Beyond visual acuity
Development of visual processing speed and quantitative assessment of visual impairment in children

Annemiek D. Barsingerhorn
Beyond visual acuity: Development of visual processing speed and quantitative assessment of visual impairment in children

Annemiek D. Barsingerhorn
Colofon

The research presented in this thesis was carried out at the Donders Institute for Brain, Cognition and Behaviour, Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, Netherlands, in collaboration with Bartiméus, Institute for the Visually Impaired, Zeist, The Netherlands, with financial support from the ODAS foundation (grant Uitzicht-2012-35, awarded to dr. F.N. Boonstra and dr. J. Goossens). Printing of this thesis was financially supported by the Radboud University Medical Center and Donders Institute for Brain, Cognition and Behaviour.

ISBN: 978-94-6284-161-1

Cover illustration: Marlies Barsingerhorn
Design/lay-out: Annemiek Barsingerhorn
Printed by: Gildeprint - Enschede, The Netherlands

© Annemiek D. Barsingerhorn, 2018, Nijmegen, The Netherlands

All rights reserved. No part of this book may be reproduced, distributed or transmitted in any form or by any means, without prior written permission of the author.
Beyond visual acuity: Development of visual processing speed and quantitative assessment of visual impairment in children

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 31 oktober 2018
om 14.30 uur precies

door
Annemiek Dorien Barsingerhorn
geboren op 9 juni 1988
te Delfzijl
Promotor
Prof. dr. A.J. van Opstal

Copromotoren
Dr. H.H.L.M. Goossens
Dr. F.N. Boonstra

Manuscriptcommissie
Prof. dr. R.J.A. van Wezel
Prof. dr. J. van der Steen (Erasmus MC, Rotterdam)
Prof. dr. F.W. Cornelissen (UMCG)
# Table of contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Introduction</td>
<td>7</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Development of symbol discrimination speed in children with normal vision</td>
<td>23</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Symbol discrimination speed in children with visual impairments</td>
<td>53</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Optics of the human cornea influence the accuracy of stereo eye-tracking methods: a simulation study</td>
<td>79</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Development and validation of a high-speed stereoscopic eye tracker</td>
<td>101</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Saccade latencies during a preferential looking task and objective scoring of grating detection in children with and without visual impairments</td>
<td>141</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Summary and General Discussion</td>
<td>165</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>References</td>
<td>181</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Nederlandse Samenvatting</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Dankwoord</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>Publications and Presentations</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>Biography</td>
<td>223</td>
</tr>
<tr>
<td></td>
<td>Donders Graduate School for Cognitive Neuroscience</td>
<td>224</td>
</tr>
</tbody>
</table>
1

Introduction
For the majority of tasks in daily life we rely on visual information. The demands on our visual system become even greater in dynamic situations, for example in traffic, where visual information changes rapidly over time. Most of young adults and children are able to deal with this information efficiently. However, if it takes too long to process and respond to visual stimuli, problems can occur. In addition, people with visual impairments often complain about the effort it takes to perform everyday tasks, and they need more time to perform these tasks. Furthermore, reduced speed of processing in children with low vision has been reported to be a limiting factor in educational settings. However, existing tests of visual acuity do not take into account the time that is needed to respond and the definition of visual impairment of the WHO is only based on visual acuity and visual field deficits.

This thesis focuses on the quantification of visual processing speed, both in children with normal vision and children with visual impairments. In addition, a new eye tracker was developed, to assess oculomotor behavior in these children. In the following sections, standard tests to measure visual performance are described, as well as the criteria for visual impairment and the prevalence and etiologies of visual impairment in children. Subsequently, I will give an overview of research on visual processing speed. In addition, I will review oculomotor functions and eye tracking methods that can be used to assess these functions. Furthermore, I will discuss different visual functions and their development throughout childhood. Finally, I will present the outline of this thesis.

1.1 Standard tests to measure visual performance

About 150 years ago, Herman Snellen and Franciscus Cornelis Donders provided the bases for standardized testing of visual acuity. Visual acuity (VA) tests measure the ability of the visual system to perceive detail (e.g., the gap in the Landolt C). The VA score is determined by the smallest optotypes that can be identified at a given distance. Numerous tests are available to measure VA, using a wide variety of optotypes, such as letters (e.g. in the Snellen test, HOTV test and EDTRS test), numbers, Landolt C, tumbling E, Lea figures and other symbols, and gratings (for an overview of different optotypes and tests see e.g. ISO 8596 standard). The International Visual Acuity Chart Design Guidelines and ISO 8596 standard indicate that the Landolt-C is the standard optotype (see Figure 1.1A, for the Landolt-C chart), which should be used as a reference when designing new optotypes and/or charts to ensure that the results are equivalent to the Landolt-C test. Most of the standard validated tests to measure VA exist of rows of black high-contrast optotypes of different sizes on a
white background, which can either be presented at a printed chart or on a computer screen with sufficient temporal and spatial resolution.

In children, the choice for a specific VA test is based on the developmental level of the child. Optotype based testing is unsuitable for infants, toddlers and non-verbal children, because they cannot reply in a verbal manner and are not capable of matching an answer to a stimulus. In these populations, the preferential looking technique is the most common clinical method to assess visual acuity\textsuperscript{15}. Preferential looking tests, such as the Teller acuity card test (TAC, see Figure 1.1B), use the infant’s preference to look at a patterned target rather than at a blank target\textsuperscript{16}. The grating with the finest stripes that produces a consistent orienting response provides an estimate of the child’s visual acuity\textsuperscript{17,18}. Most preschool children can complete VA tests that require either naming or matching of optotypes\textsuperscript{12,19,20}. For these young children charts with symbols, such as the LEA-symbols\textsuperscript{21} which consist of a circle, square, apple and house (see Figure 1.1C), are often preferred\textsuperscript{12,19,20}. For older children and adults charts that consist of letters or numbers are most used in clinical settings\textsuperscript{12,19,20}. It is difficult to directly compare the outcome of different charts, because not all charts meet the International Visual Acuity Chart Design Guidelines\textsuperscript{13,20}.

