I see	I have been confused by my vision
	I get frustrated by my eyes		
	I get distracted by my eyes		
	I have difficulty swapping to look with each eye individually	I can swap to look with each eye individually	I have been able to swap to look with each eye separately
	I have problems with my vision not being central	My vision looks central	My vision has looked central
	I have difficulty seeing things quickly	I find it difficult to see things when it is very busy	When I have been in a busy place, I have found it difficult to see (for example, a shopping centre or train station)
	I have problems with my vision not being straight	X	
	I have difficulty looking around	I have difficulty moving my eyes to look around	I have found it difficult to move my eyes to look around
	I have difficulty moving my eyes together		
	One of my eyes works much harder than the other eye	One of my eyes works much harder than the other eye	One of my eyes has been working much harder than the other eye
	One of my eyes sees much worse than the other eye		
	I have difficulty looking through both eyes at the same time	I can look through both eyes at the same time	I have been able to look through both eyes at the same time
	I have difficulty using my eyes together	I can use my eyes together	I have been able to use my eyes together
	I have difficulty coordinating my eyes together		
	I have difficulty seeing colours	X	X
	I have difficulty seeing different shades of black, grey and white	X	

Task performance	My eyes limit how hard I can work	My eyes limit my ability to do things (for example work, things I enjoy or activities)	My eyes have limited my ability (for example, at work, in activities I enjoy or in undertaking day-to-day tasks)
	My eyes cause problems with performing tasks		
	I avoid doing things because of my eyes		
	I have difficulty with eye hand coordination	I have difficulty with eye hand coordination	I have had difficulty with eye hand coordination
	I make mistakes when I am performing tasks	I make mistakes when I am performing tasks	I have made mistakes when performing day-to-day tasks
	I have difficulty performing tasks quickly	I can perform tasks quickly	I have been able to perform day-to-day tasks quickly
	I have difficulty judging the position of things	X	X
	I have difficulty looking at screens	I have difficulty looking at screens (for example computer, mobile phone or tablet devices)	I have had difficulty looking at screens (for example, computer, mobile phone or tablet screens)
	I have difficulty with the brightness of rooms or objects	X	X
	I have difficulty concentrating on tasks or activities	X	
	I have difficulty judging steps I can be clumsy and bump into things	X	
	I experience dizziness	X	
	I experience balance problems	I experience balance problems	I have had problems with my balance
Physical symptoms	I get headaches because of my eyes	I get headaches	I have had headaches
	I have to stop doing things because of my eyes	X	X
	I take time off work because of my eyes	X	
	I need to take breaks because of my eyes	I need to take breaks because of my eyes	I have needed to take breaks because of my eyes (for example, breaks from work, from activities I enjoy or when performing day-to-day tasks)
	I find it difficult to see things when it is very busy	Moved to vision theme	X
	I feel a pulling sensation in my eyes	I feel my eyes pulling	I have felt my eyes pulling
	I feel a tightness in my eyes	My eyes feel tight	My eyes have felt tight
	My eyes water	X	X
	I avoid looking at things because of my eyes	X	

	My eye position is affected by tiredness	I feel tired	I have felt tired
	My eyes make me tired		
	I can feel my eyes turning	I can feel my eyes turning	I have felt my eyes turning
Confidence and emotions	I have problems communicating with people because of my eyes	I have problems communicating with people	I have been able to talk to people
	I avoid face-to-face situations because of my eyes	I avoid face-to-face situations	I have avoided face-to-face situations
	It bothers me when people comment on my eyes	It bothers me when people comment on my eyes	People have treated me differently
	People treat me differently because of my eyes	People treat me differently	
	I am treated differently at work because of my eyes		
	I am bullied because of my eyes		
	I have problems relying on my vision	X	X
	My eyes cause me to lack confidence in my abilities	I have confidence in my eyes	I have had confidence in my vision
		I have confidence in my abilities	I have had confidence in my abilities (for example, at work, to take part in activities I enjoy or to undertake day-to-day tasks)
		I have self-confidence	I have had self-confidence
	I dislike the ways my eyes look	I like the way my eyes look	I have liked the way my eyes look
	I avoid social situations because of my eyes	X	X
	I have problems fitting in because of my eyes	X	
	I feel tense because of my eyes	I am bad tempered	X
	I am bad tempered because of my eyes		
	I feel depressed because of my eyes	I feel depressed	X
	I feel anxious because of my eyes	I feel anxious	I have felt anxious
	I am embarrassed because of my eyes	I feel embarrassed about my eyes	I have felt embarrassed about my eyes
	I feel unhappy because of my eyes	I feel happy	I have felt happy
X question removed			

Appendix I.7 Stage four – additional study questions developed

Additional Study Questions

Introduction:

Study questionnaire – Instructions for Patient

This questionnaire contains 33 statements about how your eyes, your vision and your strabismus (misaligned eyes) may affect you in your everyday life.

If you are unable to complete this questionnaire on your own, please ask for assistance from the researcher.

Instructions:

- Please answer each of the following questions by circling the statement that best describes your eyes and your experiences
- Circle only ONE response for each statement
- Please answer based your eyes and your experiences **during the last 7 days**
- If you wear glasses or contact lenses respond as if you were wearing them
- If you are not sure how to respond, please circle the response you think is most appropriate and make a comment to the researcher.

If you have any questions, please ask.

Thank you for completing this questionnaire.

Possible responses to each question:

none of the time	a little of the time	some of the time	most of the time	all of the time

Questions:

1. My eyes have been turning
2. I have been able to control my eye position
3. I have been able to focus my eyes
4. I have been confused by my vision
5. I have been able to swap to look with each eye separately
6. My vision has looked central
7. When I have been in a busy place, I have found it difficult to see (for example, a shopping centre or train station)
8. I have found it difficult to move my eyes to look around
9. One of my eyes has been working much harder than the other eye
10. I have been able to look through both eyes at the same time
11. I have been able to use my eyes together
12. My eyes have limited my ability (for example, at work, in activities I enjoy or in undertaking day-to-day tasks)
13. I have had difficulty with eye hand coordination
14. I have made mistakes when performing day-to-day tasks
15. I have been able to perform day-to-day tasks quickly

16. I have had difficulty looking at screens (for example, computer, mobile phone or tablet screens)
17. I have had problems with my balance
18. I have had headaches
19. I have needed to take breaks because of my eyes (for example, breaks from work, from activities I enjoy or when performing day-to-day tasks)
20. I have felt my eyes pulling
21. My eyes have felt tight
22. I have felt tired
23. I have felt my eyes turning
24. I have been able to talk to people
25. I have avoided face-to-face situations
26. People have treated me differently
27. I have had confidence in my vision
28. I have had confidence in my abilities (for example, at work, to take part in activities I enjoy or to undertake day-to-day tasks)
29. I have had self-confidence
30. I have liked the way my eyes look
31. I have felt anxious
32. I have felt embarrassed about my eyes
33. I have felt happy

Appendix J. AS-20 questionnaire and scoring

AS-20 QUESTIONNAIRE

Adult Strabismus Quality of Life Questionnaire (AS-20)
(May 2008 version)

Instructions for Patient

The AS-20 is a short questionnaire with statements about how strabismus (misaligned eyes) may affect you in your everyday life.

If you are unable to complete this on your own, please ask for someone to assist you.

Instructions:

- Please respond to EACH statement by circling the response that best reflects how you feel.
- Circle only ONE response for each statement.
- Please answer based on your experiences **during the past month, or since your last appointment if sooner.**
- If you wear glasses or contact lenses respond as if you were wearing them, unless otherwise instructed.
- If you are not sure how to respond, please circle the response you think is most appropriate and make a comment in the margin.

If you have any questions please ask.

Thank you for completing this questionnaire.

Adult Strabismus Quality of Life Questionnaire (AS-20)

1) I worry about what people will think about my eyes

Never	Rarely	Sometimes	Often	Always

2) I feel that people are thinking about my eyes even when they don't say anything

Never	Rarely	Sometimes	Often	Always

3) I feel uncomfortable when people are looking at me because of my eyes

Never	Rarely	Sometimes	Often	Always

4) I wonder what people are thinking when they are looking at me because of my eyes

Never	Rarely	Sometimes	Often	Always

5) People don't give me opportunities because of my eyes

Never	Rarely	Sometimes	Often	Always

6) I am self-conscious about my eyes

Never	Rarely	Sometimes	Often	Always

7) People avoid looking at me because of my eyes

Never	Rarely	Sometimes	Often	Always

8) I feel inferior to others because of my eyes

Never	Rarely	Sometimes	Often	Always

9) People react differently to me because of my eyes

Never	Rarely	Sometimes	Often	Always

10) I find it hard to initiate contact with people I don't know because of my eyes

Never	Rarely	Sometimes	Often	Always

11) I cover or close one eye to see things better

Never	Rarely	Sometimes	Often	Always

12) I avoid reading because of my eyes

Never	Rarely	Sometimes	Often	Always

13) I stop doing things because my eyes make it difficult to concentrate

Never	Rarely	Sometimes	Often	Always

14) I have problems with depth perception

Never	Rarely	Sometimes	Often	Always

15) My eyes feel strained

Never	Rarely	Sometimes	Often	Always

16) I have problems reading because of my eye condition

Never	Rarely	Sometimes	Often	Always

17) I feel stressed because of my eyes

Never	Rarely	Sometimes	Often	Always

18) I worry about my eyes

Never	Rarely	Sometimes	Often	Always

19) I can't enjoy my hobbies because of my eyes

Never	Rarely	Sometimes	Often	Always

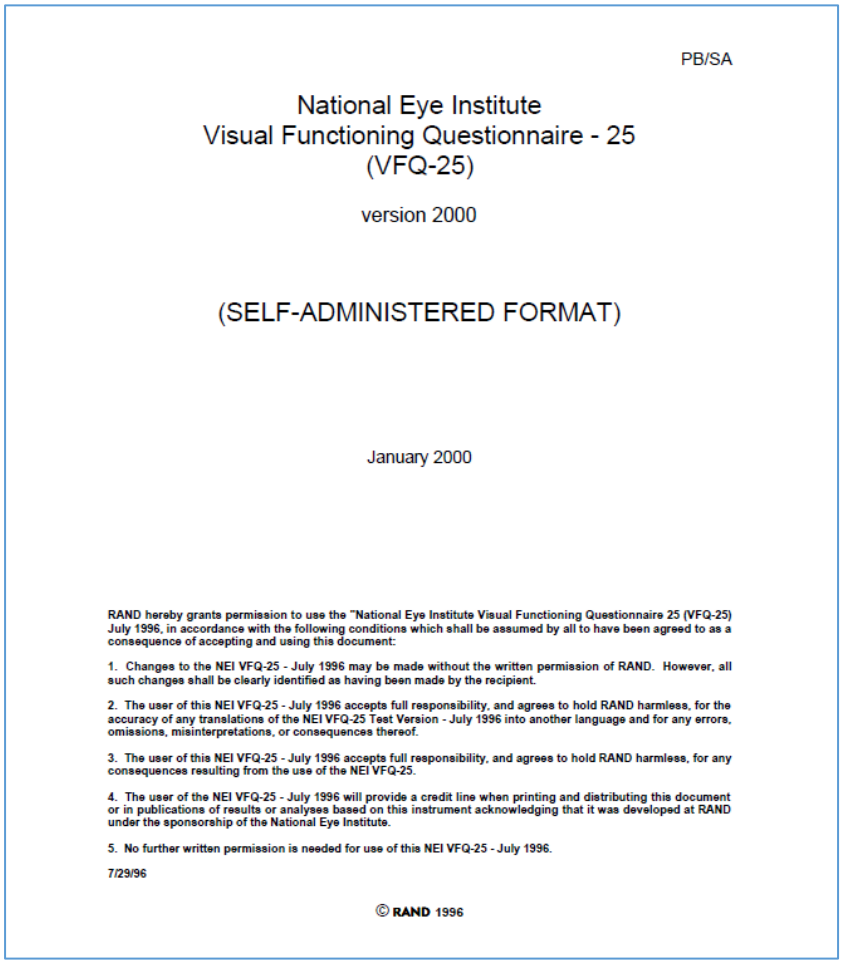
20) I need to take frequent breaks when reading because of my eyes

Never	Rarely	Sometimes	Often	Always

All questions were scored in the same direction. A lower score equated to a worse QoL and a higher score equated to a better QoL. Always = 0, Often = 25, Sometimes = 50, Rarely = 75, Never = 100.

Appendix K. VFQ-25 questionnaire

Appendix K.1 Instructions and questions



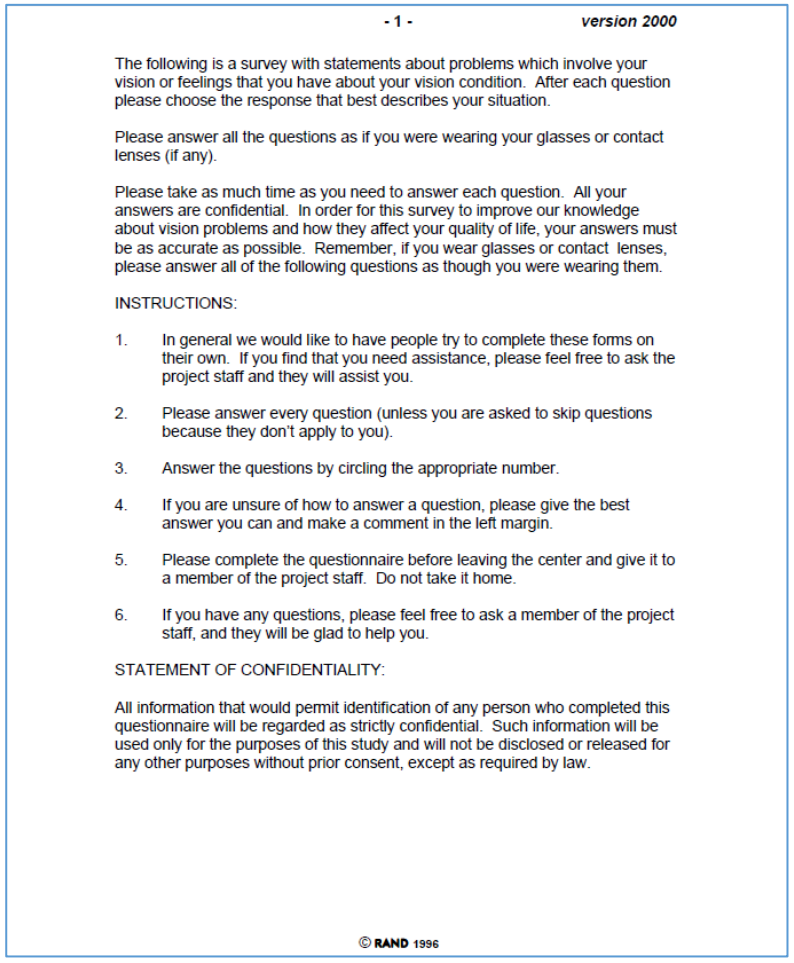
PB/SA

National Eye Institute
Visual Functioning Questionnaire - 25
(VFQ-25)

version 2000

(SELF-ADMINISTERED FORMAT)

January 2000



The following is a survey with statements about problems which involve your vision or feelings that you have about your vision condition. After each question please choose the response that best describes your situation.

Please answer all the questions as if you were wearing your glasses or contact lenses (if any).

Please take as much time as you need to answer each question. All your answers are confidential. In order for this survey to improve our knowledge about vision problems and how they affect your quality of life, your answers must be as accurate as possible. Remember, if you wear glasses or contact lenses, please answer all of the following questions as though you were wearing them.

INSTRUCTIONS:

1. In general we would like to have people try to complete these forms on their own. If you find that you need assistance, please feel free to ask the project staff and they will assist you.

2. Please answer every question (unless you are asked to skip questions because they don't apply to you).

3. Answer the questions by circling the appropriate number.

4. If you are unsure of how to answer a question, please give the best answer you can and make a comment in the left margin.

5. Please complete the questionnaire before leaving the center and give it to a member of the project staff. Do not take it home.

6. If you have any questions, please feel free to ask a member of the project staff, and they will be glad to help you.

STATEMENT OF CONFIDENTIALITY:

All information that would permit identification of any person who completed this questionnaire will be regarded as strictly confidential. Such information will be used only for the purposes of this study and will not be disclosed or released for any other purposes without prior consent, except as required by law.

Visual Functioning Questionnaire - 25

PART 1 - GENERAL HEALTH AND VISION

1. **In general, would you say your overall health is:**

(Circle One)

Excellent	1
Very Good	2
Good	3
Fair	4
Poor	5

2. **At the present time, would you say your eyesight using both eyes (with glasses or contact lenses, if you wear them) is excellent, good, fair, poor, or very poor or are you completely blind?**

(Circle One)

Excellent	1
Good	2
Fair	3
Poor	4
Very Poor	5
Completely Blind	6

3. **How much of the time do you worry about your eyesight?**

(Circle One)

None of the time	1
A little of the time	2
Some of the time	3
Most of the time	4
All of the time?	5

4. **How much pain or discomfort have you had in and around your eyes (for example, burning, itching, or aching)? Would you say it is:**

(Circle One)

None	1
Mild	2
Moderate	3
Severe, or	4
Very severe?	5

PART 2 - DIFFICULTY WITH ACTIVITIES

The next questions are about how much difficulty, if any, you have doing certain activities wearing your glasses or contact lenses if you use them for that activity.

5. **How much difficulty do you have reading ordinary print in newspapers? Would you say you have:**

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Stopped doing this because of your eyesight	5
Stopped doing this for other reasons or not interested in doing this	6

6. How much difficulty do you have doing work or hobbies that require you to see well up close, such as cooking, sewing, fixing things around the house, or using hand tools? Would you say:

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

7. Because of your eyesight, how much difficulty do you have finding something on a crowded shelf?

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

8. How much difficulty do you have reading street signs or the names of stores?

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

9. Because of your eyesight, how much difficulty do you have going down steps, stairs, or curbs in dim light or at night?

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

10. Because of your eyesight, how much difficulty do you have noticing objects off to the side while you are walking along?

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

11. Because of your eyesight, how much difficulty do you have seeing how people react to things you say?

(Circle One)

- No difficulty at all 1
- A little difficulty 2
- Moderate difficulty 3
- Extreme difficulty 4
- Stopped doing this because of your eyesight 5
- Stopped doing this for other reasons or not interested in doing this 6

12. Because of your eyesight, how much difficulty do you have picking out and matching your own clothes?

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Stopped doing this because of your eyesight	5
Stopped doing this for other reasons or not interested in doing this	6

13. Because of your eyesight, how much difficulty do you have visiting with people in their homes, at parties, or in restaurants ?

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Stopped doing this because of your eyesight	5
Stopped doing this for other reasons or not interested in doing this	6

14. Because of your eyesight, how much difficulty do you have going out to see movies, plays, or sports events?

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Stopped doing this because of your eyesight	5
Stopped doing this for other reasons or not interested in doing this	6

15. Are you currently driving, at least once in a while?

(Circle One)

Yes	1	*Skip To Q 15c*
No	2	

15a. IF NO: Have you never driven a car or have you given up driving?

(Circle One)

Never drove	1	*Skip To Part 3, Q 17*
Gave up	2	

15b. IF YOU GAVE UP DRIVING: Was that mainly because of your eyesight, mainly for some other reason, or because of both your eyesight and other reasons?

(Circle One)

Mainly eyesight	1	*Skip To Part 3, Q 17*
Mainly other reasons	2	*Skip To Part 3, Q 17*
Both eyesight and other reasons	3	*Skip To Part 3, Q 17*

15c. IF CURRENTLY DRIVING: How much difficulty do you have driving during the daytime in familiar places? Would you say you have:

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4

16. How much difficulty do you have driving at night? Would you say you have:

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Have you stopped doing this because of your eyesight	5
Have you stopped doing this for other reasons or are you not interested in doing this	6

16A. How much difficulty do you have driving in difficult conditions, such as in bad weather, during rush hour, on the freeway, or in city traffic? Would you say you have:

(Circle One)

No difficulty at all	1
A little difficulty	2
Moderate difficulty	3
Extreme difficulty	4
Have you stopped doing this because of your eyesight	5
Have you stopped doing this for other reasons or are you not interested in doing this	6

PART 3: RESPONSES TO VISION PROBLEMS

The next questions are about how things you do may be affected by your vision. For each one, please circle the number to indicate whether for you the statement is true for you all, most, some, a little, or none of the time.

(Circle One On Each Line)

READ CATEGORIES:	All of the time	Most of the time	Some of the time	A little of the time	None of the time
17. Do you accomplish less than you would like because of your vision?	1	2	3	4	5
18. Are you limited in how long you can work or do other activities because of your vision?	1	2	3	4	5
19. How much does pain or discomfort in or around your eyes, for example, burning, itching, or aching, keep you from doing what you'd like to be doing? Would you say:	1	2	3	4	5

For each of the following statements, please circle the number to indicate whether for you the statement is definitely true, mostly true, mostly false, or definitely false for you or you are not sure.