Multiple notations exist for expressing the VA, but in general they use the optotype size at which the detail subtends one minute of arc as a reference value. Snellen notations consist of the test distance divided by the distance at which the gap of the equivalent Landolt C subtends one minute of arc, and can be expressed as a fraction (e.g. 6/6) or in decimal notation (e.g. 1.0). LogMAR units (Logarithm of the Minimum Angle of Resolution) are the logarithm (base 10) values of the visual angle in minutes of arc of the gap in the equivalent Landolt C optotype. The LogMAR value of an optotype at which the detail subtend one minute of arc is 0.0 (for more

![Figure 1.1. Examples of different visual acuity charts. A. Landolt-C test, B. Teller Acuity Cards, C. LEA-symbols.](image-url)
information about the different notations see\textsuperscript{13,22}). Grating acuity is often expressed in cycle/deg, but can also be expressed in Snellen notation or in LogMAR.

The Visual Functions Committee of the International Council of Ophthalmology recommends in the International Visual Acuity Chart Design Guidelines that VA tests are performed under standard conditions with sufficient light intensity (minimum chart luminance of 80 cd/m\textsuperscript{2}) with the person seated at a fixed distance (4, 5 or 6 meters) from the presented optotypes and that the size of the optotypes decreases with a factor of 0.1 log unit per row, with preferably a minimum of 5 optotypes per row\textsuperscript{13}. The spacing between the optotypes can be proportional to the size of the optotypes, or can be a fixed distance independently of the size of the optotypes (see\textsuperscript{23} for examples of the different charts).

\section*{1.2 Visual impairment in children}

The International statistical Classification of Diseases and related health problems (ICD-10) defines low vision as visual acuity in the best eye (with correction) of better than or equal to 0.05 ($\geq 1.3$ LogMAR) and less than 0.3 (0.5 LogMAR)\textsuperscript{6}. The underlying causes and the prevalence of visual impairment in children differ widely per region\textsuperscript{15,24–26}. The prevalence of visual impairment in children aged 0-15 years is estimated to range from 0.1/1000 children in lower-income countries to 1.1/1000 children in higher-income countries\textsuperscript{24}. Most often, visual impairment results from ocular diseases and/or genetic factors, such as albinism, optic atrophies, retinal dystrophies, congenital cataract, and retinopathy of prematurity\textsuperscript{25,27,28}. However, in higher-income countries cerebral visual impairment (CVI) is considered to be the most prevalent cause of visual impairment\textsuperscript{25–27,29,30}. The current definition of CVI includes all visual dysfunctions caused by damage to, or malfunctioning of, the post-chiasmatic visual pathways\textsuperscript{31}. However, pre-chiasmatic pathways can be involved in CVI as well and CVI can occur together with ophthalmological diseases, such as cataract or retinopathy of prematurity. CVI can be caused by a wide range of developmental disorders and brain injuries; for instance, hypoxia, malformations of cortical development, preterm birth, closed head trauma, encephalitis, and genetic disorders\textsuperscript{31,32}. The diversity of underlying causes of CVI results in deficits in a wide range of visual functions\textsuperscript{31,33,34}. Problems may arise in higher visual functions, such as visual search and visual attention, and/or in lower visual functions, such as contrast sensitivity, visual field or visual acuity. Higher visual functions, such as motion processing, visual search, and visual attention, may be impaired in children with visual impairments due to ocular diseases and/or congenital disorders as well\textsuperscript{35–37}.
1.3 Visual processing speed

Non-timed visual acuity tests may not reflect the demands in activities of daily life. However, the only reference to time during VA tests in the VA measurement standard is that during validation of a new VA test the optotype has to be exposed for a maximum of 3 seconds and that the judgment period should not exceed 4 seconds. However, for validated VA tests no recommendations are made and none of the existing VA tests takes the perception time into account. Perception time is defined as the time a person needs to identify a stimulus from the moment that it is presented. Perception time in the visual domain is often referred to as visual processing speed.

The use of response times to study mental and perceptual processes has a long history, which started in 1868 with Donders’ seminal work on the time to carry out mental subprocesses. Since then, various tasks and stimuli have been used to study the speed of visual processes (for extensive overviews see). Response time, or reaction time, is defined as the time between stimulus presentation and the response to that stimulus. Classic reaction time paradigms include simple visual reaction time, in which the participant has to respond as soon as a visual stimulus is detected, and choice reaction time, in which two or more stimuli are used and the participant has to make a different response to each stimulus (e.g. push the left mouse button if the stimulus is a square, and push the right mouse button if the stimulus is a circle). In standard reaction time paradigms the stimuli are presented until a decision is made. In contrast, in inspection time paradigms the stimulus duration is varied and the minimal stimulus duration at which participants can perform a task accurately is determined. In the classic inspection time task the stimulus consists of two vertical lines of different lengths which are connected at their tops by a horizontal line. The participant has to indicate at which side the longer line appears.

Recently several studies addressed the issue of visual processing speed, especially in relation to driving performance and/or traffic accident proneness. One test often used in these studies is the Useful Field of Vision test (UFOV®, Visual Resources, Inc., Chicago, Illinois). Subtest one of the UFOV® test was designed to measure visual processing speed and subtest two and three were designed to measure aspects of divided and selective attention. In the UFOV® test, the duration of the stimulus is varied. The outcome can therefore be considered as the stimulus duration threshold instead of the perception time (cf.). For older adults, poor results on the UFOV® test have been related to an elevated risk of at-fault traffic accidents and decreased...
Introduction

Although studies using the UFOV® test advanced our knowledge about the relevance of such tests for diagnostic purposes, a recent meta-analysis revealed that a broad range of perceptual and cognitive functions is related to UFOV® performance. This suggests that the UFOV® subtests do not measure one clear construct, as they were designed to do. Furthermore, the UFOV® does not take into account the visual acuity of the subjects and the size of the stimuli remain the same during the test. Finally, visual task performance is generally characterized by speed-accuracy trade-offs, with less accurate responses when responding quickly and more accurate responses when taking more time to respond. The UFOV® does not take into account this aspect of visual performance.