(Circle One On Each Line)

	Definitely True	Mostly True	Not Sure	Mostly False	Definitely False
20. I stay home most of the time because of my eyesight.	1	2	3	4	5
21. I feel frustrated a lot of the time because of my eyesight.	1	2	3	4	5
22. I have much less control over what I do, because of my eyesight.	1	2	3	4	5
23. Because of my eyesight, I have to rely too much on what other people tell me...	1	2	3	4	5
24. I need a lot of help from others because of my eyesight.	1	2	3	4	5
25. I worry about doing things that will embarrass myself or others, because of my eyesight.	1	2	3	4	5

Appendix K.2 Scoring

Questions recoded following the VFQ-25 scoring guidance described below:

Question number	Original response category	To be recoded to a value of
1, 3, 4	1	100
	2	75
	3	50
	4	25
	5	0
2	1	100
	2	80
	3	60
	4	40
	5	20
	6	0
5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 16a,	1	100
	2	75
	3	50
	4	25
	5	0
	6	*missing
15 (contains filter Qu)	If 15b = 1	15c = 0
	If 15b = 2	15c = * missing
	If 15b = 3	15c = * missing
15c	1	100
	2	75
	3	50
	4	25
17, 18, 19, 20, 21, 22, 23, 24, 25	1	0
	2	25
	3	50

	4	75
	5	100

Items within same subscale are averaged together to create 12 subscale scores.

Subscale	Number of items	Items to be averaged (after recoding)
General health	1	1
General vision	1	2
Ocular pain	2	4, 19
Near activities	3	5, 6, 7
Distance activities	3	8, 9, 14
Vision specific: social functioning	2	11, 13
Vision specific: mental health	4	3, 21, 22, 25
Vision specific: role difficulties	2	17, 18
Vision specific: dependency	3	20, 23, 24
Driving	3	15c, 16, 16a
Colour vision	1	12
Peripheral vision	1	10

Missing data not taken into account when calculating subscale scores - the average subscale score is the average for all items in the subscale that the patient has responded to.

Mean (for each subscale) = sum of the recoded scores for each item with non-missing answer / total number of items in that subscale with non-missing answers.

Calculate overall composite score for the VFQ-25 by taking an average of the vision targeted subscale scores (excluding the general health rating question, Qu1). Mean of 11 subscales = composite score

If a subscale has all missing values (for example: driving) then the overall composite VFQ-25 score is calculated with driving subscale removed.
Example: non-driver / never driven. All driving questions recorded as missing results rather than 0. Driving subscale then recorded as missing rather than 0. Then calculate mean of remaining scored subscales (i.e. remove driving and calculate mean of remaining 10 scored subscales). Including the driving subscale in the calculation of the composite score if have never driven would falsely reduce QoL as the score would be divided by 11 rather than 10 subscales.

Appendix L. Additional study questions

Appendix L.1 Instructions and questions

Measuring the Effects of Eye Alignment Surgery

Phase two – Measuring vision and task performance before and after eye alignment surgery

Study questionnaire – Instructions for Patient

This questionnaire contains 33 statements about how your eyes, your vision and your strabismus (misaligned eyes) may affect you in your everyday life.

If you are unable to complete this questionnaire on your own, please ask for assistance from the researcher.

Instructions:

- Please answer each of the following questions by circling the statement that best describes your eyes and your experiences
- Circle only ONE response for each statement
- Please answer based your eyes and your experiences **during the last 7 days**
- If you wear glasses or contact lenses respond as if you were wearing them
- If you are not sure how to respond, please circle the response you think is most appropriate and make a comment to the researcher.

If you have any questions, please ask.

Thank you for completing this questionnaire.

1. **My eyes have been turning**

none of the time	a little of the time	some of the time	most of the time	all of the time

2. **I have been able to control my eye position**

none of the time	a little of the time	some of the time	most of the time	all of the time

3. **I have been able to focus my eyes**

none of the time	a little of the time	some of the time	most of the time	all of the time

4. **I have been confused by my vision**

none of the time	a little of the time	some of the time	most of the time	all of the time

5. **I have been able to swap to look with each eye separately**

none of the time	a little of the time	some of the time	most of the time	all of the time

6. **My vision has looked central**

none of the time	a little of the time	some of the time	most of the time	all of the time

7. **When I have been in a busy place, I have found it difficult to see (for example, a shopping centre or train station)**

none of the time	a little of the time	some of the time	most of the time	all of the time

8. **I have found it difficult to move my eyes to look around**

none of the time	a little of the time	some of the time	most of the time	all of the time

9. **One of my eyes has been working much harder than the other eye**

none of the time	a little of the time	some of the time	most of the time	all of the time

10. **I have been able to look through both eyes at the same time**

none of the time	a little of the time	some of the time	most of the time	all of the time

11. **I have been able to use my eyes together**

none of the time	a little of the time	some of the time	most of the time	all of the time

12. My eyes have limited my ability (for example, at work, in activities I enjoy or in undertaking day-to-day tasks)

none of the time	a little of the time	some of the time	most of the time	all of the time

13. I have had difficulty with eye hand coordination

none of the time	a little of the time	some of the time	most of the time	all of the time

14. I have made mistakes when performing day-to-day tasks

none of the time	a little of the time	some of the time	most of the time	all of the time

15. I have been able to perform day-to-day tasks quickly

none of the time	a little of the time	some of the time	most of the time	all of the time

16. I have had difficulty looking at screens (for example, computer, mobile phone or tablet screens)

none of the time	a little of the time	some of the time	most of the time	all of the time

17. I have had problems with my balance

none of the time	a little of the time	some of the time	most of the time	all of the time

18. I have had headaches

none of the time	a little of the time	some of the time	most of the time	all of the time

19. I have needed to take breaks because of my eyes (for example, breaks from work, from activities I enjoy or when performing day-to-day tasks)

none of the time	a little of the time	some of the time	most of the time	all of the time

20. I have felt my eyes pulling

none of the time	a little of the time	some of the time	most of the time	all of the time

21. My eyes have felt tight

none of the time	a little of the time	some of the time	most of the time	all of the time

22. I have felt tired

none of the time	a little of the time	some of the time	most of the time	all of the time

23. I have felt my eyes turning

none of the time	a little of the time	some of the time	most of the time	all of the time

24. I have been able to talk to people

none of the time	a little of the time	some of the time	most of the time	all of the time

25. I have avoided face-to-face situations

none of the time	a little of the time	some of the time	most of the time	all of the time

26. People have treated me differently

none of the time	a little of the time	some of the time	most of the time	all of the time

27. I have had confidence in my vision

none of the time	a little of the time	some of the time	most of the time	all of the time

28. I have had confidence in my abilities (for example, at work, to take part in activities I enjoy or to undertake day-to-day tasks)

none of the time	a little of the time	some of the time	most of the time	all of the time

29. I have had self-confidence

none of the time	a little of the time	some of the time	most of the time	all of the time

30. I have liked the way my eyes look

none of the time	a little of the time	some of the time	most of the time	all of the time

31. I have felt anxious

none of the time	a little of the time	some of the time	most of the time	all of the time

32. I have felt embarrassed about my eyes

none of the time	a little of the time	some of the time	most of the time	all of the time

33. I have felt happy

none of the time	a little of the time	some of the time	most of the time	all of the time

Appendix L.2 Scoring

Questions: 1, 4, 7, 8, 9, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 31, 32

The question response ‘none of the time’ equates to the best QoL or best function and ‘all of the time’ equates to the worse QoL or worse function. These questions were scored as follows, so that a higher score reflects a higher QoL or better function.

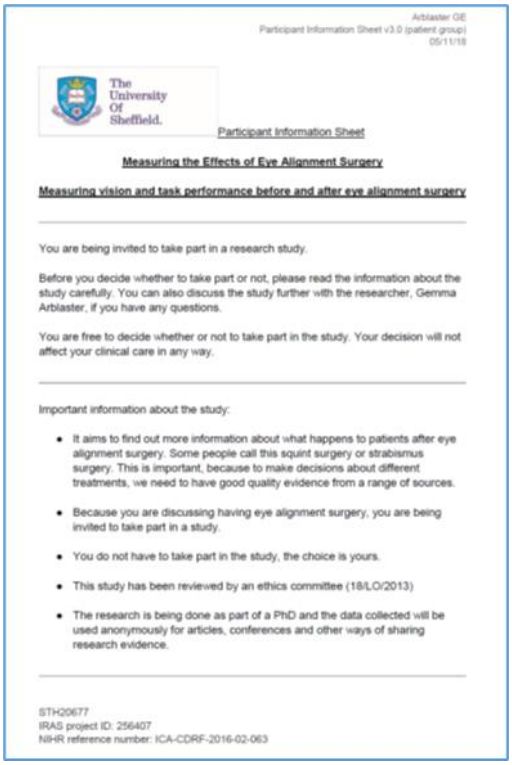
Raw Score	Reported Score (NOT converted)
None of the time = 100	None of the time = 100
A little of the time = 75	A little of the time = 75
Some of the time = 50	Some of the time = 50
Most of the time = 25	Most of the time = 25
All of the time = 0	All of the time = 0

Questions: 2, 3, 5, 6, 10, 11, 15, 24, 27, 28, 29, 30, 33

For these questions ‘none of the time’ equates to the worst QoL or worse function and ‘all of the time’ equates to the best QoL or best function. These questions are scored in reverse, so that a lower score reflects a lower QoL.

Raw Score	Reported score (CONVERTED to reverse the direction of the score)
None of the time = 0	None of the time = 100
A little of the time = 25	A little of the time = 75
Some of the time = 50	Some of the time = 50
Most of the time = 75	Most of the time = 25
All of the time = 100	All of the time = 0

Appendix M. Participant information sheet (patient group) – phase two



The University Of Sheffield.

Participant Information Sheet

Measuring the Effects of Eye Alignment Surgery

Measuring vision and task performance before and after eye alignment surgery

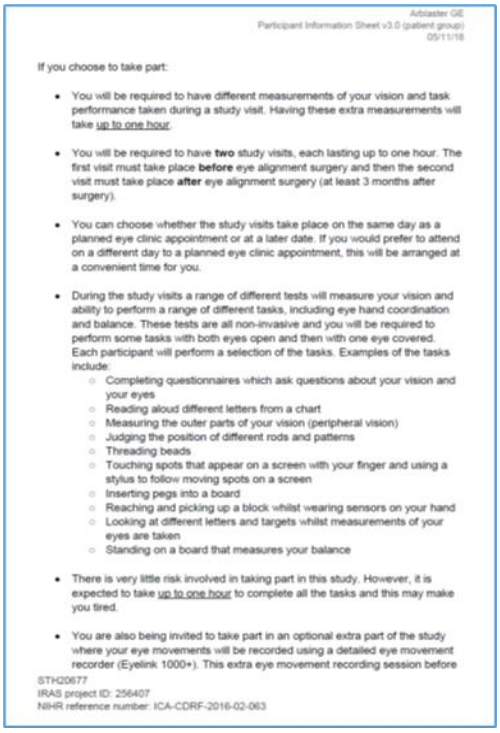
You are being invited to take part in a research study.