The presence of speed-accuracy tradeoffs in visual task performance suggest that tests which combine speed and accuracy measures provide a better quantitative assessment of visual impairment. This is supported by studies demonstrating relationships between the speed of visual processing, and stimulus size and/or visual acuity. Detecting smaller squares takes longer than detecting larger squares, and visual evoked potentials (VEP) and reaction times are faster for coarser gratings. In addition, the outcome of different visual processing speed tests, neuropsychological and cognitive tests are correlated with visual acuity. Moreover, visual degrading filters significantly increased the time participants needed to complete the Digit Symbol Substitution Test and the Trail Making Test, despite the fact that the printed targets used for the tests were still well above their visual acuity. In addition, the relationship between speed and accuracy works the other way around as well: the accuracy of responses decreases with decreasing exposure duration. Furthermore, the strength of sensory stimulation (e.g. level of contrast) affects the speed and accuracy of perceptual judgments. The same effect can be observed in reading, with decreasing reading speeds with decreasing print size.

Adults and children with visual impairments need more time for visual tasks compared to people with normal vision. Especially for children with CVI it is often reported that their visual responses are abnormally late, which was confirmed by studies demonstrating that children with CVI have delayed orienting responses. Furthermore, longer search times have been reported for children with CVI. For children with visual impairments due to ocular disorders and/or retinal abnormalities longer search times have been reported as well, and they often need more time to complete exercises at school. In addition, children with visual impairments, children with amblyopia, and adults with visual impairments have lower reading performance on other aspects of daily living. Although studies using the UFOV® test advanced our knowledge about the relevance of such tests for diagnostic purposes, a recent meta-analysis revealed that a broad range of perceptual and cognitive functions is related to UFOV® performance. This suggests that the UFOV® subtests do not measure one clear construct, as they were designed to do. Furthermore, the UFOV® does not take into account the visual acuity of the subjects and the size of the stimuli remain the same during the test. Finally, visual task performance is generally characterized by speed-accuracy trade-offs, with less accurate responses when responding quickly and more accurate responses when taking more time to respond. The UFOV® does not take into account this aspect of visual performance.

The presence of speed-accuracy tradeoffs in visual task performance suggest that tests which combine speed and accuracy measures provide a better quantitative assessment of visual impairment (cf.). This is supported by studies demonstrating relationships between the speed of visual processing, and stimulus size and/or visual acuity. Detecting smaller squares takes longer than detecting larger squares, and visual evoked potentials (VEP) and reaction times are faster for coarser gratings. In addition, the outcome of different visual processing speed tests, neuropsychological and cognitive tests are correlated with visual acuity. Moreover, visual degrading filters significantly increased the time participants needed to complete the Digit Symbol Substitution Test and the Trail Making Test, despite the fact that the printed targets used for the tests were still well above their visual acuity. In addition, the relationship between speed and accuracy works the other way around as well: the accuracy of responses decreases with decreasing exposure duration. Furthermore, the strength of sensory stimulation (e.g. level of contrast) affects the speed and accuracy of perceptual judgments. The same effect can be observed in reading, with decreasing reading speeds with decreasing print size.

Adults and children with visual impairments need more time for visual tasks compared to people with normal vision. Especially for children with CVI it is often reported that their visual responses are abnormally late, which was confirmed by studies demonstrating that children with CVI have delayed orienting responses. Furthermore, longer search times have been reported for children with CVI. For children with visual impairments due to ocular disorders and/or retinal abnormalities longer search times have been reported as well, and they often need more time to complete exercises at school. In addition, children with visual impairments, children with amblyopia, and adults with visual impairments have lower reading performance on other aspects of daily living. Although studies using the UFOV® test advanced our knowledge about the relevance of such tests for diagnostic purposes, a recent meta-analysis revealed that a broad range of perceptual and cognitive functions is related to UFOV® performance. This suggests that the UFOV® subtests do not measure one clear construct, as they were designed to do. Furthermore, the UFOV® does not take into account the visual acuity of the subjects and the size of the stimuli remain the same during the test. Finally, visual task performance is generally characterized by speed-accuracy trade-offs, with less accurate responses when responding quickly and more accurate responses when taking more time to respond. The UFOV® does not take into account this aspect of visual performance.
Chapter 1

speeds\textsuperscript{79}. However, these factors are not taken into account in criteria for rehabilitation services\textsuperscript{80}. It has been acknowledged that criteria purely based on visual acuity and visual field often fail to capture relevant vision-related problems in daily life\textsuperscript{1,80–82}. But, if people appear to have reduced visual processing speed and their visual acuity and visual field are within normal limits, they cannot receive assistance from institutes for the visually impaired or (visual) rehabilitation facilities\textsuperscript{80}.