Before you decide whether to take part or not, please read the information about the study carefully. You can also discuss the study further with the researcher, Gemma Arblaster, if you have any questions.

You are free to decide whether or not to take part in the study. Your decision will not affect your clinical care in any way.

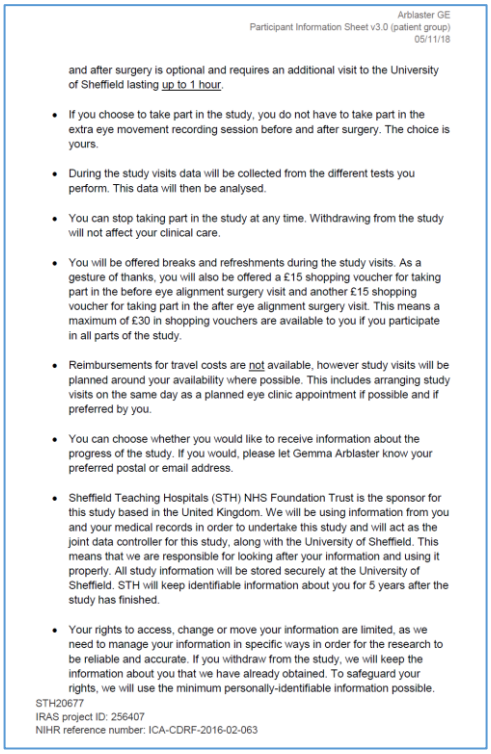
Important information about the study:

- It aims to find out more information about what happens to patients after eye alignment surgery. Some people call this squint surgery or strabismus surgery. This is important, because to make decisions about different treatments, we need to have good quality evidence from a range of sources.
- Because you are discussing having eye alignment surgery, you are being invited to take part in a study.
- You do not have to take part in the study, the choice is yours.
- This study has been reviewed by an ethics committee (18/LO/2013)
- The research is being done as part of a PhD and the data collected will be used anonymously for articles, conferences and other ways of sharing research evidence.



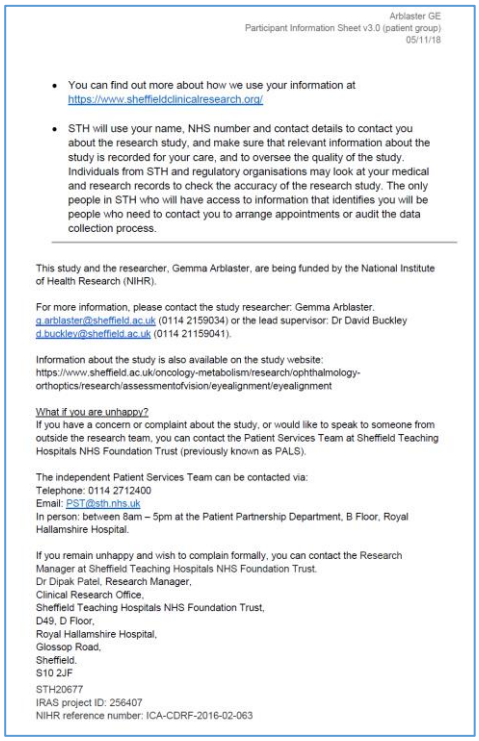
If you choose to take part:

- You will be required to have different measurements of your vision and task performance taken during a study visit. Having these extra measurements will take up to one hour.
- You will be required to have **two** study visits, each lasting up to one hour. The first visit must take place **before** eye alignment surgery and then the second visit must take place **after** eye alignment surgery (at least 3 months after surgery).
- You can choose whether the study visits take place on the same day as a planned eye clinic appointment or at a later date. If you would prefer to attend on a different day to a planned eye clinic appointment, this will be arranged at a convenient time for you.
- During the study visits a range of different tests will measure your vision and ability to perform a range of different tasks, including eye hand coordination and balance. These tests are all non-invasive and you will be required to perform some tasks with both eyes open and then with one eye covered. Each participant will perform a selection of the tasks. Examples of the tasks include:
  - Completing questionnaires which ask questions about your vision and your eyes
  - Reading aloud different letters from a chart
  - Measuring the outer parts of your vision (peripheral vision)
  - Judging the position of different rods and patterns
  - Threading beads
  - Touching spots that appear on a screen with your finger and using a stylus to follow moving spots on a screen
  - Inserting pegs into a board
  - Reaching and picking up a block whilst wearing sensors on your hand
  - Looking at different letters and targets whilst measurements of your eyes are taken
  - Standing on a board that measures your balance
- There is very little risk involved in taking part in this study. However, it is expected to take up to one hour to complete all the tasks and this may make you tired.
- You are also being invited to take part in an optional extra part of the study where your eye movements will be recorded using a detailed eye movement recorder (Eyelink 1000+). This extra eye movement recording session before



and after surgery is optional and requires an additional visit to the University of Sheffield lasting up to 1 hour.

- If you choose to take part in the study, you do not have to take part in the extra eye movement recording session before and after surgery. The choice is yours.
- During the study visits data will be collected from the different tests you perform. This data will then be analysed.
- You can stop taking part in the study at any time. Withdrawing from the study will not affect your clinical care.
- You will be offered breaks and refreshments during the study visits. As a gesture of thanks, you will also be offered a £15 shopping voucher for taking part in the before eye alignment surgery visit and another £15 shopping voucher for taking part in the after eye alignment surgery visit. This means a maximum of £30 in shopping vouchers are available to you if you participate in all parts of the study.
- Reimbursements for travel costs are not available, however study visits will be planned around your availability where possible. This includes arranging study visits on the same day as a planned eye clinic appointment if possible and if preferred by you.
- You can choose whether you would like to receive information about the progress of the study. If you would, please let Gemma Arblaster know your preferred postal or email address.
- Sheffield Teaching Hospitals (STH) NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from you and your medical records in order to undertake this study and will act as the joint data controller for this study, along with the University of Sheffield. This means that we are responsible for looking after your information and using it properly. All study information will be stored securely at the University of Sheffield. STH will keep identifiable information about you for 5 years after the study has finished.
- Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible.



- You can find out more about how we use your information at https://www.sheffieldclinicalresearch.org/
- STH will use your name, NHS number and contact details to contact you about the research study, and make sure that relevant information about the study is recorded for your care, and to oversee the quality of the study. Individuals from STH and regulatory organisations may look at your medical and research records to check the accuracy of the research study. The only people in STH who will have access to information that identifies you will be people who need to contact you to arrange appointments or audit the data collection process.

This study and the researcher, Gemma Arblaster, are being funded by the National Institute of Health Research (NIHR).

For more information, please contact the study researcher: Gemma Arblaster. g.arblaster@sheffield.ac.uk (0114 2159034) or the lead supervisor: Dr David Buckley d.buckley@sheffield.ac.uk (0114 2159041).

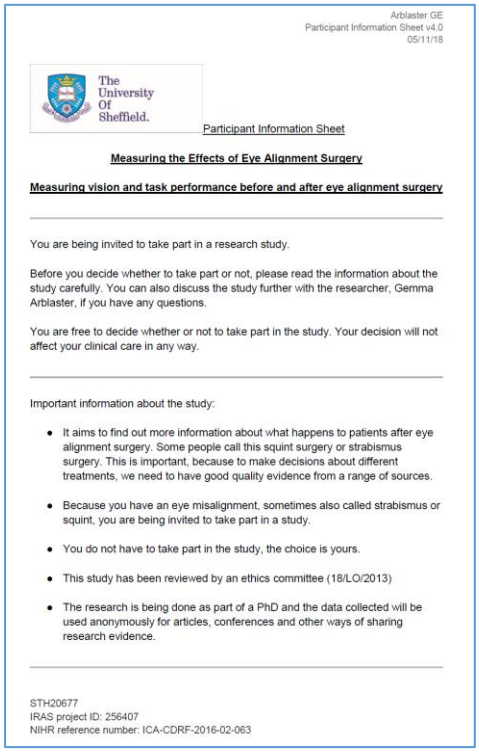
Information about the study is also available on the study website: https://www.sheffield.ac.uk/oncology-metabolism/research/ophthalmology-orthoptics/research/assessmentofvision/eyealignment/eyealignment

What if you are unhappy?
If you have a concern or complaint about the study, or would like to speak to someone from outside the research team, you can contact the Patient Services Team at Sheffield Teaching Hospitals NHS Foundation Trust (previously known as PALS).

The independent Patient Services Team can be contacted via:
Telephone: 0114 2712400
Email: PST@sth.nhs.uk
In person: between 8am – 5pm at the Patient Partnership Department, B Floor, Royal Hallamshire Hospital.

If you remain unhappy and wish to complain formally, you can contact the Research Manager at Sheffield Teaching Hospitals NHS Foundation Trust.
Dr Dipak Patel, Research Manager,
Clinical Research Office,
Sheffield Teaching Hospitals NHS Foundation Trust,
D49, D Floor,
Royal Hallamshire Hospital,
Glossop Road,
Sheffield.
S10 2JF

Appendix N. Participant information sheet (control group) – phase two



The University Of Sheffield.

Participant Information Sheet

Measuring the Effects of Eye Alignment Surgery

Measuring vision and task performance before and after eye alignment surgery

---

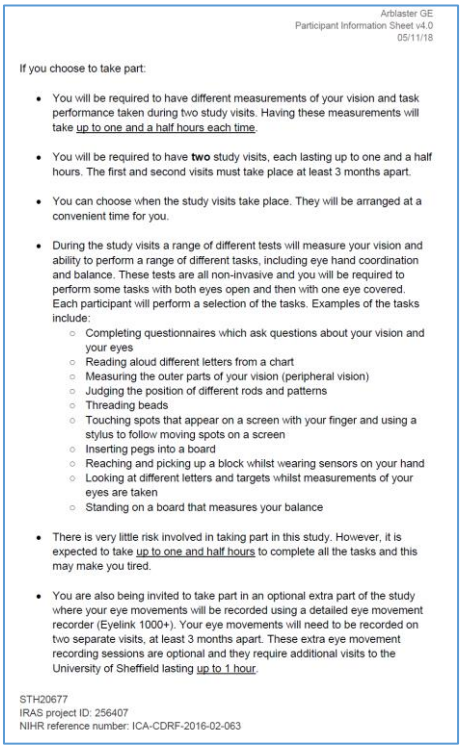
You are being invited to take part in a research study.

Before you decide whether to take part or not, please read the information about the study carefully. You can also discuss the study further with the researcher, Gemma Arblaster, if you have any questions.

You are free to decide whether or not to take part in the study. Your decision will not affect your clinical care in any way.

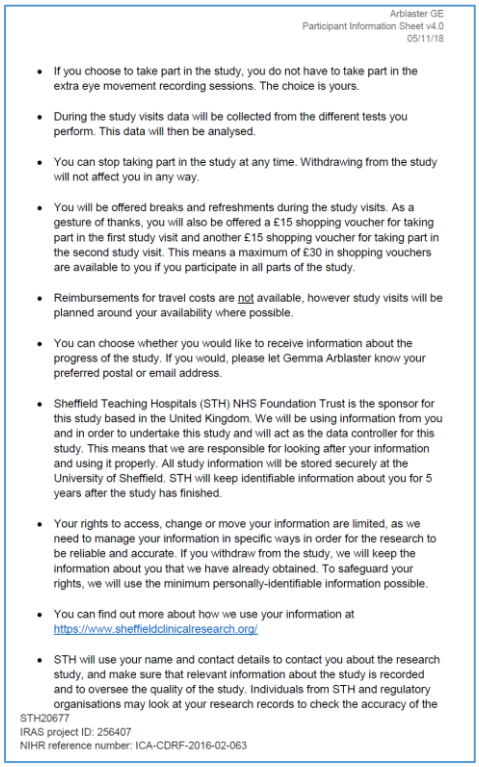
Important information about the study:

- It aims to find out more information about what happens to patients after eye alignment surgery. Some people call this squint surgery or strabismus surgery. This is important, because to make decisions about different treatments, we need to have good quality evidence from a range of sources.
- Because you have an eye misalignment, sometimes also called strabismus or squint, you are being invited to take part in a study.
- You do not have to take part in the study, the choice is yours.
- This study has been reviewed by an ethics committee (18/LO/2013)
- The research is being done as part of a PhD and the data collected will be used anonymously for articles, conferences and other ways of sharing research evidence.

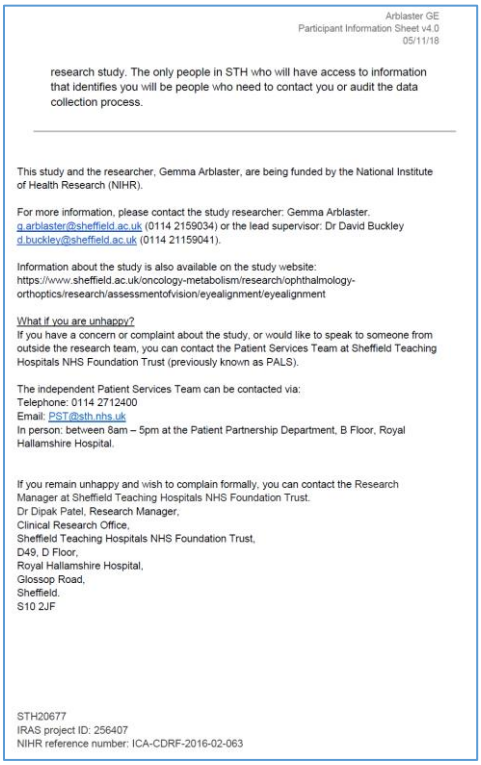


If you choose to take part:

- You will be required to have different measurements of your vision and task performance taken during two study visits. Having these measurements will take up to one and a half hours each time.
- You will be required to have **two** study visits, each lasting up to one and a half hours. The first and second visits must take place at least 3 months apart.
- You can choose when the study visits take place. They will be arranged at a convenient time for you.
- During the study visits a range of different tests will measure your vision and ability to perform a range of different tasks, including eye hand coordination and balance. These tests are all non-invasive and you will be required to perform some tasks with both eyes open and then with one eye covered. Each participant will perform a selection of the tasks. Examples of the tasks include:
  - Completing questionnaires which ask questions about your vision and your eyes
  - Reading aloud different letters from a chart
  - Measuring the outer parts of your vision (peripheral vision)
  - Judging the position of different rods and patterns
  - Threading beads
  - Touching spots that appear on a screen with your finger and using a stylus to follow moving spots on a screen
  - Inserting pegs into a board
  - Reaching and picking up a block whilst wearing sensors on your hand
  - Looking at different letters and targets whilst measurements of your eyes are taken
  - Standing on a board that measures your balance
- There is very little risk involved in taking part in this study. However, it is expected to take up to one and half hours to complete all the tasks and this may make you tired.
- You are also being invited to take part in an optional extra part of the study where your eye movements will be recorded using a detailed eye movement recorder (Eyelink 1000+). Your eye movements will need to be recorded on two separate visits, at least 3 months apart. These extra eye movement recording sessions are optional and they require additional visits to the University of Sheffield lasting up to 1 hour.



- If you choose to take part in the study, you do not have to take part in the extra eye movement recording sessions. The choice is yours.
- During the study visits data will be collected from the different tests you perform. This data will then be analysed.
- You can stop taking part in the study at any time. Withdrawing from the study will not affect you in any way.
- You will be offered breaks and refreshments during the study visits. As a gesture of thanks, you will also be offered a £15 shopping voucher for taking part in the first study visit and another £15 shopping voucher for taking part in the second study visit. This means a maximum of £30 in shopping vouchers are available to you if you participate in all parts of the study.
- Reimbursements for travel costs are not available, however study visits will be planned around your availability where possible.
- You can choose whether you would like to receive information about the progress of the study. If you would, please let Gemma Arblaster know your preferred postal or email address.
- Sheffield Teaching Hospitals (STH) NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from you and in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. All study information will be stored securely at the University of Sheffield. STH will keep identifiable information about you for 5 years after the study has finished.
- Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible.
- You can find out more about how we use your information at https://www.sheffieldclinicalresearch.org/
- STH will use your name and contact details to contact you about the research study, and make sure that relevant information about the study is recorded and to oversee the quality of the study. Individuals from STH and regulatory organisations may look at your research records to check the accuracy of the



research study. The only people in STH who will have access to information that identifies you will be people who need to contact you or audit the data collection process.

This study and the researcher, Gemma Arblaster, are being funded by the National Institute of Health Research (NIHR).

For more information, please contact the study researcher: Gemma Arblaster. g.arblaster@sheffield.ac.uk (0114 2159034) or the lead supervisor: Dr David Buckley d.buckley@sheffield.ac.uk (0114 21159041).

Information about the study is also available on the study website: https://www.sheffield.ac.uk/oncology-metabolism/research/ophthalmology-orthoptics/research/assessmentofvision/eyealignment/eyealignment

What if you are unhappy?
If you have a concern or complaint about the study, or would like to speak to someone from outside the research team, you can contact the Patient Services Team at Sheffield Teaching Hospitals NHS Foundation Trust (previously known as PALS).

The independent Patient Services Team can be contacted via:
Telephone: 0114 2712400
Email: PST@sth.nhs.uk
In person: between 8am – 5pm at the Patient Partnership Department, B Floor, Royal Hallamshire Hospital.

If you remain unhappy and wish to complain formally, you can contact the Research Manager at Sheffield Teaching Hospitals NHS Foundation Trust.
Dr Dipak Patel, Research Manager,
Clinical Research Office,
Sheffield Teaching Hospitals NHS Foundation Trust,
D49, D Floor,
Royal Hallamshire Hospital,
Glossop Road,
Sheffield.
S10 2JF

Appendix O. Consent form – phase two

Arblaster GE
Consent form v1.0
24/10/18

The University Of Sheffield.

Participant Identification Number:

Consent Form

Measuring the Effects of Eye Alignment Surgery.

Measuring vision and task performance before and after eye alignment surgery

Name of Researcher: Gemma Arblaster

Please initial each box:

1. I confirm that I have read the information sheet dated __/__/20__ (version__) and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my medical notes may be looked at by individuals from the University of Sheffield or from Sheffield Teaching Hospitals NHS Foundation Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I understand the information I provide, as part of this study, will only be accessed by members of the research team and this will be stored securely at the University of Sheffield. ☐
5. I understand that my information, as part of this study, will be used to support other research in the future, and may be shared anonymously with other researchers. ☐
6. I understand that I can receive information about the progress of the study if I wish, and this will require me to provide the researcher with my preferred postal or email address. ☐
7. I agree to take part in the above study. ☐

Name of Participant Date Signature

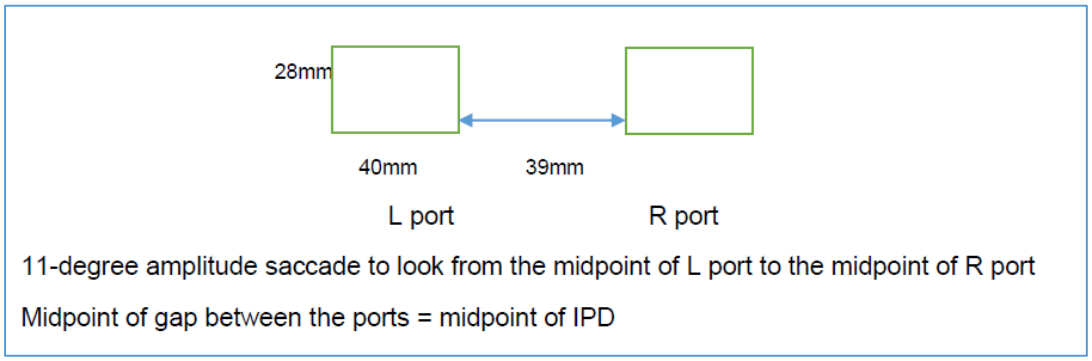
Name of Person taking consent Date Signature

IRAS ID: 256407
Study Number: STH20677
NIHR reference number: ICA-CDRF-2016-02-063