1.4 Oculomotor behavior

Visual information processing and oculomotor control are inseparable\textsuperscript{83}. High visual acuity is restricted to the fovea: a small region close to the center of the retina with the highest density of cone receptors\textsuperscript{84}. We continuously move our eyes, to redirect our fovea to relevant features in our environment. In addition, we need to stabilize our gaze during head movements. Stabilization is achieved by the vestibular-ocular reflex (VOR), and the optokinetic nystagmus (OKN)\textsuperscript{83,85}. The VOR keeps the retinal image stable during brief and accelerating head movements, by moving the eyes at the same speed in the opposite direction of the head movement\textsuperscript{83,85,86}. The OKN holds the retinal image stable during prolonged head movements, or when our surround moves rapidly over the retina (e.g. during self-motion), by alternating between periods in which the retinal pattern is followed (slow phase), and quick jumps in the opposite direction to reorient the gaze (quick phase)\textsuperscript{83,85}. Eye movements which allow us to perceive objects of interest with the highest visual acuity are smooth pursuit, saccades and fixations. Smooth pursuit movements are used to track moving objects, to keep the object within, or close to the foveal region\textsuperscript{85,87}. Saccades are rapid gaze shifts, to direct the fovea sequentially from one selected location to another\textsuperscript{83,85,87}. During fixations the gaze is focused on a specific target for at least 80-100 ms\textsuperscript{88}. Even during fixation the eyes are not completely stationary\textsuperscript{87,89,90}. The classical fixation pattern consists of microsaccades, drift and tremor\textsuperscript{87,89,90}. Microsaccades are small saccades of \textasciitilde12-15 minutes of arc. Drifts are slow movements (4 minutes of arc/s) occurring in the intersaccadic interval with an amplitude of \textasciitilde1.5 to 4 minutes of arc. Tremors are rapid oscillatory movements with amplitudes of \textasciitilde5 to 30 seconds of arc with frequencies up to about 200 Hz\textsuperscript{87,89,90}.

Abnormal eye movements are observed in many diseases of the central nervous system, such as in Parkinson’s disease, Alzheimer’s disease, Huntington’s disease, and spinocerebellar ataxias\textsuperscript{91–94}. Properties of saccades and fixations can provide diagnostic data for the identification of these diseases, can help to distinguish between different subtypes (e.g. different spinocerebellar ataxias), and/or allow
to localize the disturbance to a specific area in the brain, such as the cerebellum or brainstem. Saccade abnormalities include hypometric (undershooting) and hypermetric (overshooting) saccades, reduced accuracy in memory guided saccades, reduced saccade velocity, and increased saccade latencies. Impaired visual fixation include saccadic intrusions, saccadic oscillations and nystagmus. Square wave jerks are the most common type of saccadic intrusions and consist of small saccades ranging from 0.5 to 5 deg, which take the eye from the fixation point and return gaze to the target after about 200ms. Saccadic oscillations consist of ocular flutter, which are intermittent bursts of high frequent horizontal oscillations with amplitudes ranging from 1 to 5 deg, and opsinclonus, which are multidirectional high frequent oscillations of varying amplitudes. Nystagmus is another form of repetitive involuntary ocular oscillations. Contrary to saccadic intrusions and saccadic oscillations, nystagmus is initiated by slow phases instead of fast saccadic movements. Nystagmus can be idiopathic, or can be related to albinism or ocular diseases such as retinal dystrophies, or caused by a wide range of neurological diseases.

In clinical practice, assessment of eye movements is in general based on visual inspection, often resulting in a qualitative description of abnormalities (e.g. type of nystagmus, smooth or saccadic smooth pursuit). Objective eye movement recordings allows clinicians to quantify characteristics of eye movements (e.g. saccade velocity), and to obtain objective records which can be used to monitor changes in oculomotor behavior over time. However, in current clinical practice objective eye movement recordings remain relatively rare, mainly because most clinicians do not have access to eye tracking methods, or because the available methods are not suitable for use in patients and/or children. In the following section, different eye tracking methods will be discussed.

1.5 Eye tracking methods

Near the end of the 19th century the first objective eye recording techniques were developed, which used a variety of mechanical attachments to the eye to translate the movements of the eye to a recording surface. Variants of attachment eye trackers that monitored light reflected from mirrors which were attached to the eye were developed throughout the 20th century, including Yarbus’ suction caps which he used in his seminal work and contact lens methods (see e.g. for an extensive overview). In 1963 Robinson introduced a novel contact lens method: the scleral search coil, which was further refined by Collewijn and colleagues. A copper
Chapter 1

coil is embedded in a contact lens and the scleral search coil technique measures electromagnetic induction in this coil\textsuperscript{104,105}. The scleral search coil is still one of the most precise and accurate eye tracking methods, which allows studying even the smaller eye movements. However, due to its invasive and uncomfortable nature and a limited recording time (typically about 30-45 min) it is unsuitable for eye tracking in children and for most clinical settings\textsuperscript{106}. In electro-oculography (EOG) there is no attachment to the eye, but electrodes are placed on the skin. These electrodes measure the electrical potential of the cornea-retinal dipole using electrodes placed on the skin\textsuperscript{100,107}. The advantage of EOG is that it can also monitor eye movements when the eyes are closed and has therefore been used widely in studies that have assessed eye movements during sleep\textsuperscript{108,109}. However, EOG has relatively low accuracy and precision, because it is prone to noise\textsuperscript{100}.

An alternative to attachment eye trackers are optical eye trackers, which record the light that is reflected from the surface of the eye. In 1901, Dodge and Cline were the first to develop photographic techniques, which recorded the movement of the corneal light reflection on a moving photographic plate\textsuperscript{100,110,111}. At present, most commercial eye trackers, such as the Eyelink and Tobii, are video-based. They use the reflection of (IR) light on the cornea, often in combination with pupil tracking (see e.g.\textsuperscript{88,112,113} for extensive overviews). A special type of video-based eye tracker is the dual Purkinje eye tracker (DPI)\textsuperscript{114–116}. Light is not only reflected by the cornea, but also by other structures within the eye, such as the external and internal surface of the lens. These reflections are called Purkinje images. The DPI tracks the first (reflection external surface of the cornea), and fourth (reflection internal surface of the lens) Purkinje images, which result in highly accurate eye movements recordings\textsuperscript{114–116}.

Video-based methods are non-invasive. However an individual calibration procedure is needed, in which participants have to fixate several small visual targets at known locations, in order to convert the image features of the eyes into estimates of the point of gaze (POG) on the screen\textsuperscript{88}. Most normally-sighted adults are able to perform such a calibration procedure effortlessly. However, if participants are unable to reliably fixate small stimuli, for example in the case of oculomotor deficits, low vision, or reduced attention, accurate calibration is not possible. To simplify the calibration procedure, researchers have developed video-based eye trackers with two cameras: stereo eye trackers\textsuperscript{117–122}. With these eye trackers the position of the eye and the orientation of its optical axis can be estimated directly based on the eye features and the known geometry of the set-up. As a result, only the
deviation between the optical and visual axes needs to be determined, which can be done through a one point calibration procedure. The spatial accuracy of the developed prototypes is ~1 deg or lower\textsuperscript{120–124}, which is acceptable for a wide range of eye tracking applications. However, the sampling rate is 20-30 Hz\textsuperscript{120–124}, which is insufficient for most eye tracking applications as it does not allow studying the kinematics of eye movements\textsuperscript{125–127}.

### 1.6 Visual development

Human infants are born with an immature visual system, both functionally as well as anatomically\textsuperscript{15,128}. Especially during the first few months after birth, the visual system develops rapidly, but this maturation continues during childhood\textsuperscript{129,130}. Development of ocular structures includes differentiation and maturation of the fovea\textsuperscript{131,132} and growth of the eyeball\textsuperscript{133}. In addition to structural developments, visual functions improve with age. For example, VA increases rapidly from ~0.01 - 0.05 (1.3-2.0 LogMAR) at birth to ~0.40 (0.4 LogMAR) at 12 months\textsuperscript{15,134}, and fully matures between 6 and 10 years of age\textsuperscript{135–137} (Figure 1.2). Maturation of VA is faster

![Figure 1.2. The development of visual acuity throughout childhood. Different tests were used to test visual acuity: TAC (Salomao et al., 1995\textsuperscript{134}), HOTV monocular (Pan et al., 2010\textsuperscript{166}), EDTRS monocular (Dobson et al., 2009\textsuperscript{137}), tumbling E binocular (Huurneman et al., 2012\textsuperscript{219}, and Jeon et al., 2010\textsuperscript{218}), and binocular Landolt-C test (Lai et al., 2011\textsuperscript{157}).](image-url)
for isolated optotypes compared to optotypes surrounded by other optotypes or flanking bars\textsuperscript{138,139}. This phenomenon, in which recognition is impaired due to surrounding contours, is called crowding. In addition, the visual field expands as children grow older, with a rapid increase of the visual field between 2-8 months of age and a slower increase throughout childhood until adult visual field sizes are reached at approximately 12 years of age\textsuperscript{15,140,141}. Contrast sensitivity\textsuperscript{135,142} and visual motion perception\textsuperscript{143,144} improve with age as well. Not only the acuity and sensitivity of the visual system improve with age, children also become faster on visual tasks. For instance, reading speed increases with age\textsuperscript{66,67} and the time needed for visual search tasks\textsuperscript{145–147} and visual matching of letters or numbers\textsuperscript{148–150} decreases as children grow older. In addition, stimulus-duration thresholds on inspection time tasks, including the UFOV, decrease with age\textsuperscript{151–153}. Furthermore, simple visual reaction times\textsuperscript{153–157} become faster as children grow older, as well as choice reaction times for identifying animal pictures\textsuperscript{154} or numbers\textsuperscript{156}.

Developmental aspects of oculomotor behaviour are widely documented for different characteristics of eye movements (for an extensive overview see\textsuperscript{158}). The stability of fixations increase with age due to a decrease in the occurrence of intruding saccades\textsuperscript{159,160}. In addition, saccade latencies decrease with age through childhood, until adult levels are reached at approximately 10 to 12 years of age\textsuperscript{161–164}. Furthermore, error-rates in antisaccade tasks, in which participants have to make a saccade in the opposite direction of the target, decrease with age until approximately 15 years of age\textsuperscript{162,163,165}.

1.7 Thesis outline

A child’s ability to distinguish visual details fast and accurately is important for his/her participation in school and society. Although research has revealed that children become faster on visual tasks as they grow older (section 1.6), it is unknown how visual processing speed develops during childhood. In addition, reduced speed of processing has been reported to be a limiting factor in educational settings for children with visual impairment, and children with visual impairments need more time on visual tasks (section 1.3). Furthermore, despite the accumulating evidence that purely spatial approaches to vision are insufficient to screen for visual impairment, the temporal aspects of vision still receive very little attention in clinical practice. Thus far there are no studies that quantified visual processing speed in children with visual impairments, to assess whether they are slower in discerning visual details than children with normal vision. This thesis aims to answer these questions. That
is, i) to quantify the development of visual processing speed in children with normal vision and, ii) to determine whether children with visual impairment are slower in discerning visual details than children with normal vision. To that end, new methods were developed to assess the speed and accuracy of visual processes simultaneously and to quantify oculomotor behavior in these children.

The thesis consists of three parts. **Part 1** addresses symbol discrimination speed in children with and without visual impairments, **Part 2** describes the development of a stereoscopic eye tracker, and **Part 3** focuses on saccade latencies during a preferential looking task in children with and without visual impairments. The eye tracker which was developed and tested in **Part 2** was used to quantify the saccade latencies in **Part 3**.

**Part 1: Symbol discrimination speed in children with and without visual impairments**

In the first part of the thesis we investigated the speed of symbol discrimination in children with and without visual impairments. Because tests that combine speed and accuracy measures could provide a better quantitative assessment of visual impairment (section 1.3), we used a combined symbol-discrimination reaction-time test to assess visual acuity and visual discrimination speed simultaneously.

The developmental effects on reading speed and the time children need for other visual tasks (section 1.6) could in part result from the development of the speed at which basic symbols are distinguished. However, to our knowledge, it is unknown how symbol discrimination speed develops during childhood. Therefore, we determined how fast children with normal vision can discern foveal stimuli and how this ability improves with age in **Chapter 2**, using advanced psychophysical analyses.