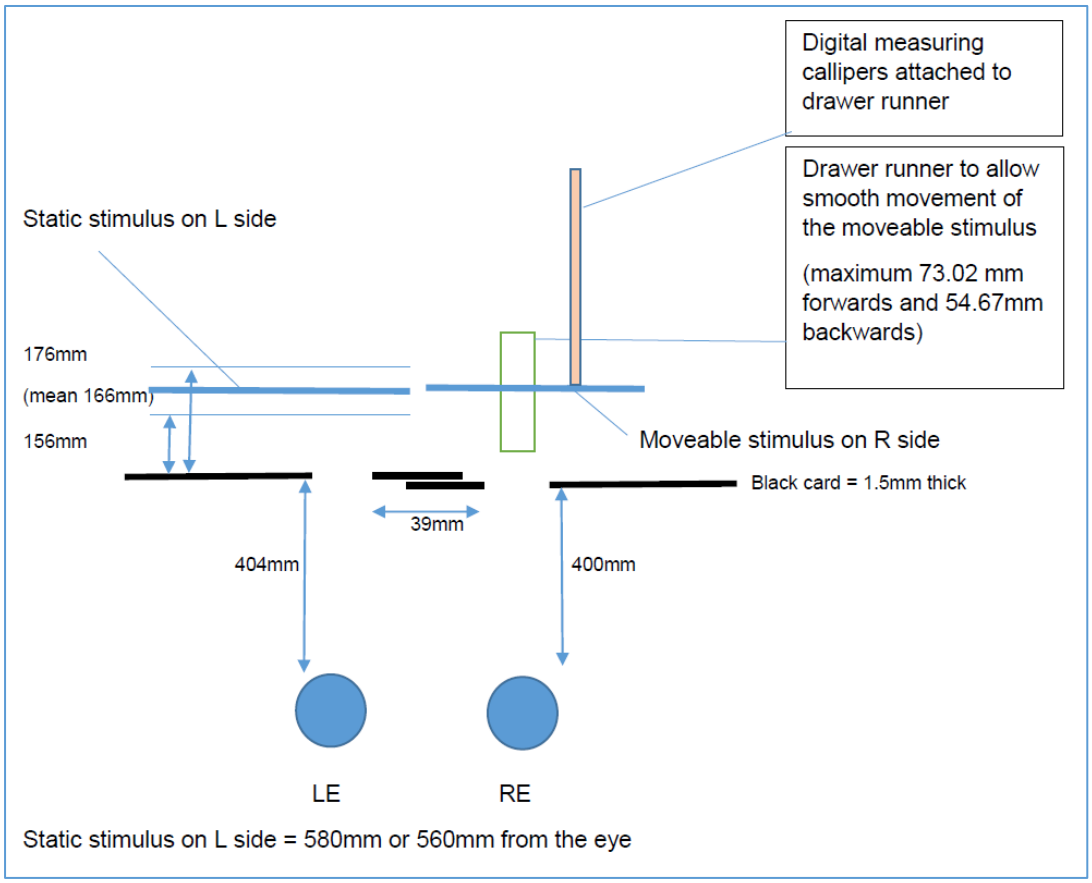
Copies to: patient
investigator site file
medical notes

Appendix P. Sequential stereopsis set up and testing method

Viewing ports



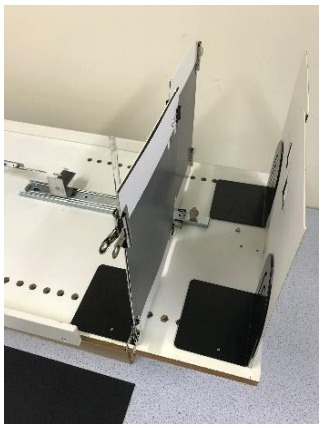
Schematic drawing of the sequential stereopsis rig (not to scale)



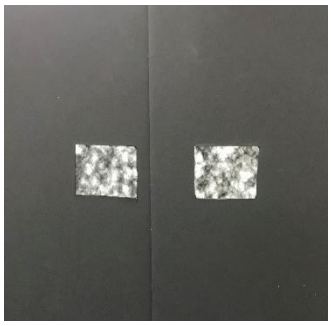
Sequential stereopsis set up and instructions

The testing rig front surface, consisted of two pieces of stiff black card (1.5mm thick) mounted vertically on book ends. Cut into the black pieces of card were rectangular viewing

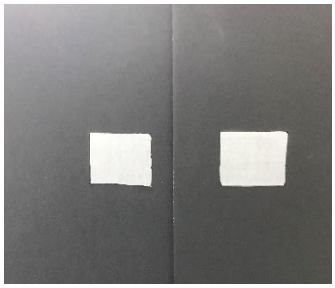
ports, through which the participant viewed the stimuli placed behind the ports. The cut edges of the viewing ports were coloured black to prevent a visible edge or border effect.



Unfiltered stimuli visible through the viewing ports



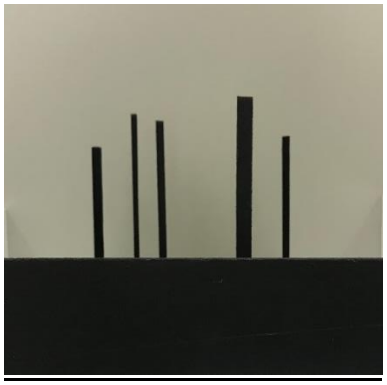
High pass filtered stimuli visible through the viewing ports



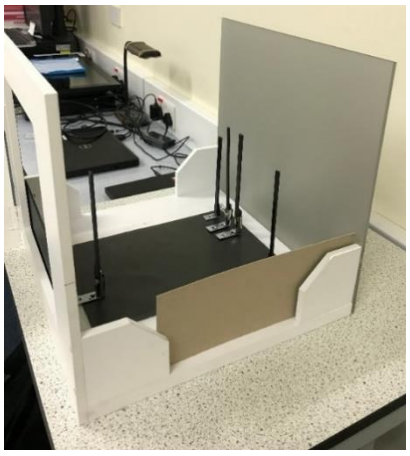
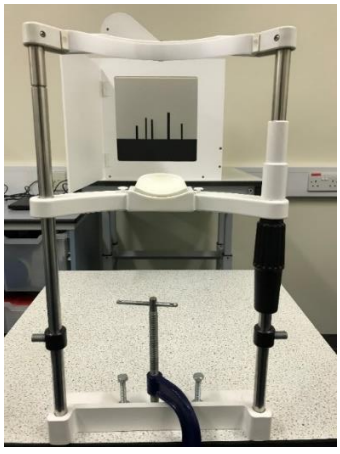
The patterned stimuli behind each of the viewing ports were the same pattern (either unfiltered or high pass filtered), but they did not match exactly as the stimulus on the right was inverted compared to the stimulus on the left side. This was presented deliberately to prevent the participant matching the patterns.

Appendix Q. Coarse stereotest set up and testing method

The CST was a white box (35cm high x 40cm wide) with a door at the front, to allow viewing of the stimuli within the box when opened. The back of the box was opaque plastic and the inside of the box was lined with black card to prevent shadows or reflections forming on the stimuli. Inside the box the base was magnetic to allow the stimuli to be positioned within the box and held in place securely.


The stimuli were 10 different sized vertical black rods attached to a metal base. The rods were slightly varying shades of black to prevent colour matching or colour change being used as a monocular clue. The 2mm thick rods were made of card, with no visible edge.


<u>Stimuli presented at 3094” of arc</u>

Angle of disparity (seconds of arc) = C x IPD (mm) x distance between the stimuli (mm) / viewing distance $(mm)^2$

Constant C = 206264.88 (allowed disparity in radians to be converted to disparity in seconds of arc)

1 radian = 57.2958 degrees, x 60 = minutes of arc, x 60 = seconds of arc

57.2958 x 60 x 60 = 206264.88

IPD measured in mm (average of 60mm used)

Distance between stimuli (background and foreground stimuli) (mm) = 250mm

Viewing distance, measured from the eye to the background stimuli (mm) = 1000mm

Vertical rod

A: 8.8mm wide x 152.0mm tall

B: 9.3mm wide x 157.3mm tall

C: 8.8mm wide x 159.0mm tall

D: 9.1mm wide x 161.0mm tall

E: 5.0mm wide x 160.0mm tall

F: 6.7mm wide x 160.0mm tall

G: 6.5mm wide x 161.0mm tall

H: 7.1mm wide x 176.6mm tall

I: 8.6mm wide x 185.3mm tall

J: 5.1mm wide x 170.0mm tall

The 10 rods were a minimum height of 13cm to ensure the rod was visible in the different positions (back and forwards)

Rods used in each test order

Test order	Forward rod	Position number	Rod 1	Rod 2	Rod 3	Rod 4	Rod 5
1	H	3	E	C	**H**	F	J
2	D	4	G	A	B	**D**	I
3	B	5	H	F	I	C	**B**
4	G	2	A	**G**	D	E	J

5	A	1	**A**	B	H	I	C
6	F	4	D	G	J	**F**	E
7	I	1	**I**	C	H	G	A
8	J	2	E	**J**	F	D	B
9	C	3	D	A	**C**	H	F
10	E	5	J	I	G	B	**E**

Counterbalancing – of the test order

Participant number	**Counterbalanced test order**									
1	1	2	10	3	9	4	8	5	7	6
2	2	3	1	4	10	5	9	6	8	7
3	3	4	2	5	1	6	10	7	9	8
4	4	5	3	6	2	7	1	8	10	9
5	5	6	4	7	3	8	2	9	1	10
6	6	7	5	8	4	9	3	10	2	1
7	7	8	6	9	5	10	4	1	3	2
8	8	9	7	10	6	1	5	2	4	3
9	9	10	8	1	7	2	6	3	5	4
10	10	1	9	2	8	3	7	4	6	5

Appendix R. CKAT set up and testing method

CKAT

Data from the CKAT was extracted using the inbuilt ‘data manager’ function. Data was presented in Excel. Data errors appeared in the Excel spreadsheet as ‘error’. Data from all tasks was additionally checked manually, by viewing the raw CKAT data, prior to including the data in the subsequent analysis. This was to ensure no errors had been incorrectly included in the recorded data.

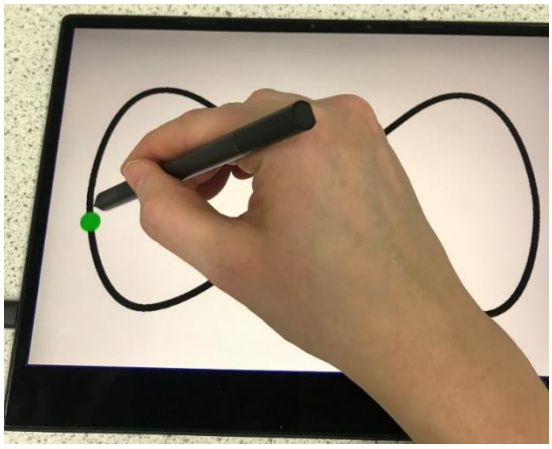
Tracking

5mm diameter circular target used during tracking task

3mins total tracking time

Tracking accuracy (RMSE mm) = straight line distance (mm) between the centre of the moving target and the tip of the stylus (120Hz sampling rate).

A smaller RMSE (mm) represented better tracking accuracy.



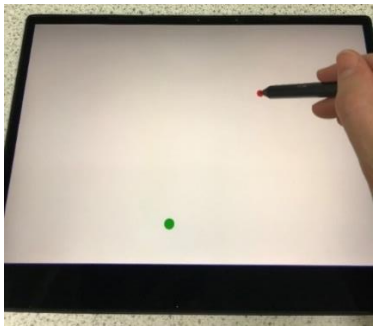
Tracking task shown with background guide

Aiming

75 spots (5mm diameter) appeared during the aiming task.

Mean Path Length Time (seconds) was calculated from the path length time (seconds) to touch the first 50 spots.

If data from less than 50 spots was available, a minimum of 8 spots was required to calculate the mean.