Reductions in visual processing speed might explain at least part of the problems in higher visual functions in children with visual impairments (section 1.2 and 1.3). Therefore, we determined whether the symbol discrimination of children with visual impairment is slower than in children with normal vision in **Chapter 3**. In addition, we investigated whether delays in symbol discrimination may be explained by their reduced acuity alone. This could have important implications for rehabilitation and educational settings, as it provides insight in whether or not magnification may help the children to compensate for reduced visual processing speed. The results of **Chapter 2** were used as normative data this study.
Part 2: Development of a high speed stereo eye-tracker

Recording eye movements in children with visual impairments is challenging. Individual calibration is often not possible or unreliable, due to problems with fixation of the targets on the screen as a result of low vision and/or fixation instability in case of nystagmus. We were interested in the oculomotor behavior of children with visual impairments, but commercially available eye trackers did not serve our needs due to the need for extensive individual calibration procedures, low temporal resolution and/or insufficient precision (section 1.5). Therefore, we decided to develop our own stereoscopic eye tracker.

Current stereo eye-tracking methods model the cornea as a sphere. However, the human cornea is slightly aspheric. In Chapter 4 we used simulations to investigate how the optics of the aspheric cornea influence the accuracy of stereo eye tracking methods. The development and validation of the stereoscopic eye tracker is described in Chapter 5. Our aim was to develop a stereo eye tracker with a sampling rate of at least 250 Hz, which is sufficient to analyze kinematics of eye movements and to determine saccade latencies reliably. In addition, we validated the accuracy of our stereo eye tracker against the accuracy of an established eye tracker.

Part 3: Saccade latencies in children with and without visual impairments

In Part 3 we further explored how visual processing speed develops during childhood and whether visual processing is slower in children with visual impairments. Because eye movements and visual processing are inseparable and eye movement recordings can provide valuable diagnostic insights (section 1.4), we assessed the oculomotor behavior of children with and without visual impairments during a preferential looking task (section 1.1). In Chapter 6, our goal was to determine how saccade latencies improve with age, and whether the onset of orienting responses was delayed in the children with visual impairments. The same children with and without visual impairments as in Chapter 2 and 3 participated in this study, and the saccade latencies were measured using the stereoscopic eye tracker from Chapter 5.
Development of symbol discrimination speed in children with normal vision

Development of symbol discrimination speed

Introduction

For many tasks in daily life we rely strongly on visual information. In children the visual system is still developing. Developmental effects have been found on a wide range of visual functions, such as visual acuity\textsuperscript{130,135,137,168}, visual field\textsuperscript{140}, contrast sensitivity\textsuperscript{135}, visual span\textsuperscript{169}, crowding\textsuperscript{138}, and perception of movement\textsuperscript{143} as well as on anatomical structures, such as the retinal thickness and the size of the optic disk\textsuperscript{170}. Furthermore, reading acuity and reading speed, as measured with the MNREAD, develop with age\textsuperscript{66}. Not only is reading slower in children, also for other visual tasks children need more time than adults. For instance, young children need more time for visual search tasks\textsuperscript{145,146} and visual matching\textsuperscript{148,149} compared to older children or adults. The question we address in this study is whether these developmental effects on the time children need for reading and other visual tasks could in part result from the development of the basic ability to discern visual details fast.

When assessing the speed of visual processing, it is important to take into account the visual acuity of the participant and the size of the stimuli. Visual acuity correlates with the outcome of different visual processing speed tests, as well as neuropsychological and cognitive tests\textsuperscript{55–58}. Furthermore, both reaction times and visual evoked potentials (VEP) are slower for higher spatial frequencies of visual grating stimuli\textsuperscript{52,53} and reaction times for detecting large squares are shorter than reaction times for detecting smaller squares\textsuperscript{51}. Not only is the speed of the perceptual judgement influenced by the strength or size of the stimulus, the accuracy of the response is also influenced by the exposure duration\textsuperscript{60,61}. Moreover, a large number of studies have demonstrated that both the speed and the accuracy of a perceptual judgment depend on the strength of the sensory stimulation\textsuperscript{62–65}. The same effect can be observed for reading speed. A large number of studies in children and adults showed an increase in reading speed as print size increased\textsuperscript{66–68}.

In clinical practice the temporal aspects of vision receive little attention. The definition of visual impairment (VI) is based on visual acuity (VA) and visual field size\textsuperscript{6} and consequently these aspects of vision are tested most often. However, purely spatial approaches to vision or purely temporal approaches to vision are insufficient, as they cannot be easily separated\textsuperscript{171}. Furthermore, it has been argued that non-timed visual acuity tests may not reflect the demands in activities of daily life\textsuperscript{38}. Studies on visual processing speed in older adults have revealed that slower visual processing is related to an elevated risk of at-fault traffic accidents and decreased performance on other aspects of daily living (for an overview, see e.g.\textsuperscript{1,45}). Therefore,
it has been argued that measurements which take speed and accuracy into account are key to a better quantitative assessment of visual impairment\textsuperscript{30}.

The studies mentioned above stress the importance of assessing visual acuity and processing speed simultaneously. The advantage of the MNREAD is that it allows the assessment of reading speed and reading acuity simultaneously. However, younger children are unable to read, which limits the age range for which the test can be used. Moreover, factors such as visual span and crowding which develop in childhood are predictors of reading speed\textsuperscript{169,172} and therefore likely to influence the results on the MNREAD in school-aged children. Therefore, a test that allows assessment of visual recognition acuity and visual recognition speed is more suitable for children as it could be used in a wider age range and would eliminate confounding factors such as crowding, visual span, or reading fluency. Additionally, such a test provides insight in whether one of the underlying mechanisms for the developmental effects on reading speed could be the development of fast visual discrimination abilities.

Our goal was to investigate developmental effects on the ability to discern visual details fast by measuring visual acuity and visual discrimination speed simultaneously. To that end, we quantified how fast and how accurate 5-12 year-old children discriminate optotypes of different sizes using a speed-acuity test in which the children had to indicate, as fast and accurately as possible, the orientation of a Landolt-C symbol. To quantify the effects of age on the reaction times, we used a drift-diffusion model\textsuperscript{62} to account for the nonlinear relationship between reaction time and optotype size. We also measured reaction times on a visual and auditory detection task to investigate to what extent the reaction times in the speed-acuity test could be explained by delays in detecting stimuli and executing the motor response. Our results reveal strong developmental improvements in visual discrimination speed that have not been reported previously. They also provide normative reaction time data for children between 5 and 12 years old against which clinical results can be compared.

\textbf{Methods}

\textit{Participants}

Ninety-four children (9.4 ± 2.0 years) with normal vision participated. Inclusion criteria were: age 5 to 12 years, normal birth weight (>2500 g), birth at term (>36 weeks), no perinatal complications, no complaints of slow visual processing, crowded visual acuity of 0.1 LogMAR or better and normal development. The children were
Development of symbol discrimination speed

recruited from schools around the Bartiméus Institute, a Dutch institute for the visually impaired and the children were tested at their own primary schools. Children with glasses had to wear them during all tests. The children were considered to have no eye problems by their parents and teachers. To verify that the children had a crowded visual acuity of 0.1 LogMAR or better, we measured crowded distance visual acuity (DVA) at the start of the experiments (see ‘standard acuity measurements’ below).

Informed consent was obtained from the parents of all participants. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands) and conducted according to the principles of the Declaration of Helsinki.

Standard acuity measurements

The Freiburg visual acuity test (FrACT) software was used to measure distance visual acuity (DVA; crowded and uncrowded) at 5m with the Landolt C-test on a 23-inch LCD screen. A four alternative forced choice procedure was used. The children were seated and instructed to indicate the orientation of a black Landolt-C on a white background, either verbally or by pointing in that direction. The children could take as long as they wanted to respond. Uncrowded DVA was measured mono- and binocularly. Crowded DVA was measured binocularly using flanking rings with fixed inter-letter spacing of 2.6 arcmin. Each test consisted of 24 trials with a staircase procedure (best PEST) to determine the thresholds. The difference between crowded and uncrowded acuities in logMAR was calculated to obtain the crowding intensity.

Speed-acuity test

The speed-acuity test (Figure 2.1A) was administered binocularly with the children still seated at 5m from the screen. Each trial consisted of a single high-contrast (98.2% Michelson) black Landolt-C (2.1 cd/m²) presented at the center of the screen against a white background (235.6 cd/m²). Children had to indicate as fast and accurately as possible, on which side, right or left, the opening of the Landolt-C was located by pressing one of two mouse buttons in a two alternative forced choice task (2AFC). We chose Landolt-C’s with their opening to the left/right because of the intuitive link with left and right mouse buttons. The stimulus was presented until the child responded. The opening of the Landolt-C was unrendered to have clear edges. The round parts of the Landolt-C were rendered to prevent pixelated edges. Task difficulty was manipulated by presenting nine different optotype sizes: -0.43, -0.25,
-0.13, -0.03, 0.05, 0.27, 0.50, 0.68 and 1.09 LogMAR (method of constant stimuli). The different sizes were presented pseudo randomly in such a way that there was always at least 0.2 LogMAR difference between subsequent trials. Between trials a white screen was presented for a random period of 0.5-1.5s. The test consisted of 90 trials, 10 trials per optotype size. Children needed on average about 3 minutes to finish the test (range ~2.5 – 4.5 minutes). Children took the test twice with a short break in between. Including this break and the instruction, testing took between 10 and 15 minutes in total.

Detection tasks

The children also performed a visual detection task (VDT) and an auditory detection task (ADT) in which we measured the time children needed to respond to the presentation of a supra-threshold stimulus (Figure 2.1B). We collected these measures to investigate how other task factors, such as stimulus detection and execution of the motor response, influence the reaction time. Both tests consisted of 20 trials with random inter-trial durations of 1-4 s. In the VDT the children had to press the mouse button as soon as they saw the visual stimulus. The children were seated at 65 cm from the screen. The stimulus was a high contrast (98.2% Michelson) black letter ‘O’ (2.1 cd/m²) against a white background (235.6 cd/m²) of 1.3 LogMAR. In the ADT the children had to press the mouse button as soon as they heard the stimulus. The sound stimulus consisted of a white noise burst of approximately 75 dBA that lasted

---

**Figure 2.1. A.** The speed-acuity test. In each trial one Landolt-C was presented at the center of the screen. The children had to indicate as fast and accurately as possible, on which side the opening of the Landolt-C was located (right or left) by pressing the corresponding mouse button. Between trials a blank screen was displayed with random durations of 0.5 to 1.5 s. **B.** The detection tasks. In the visual detection task (VDT) the children had to press the button as soon as a big “O” appeared on the screen. In the auditory detection task (ADT) they had to press the button as soon as they heard a loud sound.
Development of symbol discrimination speed

500 ms. The sound stimulus was played through the speakers of a laptop which was placed approximately 65 cm in front of the children.

**Equipment**

The software for the speed-acuity test, the VDT and the ADT was written in Matlab (version 2013b; MathWorks, Inc., Natick, MA, USA) using the Psychophysics Toolbox (version 3.0.12)\(^\text{175}\). Stimulus timing and button presses were recorded and stored at 1 ms precision for offline analysis. The stimulus software was executed on a laptop (Dell M3800; Dell Inc., Round Rock, TX, USA) equipped with an OpenGL graphics card (Nvidia Quadro K1100M; Santa Clara, CA, USA). The visual stimuli were presented on a 23-inch LCD screen (Dell U2412M, 1920 x 1200 pixels, pixel pitch 0.27 mm). Visual stimulus properties were measured with a luminance meter (Minolta LS-100; Minolta Co. Ltd., Osaka, Japan). Ambient light conditions were controlled by shutting blinds or covering windows, and ranged from 100 to 350 lux as measured with a lux meter (Voltcraft MS-1500; Hirschau, Germany). Sound intensity was measured with a sound level meter (ISO-TECH SLM 1352P, ISO-Tech, Taipei, Taiwan) at the location of the subjects’ ears.