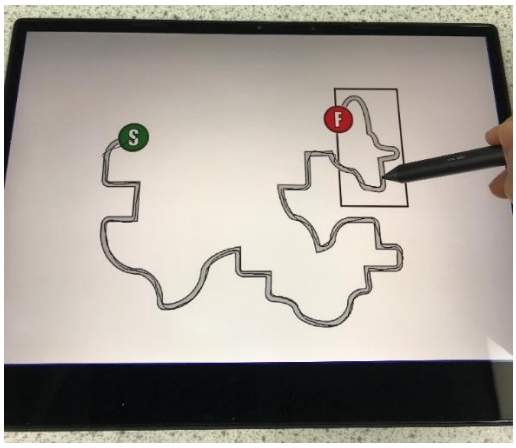
Steering

4mm wide tracing path made up of straight line and curved trajectories.

Penalised Path Accuracy (pPA) = Path Accuracy x (1+((ABS(Path Length Time – 36)) / 36))

Path Accuracy = mean error (mm) between stylus position and the centre of the idealised reference path at each sampled point (120Hz)

Path Length Time = time to complete the tracing path (seconds)



(Culmer et al., 2009; Cunningham et al., 2019)

Appendix S. TSL screen validation

Screen

iiyama prolite TC2735MSC-B2 27” multi-touch monitor

Response time of the screen = 5 msec (or 0.005 seconds)

1920x1080 full HD resolution with projective capacitive 10-point touch technology to accurately display the stimuli and measure touch responses. The screen was a capacitive type device, which used the centre of skin/surface contact area to define the contact point.

Validation of the touchscreen prior to data collection

The purpose of the validation was to validate that the 2D contact position (of the finger) reported by the touchscreen was accurate.

The two methods of measuring the point on the screen touched by the finger were compared (method 1. Screen reported coordinates when the finger touched the screen and method 2. Camera-based method (Check2D) recording where the finger touched the screen).

A camera-based spatial reconstruction method (Check2D) was used to gain a plan view image of the screen and record an unbiased measurement of the finger touching the screen, during the TSL task. It was assumed that the contact area (finger touching the screen) was nominally circular. Validation steps:

1. Check 2D, camera calibration software, for 2D kinematic analysis was used with a checkerboard pattern to calibrate the intrinsic and extrinsic parameters of the camera using the planar calibration method (Goodwill, 2013)
2. 10 trials of data (100 spot touches in total) were collected using spots presented in a random location on the screen. GA was used as the participant in the validation.
3. During the data collection (100 spots) the TSL program recorded the location on the screen where the spot was presented and the location where the finger touched the screen. A camera was used to record all the trials, so that there was visual footage of the finger touching the screen each time a spot was presented.
4. The camera was located 2m above the screen (to minimise parallax errors) with a plumb line to ensure the camera was located centrally above the screen.

5. Camera footage was digitised and each frame reviewed to ensure the optimum image (frame number) was selected to represent the point the screen was touched. It was assumed that the contact area (finger touching the screen) was nominally circular. A circular cursor was placed over the digital image of the fingertip that was aligned to the skin/surface contact area. The 2D position of the centre of the marker was calculated. This was repeated for each spot in each trial (total 100 spots).
6. The screen was presented in landscape. A value of 3.2 pixels / mm was generated from the screen dimensions. Screen = 600mm x 338mm, 1920 pixels x 1080 pixels. (1920 / 600 = 3.2. 1080 / 338 = 3.2)
7. Data was presented in Excel as x and y coordinates (mm). Data from the touchscreen was the location of the spot presented and the screen location touched by the finger. Data from the camera recording was the screen location touched by the finger.
8. The difference between the reported screen locations touched from the touchscreen and the camera, x and y coordinates for each spot (mm), were calculated (n=100 spots). The mean difference in the x coordinate = 1.44mm (SD 1.20). The mean difference in the y coordinate = 1.47mm (SD 0.96).
9. The difference between the data reported from the touchscreen and the camera was considered to be minimal. The touchscreen data was considered to be an accurate representation of where the participant touched the screen.
10. On a different day, validation steps 1-7 were repeated to generate a second calibration and validation of the camera data. The data from the first and second camera analysis were compared to ensure the process followed was robust and reviewer test-retest variability was minimal.
11. The difference between the reported screen locations touched from the camera at analysis 1 and analysis 2, x and y coordinates for each spot (mm), were calculated (n=100 spots). The mean difference in the x coordinate = 0.49mm (SD 0.35mm). The mean difference in the y coordinate = 0.41mm (SD 0.35mm).
12. The difference between camera analysis 1 and 2 was minimal and was considered repeatable.

Appendix T. Eye movement recordings set up, testing method and analysis

Specifications of the Eyelink 1000+

Used with tower mount and head support

Laptop running eyelink recorder = Dell Precision m4800

PC running stimuli and presentation on the screen = Dell Precision t1700

Screen to display stimuli = Asus nvidia Vg248qe

Screen dimensions = 532mm x 298mm, 1920 x 1080 pixels

Sampling rate = 1000Hz monocularly. Monocular recordings were taken from the fixing eye

Accuracy = typically 0.25-0.50 degrees (of visual angle)

Resolution = 0.01 degrees RMS (SR Research, 2020)

Set up

Camera focussed to ensure eye in focus and corneal reflection small and in focus.

Pupil threshold manually adjusted (and ensured it was >70 and <120 (greyscale values)).

Corneal Reflection Threshold manually adjusted (and ensured it was 200-240).

Calibration

Target = 2 degree (diameter) black cross with a white centre presented on a grey background. Measured in degrees of visual angle. Participants were instructed to look in the centre of the cross during the calibration.

A 5-point calibration was used (centre, horizontal ±10° and vertical ±8°).

Validation performed using the same target as calibration and in the same 5 positions as the calibration. If average error was not <0.5, calibration and validation were repeated.

Calibration was performed before fixation, horizontal smooth pursuit, vertical smooth pursuit, horizontal saccades and vertical saccades.

Analysis

Experimental trials analysed only, not practice trials.

Data was parsed into fixation, a saccade or a blink during the recording, automatically by the Eyelink.

Blink = no corneal reflection or pupil detection possible, eyelids closed.

Saccade = when the eye is detected to be moving with velocity >30°/second, with acceleration >8000°^{2}/second and of amplitude >0.15°.

Fixation = when the eye is not in a blink or in a saccade.

Smooth pursuit sample report in data viewer to extract:

RE H data	RE V data	LE H data	LE V data
Extract: Resolution X R fixation index R gaze X R in blink R in saccade R saccade index R velocity X Sample message Target velocity X Target X	Extract: Resolution Y R fixation index R gaze Y R in blink R in saccade R saccade index R velocity Y Sample message Target velocity Y Target Y	Extract: Resolution X L fixation index L gaze X L in blink L in saccade L saccade index L velocity X Sample message Target velocity X Target X	Extract: Resolution Y L fixation index L gaze Y L in blink L in saccade L saccade index L velocity Y Sample message Target velocity Y Target Y

Resolution X - for horizontal data and Resolution Y - for vertical data

R for RE data and L for LE data

The sample report was processed using code written in R. This took the sample report and analysed pursuit gain, pursuit accuracy and length of fixation for each part of the trial considered a fixation (and not a blink or a saccade).

Appendix U. Summary of phase two measures completed by each participant

Table U-1. Phase two measures completed

	Core clinical				PROMs	Vision					Task					EMR
Participant	VA Dist	VA Nr	Fixation	PCT	PROMs	Binocular Summatio	CS	Sequential stereopsis	CST	VF	Grooved pegboard	Purdue pegboard	Bead threading	TSL	CKAT	EMR
01	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
02	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
03	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
04	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
05	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
06	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
07	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
08	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
09	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
10	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
11	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
13	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
14	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
15	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
16	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N

17	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
18	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	Y
20	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
21	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	Y
22	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
23	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	N
24	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
25	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
27	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
29	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
31	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N
Total	27	26	27	27	27	27	27	27	27	27	26	17	17	16	13	11

Appendix V. Phase two - surgical procedures and postoperative outcome in the surgery group

Table V-1. Surgery group procedures and outcomes

Participant number	Gender	Visit 1 strabismus	Surgical procedure	Visit 2 strabismus	Post-operative sensory status
01	F	Consecutive XT	L LR recession 9mm	Consecutive residual ET	Suppression
03	M	Consecutive XT	L MR resection 5mm L LR recession 6mm	Consecutive residual ET with accommodative element	Nr: Unexpected gross BSV (sensory fusion BG cross with central gap, Wirt fly intermittently) Dist: Suppression
05	F	Consecutive XT, DVD	L SR recession 4.5mm L LR recession 4mm L LR scar tissue excised	Consecutive residual ET, DVD	Nr: mostly suppression, occasional abnormal BSV (BG cross) & occasional diplopia Dist: suppression
07	F	Consecutive XT	L LR recession 8mm L MR resection 8mm	Consecutive residual ET	Nr: Suppression Dist: Overcorrected & wearing 10PD Fresnel prism to put back into suppression area, gradually reducing prism
09	M	Primary XT & HT	R MR resection 5.5mm R LR recession 5mm	Consecutive ET & HT	Suppression
10	F	Consecutive XT	L LR recession 7mm L MR stretched scar resected 3mm & advanced 6mm	Consecutive residual ET with accommodative element	Nr: Unexpected sensory fusion (BG cross), -ve Wirt fly Dist: suppression
13	F	Consecutive XT	L MR advancement 5mm & pseudotendon resection 5mm	Residual consecutive XT	Suppression
15	M	Residual XT	L MR resection 9mm L LR recession 10mm (LR pulled up 1mm at adjustment)	Residual X	Unexpected BSV: Sensory fusion: Near & Dist: BG cross Motor fusion: Near 35 BO – 14 BI, Dist 12 BO – 12 BI (with BG control) Stereo: Near: 110” FNS

16	M	Consecutive XT & HT	R MR resection 3mm & advanced 5mm (8mm total) R LR recession 6mm R MR & LR insertions moved down ½ muscle width R medial conjunctiva trimmed (R LR recessed 1mm at adjustment)	Residual consecutive XT	Suppression
18	F	Consecutive XT	L MR advanced 5mm	Residual consecutive XT	Suppression
20	F	Consecutive XT	R LR recession 9mm R MR resection 8mm	Residual consecutive XT	Suppression
21	M	Consecutive XT	R LR recession 10mm R lateral conjunctiva recession 5mm R MR advanced 5mm	Residual consecutive XT	Suppression

Key to table:

M: male, F: female. L: left, R: right.

Consecutive: strabismus has changed direction, typically following surgery, example from ET to XT or from XT to ET.

Residual: strabismus remains in the same direction following surgery, but is typically a smaller angle, example from large XT to smaller XT.

Consecutive residual ET: original strabismus was an ET, then became consecutively XT, during study had Sx for XT and became consecutively ET but with a residual ET

Residual consecutive XT: original strabismus was an ET, then became consecutively XT, during study had Sx for XT and becomes a smaller XT

Nr: near, Dist: distance, BG: Bagolini glasses, PD: prism dioptres, BO: base out, BI: base in.

Appendix W. CKAT technical difficulties experienced during recording and analysis

Table W-1. CKAT technical difficulties					
	Participant number	Recording Visit 1	Recording Visit 2	Analysis Visit 1	Analysis Visit 2
Surgery	01	Y	Y *	C	P
	03	Y	Y	C	C
	05	Y	N	P	N
	07	N		N	
	09	N		N	
	10	Y *	Y	P	C
	13	Y *	Y *	P	P
	15	Y *	Y *	P	P
	16	N		N	
	18	Y *	N #	P	N #
Control	02	Y	Y *	P	P
	04	Y *	Y	P	C
	06	N		N	
	08	Y	Y *	C	P
	11	Y	Y	P	P
	12	Y *	Y	P	C
	14	Y *	Y *	P	P
	17	Y *	Y *	C	P
	23	Y	Y *	P	P
	24	N		N	
Total (data available for analysis Yes / No)		15 Yes / 5 No	13 Yes / 7 No	15 Yes / 5 No	13 Yes / 7 No

Recording: Y = yes, testing completed, no technical difficulties noted (green)

Y * = yes, testing completed, but technical difficulties noted during testing (orange)

N = unable to complete testing due to technical difficulties (red)

N # = unable to complete testing due to arthritis (red)

Analysis: C = complete data set (no errors documented by CKAT) (green)

P = partial data set (some errors documented by CKAT) (orange)

N = no data used during analysis (red)

If difficulties were experienced during the recording, these were documented in the recordings columns. If recording was completed, but on analysis of the data errors were documented by the CKAT these are documented in the analysis columns.

Results available for analysis included:

- Tracking (accuracy, RMSE mm) for the slow, medium and fast target speed without a background guide (control group n=7 slow, medium and n=6 fast; surgery group n=4) and for the slow, medium and fast target speed with a background guide (control group n=3, surgery group n=2).
- Aiming (mean path length time (seconds)) (control group n=5, surgery group n=5)
- Steering (penalised path accuracy (pPA) score) (control group n=6, surgery group n=3).

Appendix X. Sample size calculation

Performed using G*Power (Faul et al., 2007)

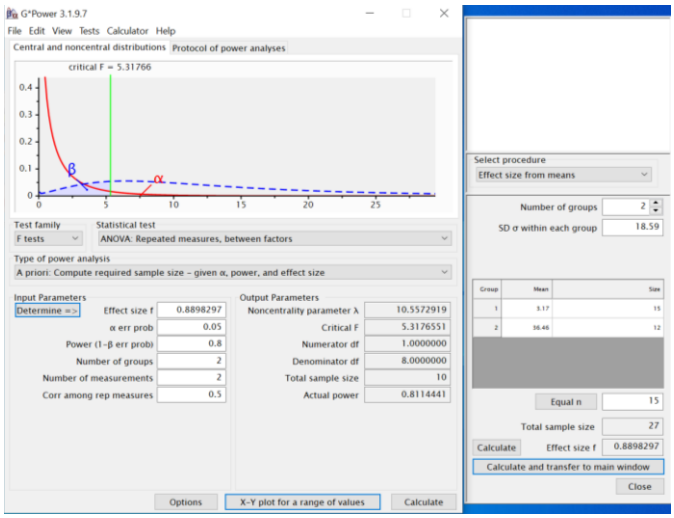
Using a two-way mixed ANOVA statistical test, with repeated measures, between factors gave the larger sample size and are presented here. Effect size was calculated from the mean change in the dependent variable from visit one to visit two in each group. The SD of the surgery group was used. α= 0.05, power= 0.80, correlation among repeated measures=0.5. Two groups (control and surgery) and two measurements were used (before and after surgery).

Using change in AS-20 (overall score) values from visit one to visit two.

Descriptive Statistics

Dependent Variable: AS20_overall_V2 - V1

Group	Mean	Std. Deviation	N
Control	3.1667	5.66684	15
Surgery	36.4583	18.58513	12
Total	17.9630	21.15689	27

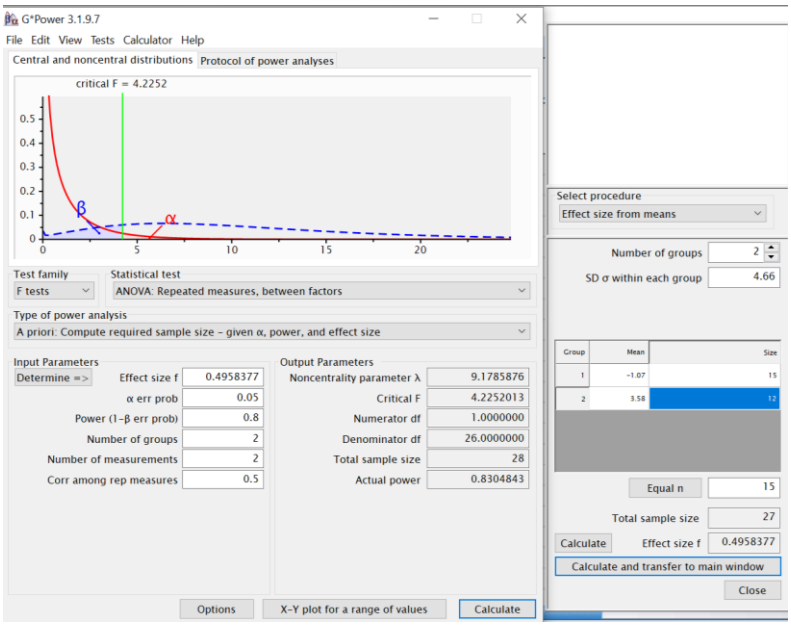


Sample size of n=10, split between two groups (control group n=5 and surgery group n=5)

Using change in binocular summation score (100% contrast) from visit one to visit two

Descriptive Statistics

	Group	Mean	Std. Deviation	N
Binocular_summation_100%_V2 - V1	Control	-1.0667	3.45309	15
	Surgery	3.5833	4.66044	12
	Total	1.0000	4.59933	27
Binocular_summation_10%_V2 - V1	Control	.5333	3.02056	15
	Surgery	-.6667	2.70801	12
	Total	.0000	2.89562	27
Binocular_summation_5%_V2 - V1	Control	1.1333	3.70071	15
	Surgery	-.0833	4.20948	12
	Total	.5926	3.90522	27

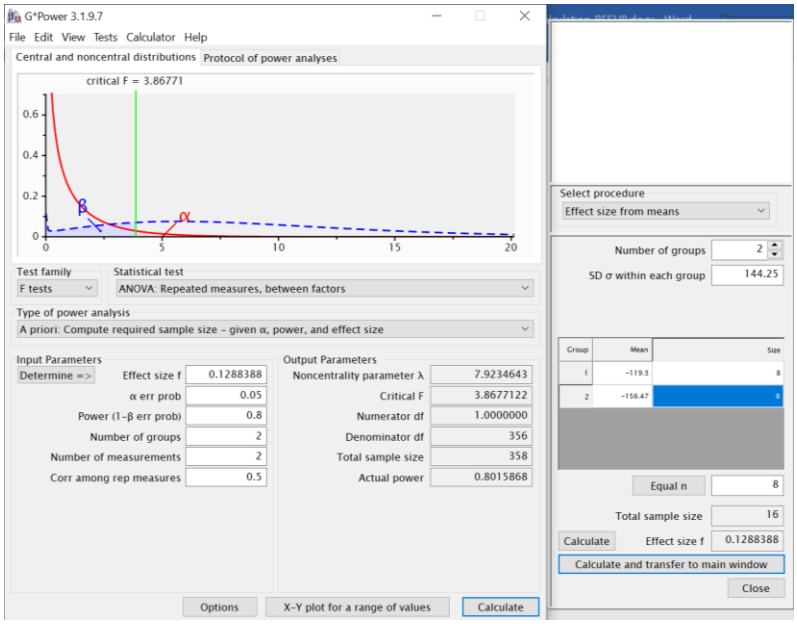


Sample size of n=28, split between two groups (control group n=14, surgery group n=14)

Using change in TSL time (msec) from visit one to visit two.

Descriptive Statistics

Group		N	Mean	Std. Deviation
Control	Mean_change_TSL_time(V2-V1)	8	-119.2992	198.28562
	Valid N (listwise)	8		
Surgery	Mean_change_TSL_time(V2-V1)	8	-156.4662	144.24988
	Valid N (listwise)	8		

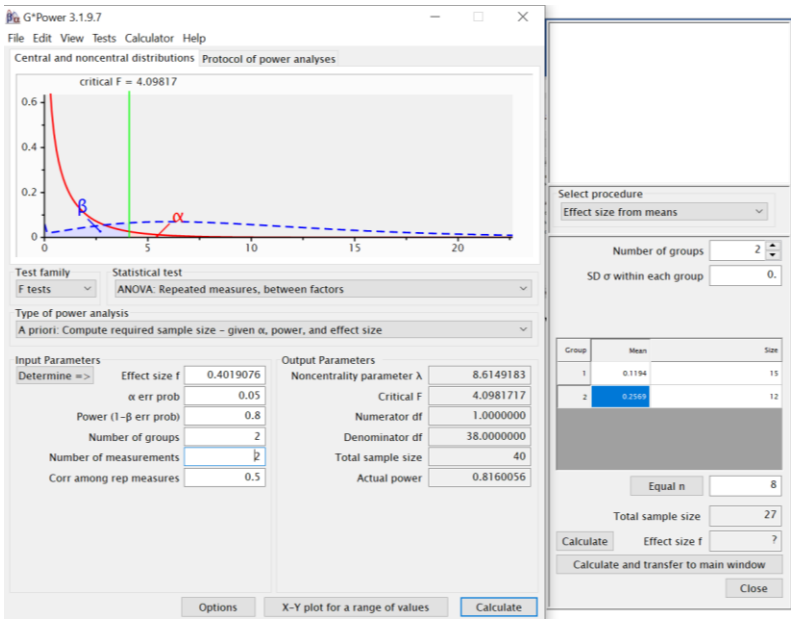


Sample size of n=358, split between two groups (control group n=179, surgery group n=179)

Using change in CST RCS BEO from visit one to visit two

Descriptive Statistics

Group		N	Mean	Std. Deviation
Control	Change_CST_RCS_BEO V2 - V1 (for sample size calc)	15	.1194	.16625
	Valid N (listwise)	15		
Surgery	Change_CST_RCS_BEO V2 - V1 (for sample size calc)	12	.2569	.17357
	Valid N (listwise)	12		



Sample size of n=40, split between two groups (n=20 control, n=20 surgery)