**Data analyses**

The offline analysis of the results was performed in Matlab. We first computed mean reaction times for the VDT, the ADT and the speed-acuity tests after removing atypically long or short reaction times. For each test and each optotype size, trials were excluded from the mean if the reaction time deviated more than 3 times the median absolute deviation (MAD) from the median\(^\text{176}\) after discarding reaction times <0.1 sec. On average 7% of the trials were excluded for the VDT, 6% for the ADT and 3% for the speed-acuity test. For two six-year-old children, we had to exclude the results of the VDT and ADT. Their behavior in the speed-acuity test was normal, but their reaction times in both detection tests were atypically long (>3 times the SD), presumably because they had not understood the instructions well enough. We then performed linear regression analyses to investigate developmental effects on visual acuity and reaction time measures. Regression parameters were obtained with a linear least-squares algorithm (fitlm, Matlab statistical toolbox). To estimate the range of reaction times that can be considered normal, we also calculated the 95% prediction intervals, that is, the interval in which one can expect 95% of the future observations to fall, given the current data from normally-sighted children\(^\text{177}\).
Chapter 2

For the speed-acuity test, the results consisted of the accuracy of the responses (percent correct) and the mean reaction time per optotype size. The speed-acuity test was run twice, resulting in two psychometric response functions for the accuracy data and two chronometric response curves for the reaction time data for all but six participants. In one 11-year-old the measurements failed due to technical problems, one 5-year-old refused to perform the speed-acuity test, one 7-year-old did not complete the second test, and for the other three children (6,7 and 10 year old) data from the second test had to be excluded from further analysis because they were no longer performing the task (as inferred from the fact that the median of their accuracy scores for the optotypes > 0.2 LogMAR was lower than 87.5% correct).

The accuracy data and reaction time data obtained in speed-acuity test were analyzed separately (see Supplement 2.1 for a detailed description). To determine the visual acuity, a cumulative Gaussian was fitted to the psychometric response curves with the psignifit toolbox for Matlab version 4.0\textsuperscript{178}. The threshold was taken at 75% correct, which is halfway between chance level performance for a 2AFC task and the 100% correct rate.

To quantify the average reaction times as function of optotype size $x$ (pooled across correct and incorrect choices) we used a well-documented model from the literature which uses a hyperbolic tangent function to describe chronometric response functions obtained in 2AFC sensory discrimination tasks\textsuperscript{62}:

$$ RT(x) = \begin{cases} \frac{A'}{k'(x-x_o)} \tanh\left[A'k'(x-x_o)\right] + t_R & x > x_o \\ A'^2 + t_R & x \leq x_o \end{cases} \quad (2.1) $$

This reaction time model is based on a body of literature, see e.g.\textsuperscript{40,62,179–181}, which suggests that the brain accumulates noisy sensory evidence over time and that a decision is made when the accumulated evidence scores reach a fixed decision bound (see Supplement 2.1). The model has parameters $x_o$ (critical optotype size), $t_R$ (residual time), $A'$ ($A'^2$ is the choice delay limit) and $k'$ (sensitivity). The critical optotype size, $x_o$, is the largest optotype size at which a child performs at chance level. The residual time, $t_R$, is the minimum time a child needs to respond and provides the lower bound of the chronometric curve. The residual time is thought to reflect the sum of sensory afferent delays, efferent motor delays, and other fixed delays\textsuperscript{39,40,62}. The upper bound of the chronometric curve is reached at the critical optotype size. It is the sum of the residual time and $A'^2$, which we refer to as the
choice delay limit. The choice delay limit reflects how much more time a child needs for optotype sizes at which he or she performs at chance level compared to the largest optotype size. The sensitivity parameter, $k^\prime$, is a scaling factor for the decrease in reaction times with increasing optotype sizes. Fit parameters for individualized fits were determined with a Levenberg-Marquardt nonlinear least squares algorithm (fitnlm, Matlab statistical toolbox). In these fits, we fixed the critical optotype size, $x_o$, to the value of -0.43 LogMAR based on the observation that subjects approached chance-level performance at this smallest optotype size present in our stimulus set.

To assess the effect of age on the reaction times in the speed-acuity test, and to obtain 95% prediction intervals for newly measured reaction times, we analyzed the chronometric functions with a mixed nonlinear regression model in which the parameters $A^\prime$, $k^\prime$ and $t_R$ of the reaction time model were a function of age (see Supplement 2.1). This allowed us to investigate whether these three parameters of the reaction time model ($t_R$, $A^\prime$ and $k^\prime$) were age dependent. The values of the parameters that were obtained with this mixed model analysis were then used to predict the reaction time curves one may expect for an average child of a certain age. Bootstrap procedures using the data of all children were used to obtain 95% prediction intervals for the individual fits and for the predicted reaction times curves (see Supplement 2.1).

As an indication of the repeatability of the speed-acuity test we calculated the absolute intraclass correlation coefficient (ICC$^{182}$) between response curves from the two test runs. Subsequently, the absolute ICC was calculated for the visual acuities for the two runs and a paired-samples t-test was performed to test for differences between the two runs.

Unless stated otherwise, values in the text are reported as means ± 1 standard deviation (SD). The type 1 error (alpha) was set to 0.05 for all analyses.
Results

Standard acuity measurements

We first assessed the children’s uncrowded acuity using a Landolt C-test in which there was no time pressure on the discrimination process. The average uncrowded DVA measured with the FrACT was $-0.23 \pm 0.08$ LogMAR binocularly, $-0.14 \pm 0.14$ LogMAR for the right eye, $-0.16 \pm 0.12$ LogMAR for the left eye. We also tested their crowded acuity because this measure is more sensitive in detecting visual problems than uncrowded visual acuity. The average crowded DVA was $-0.17 \pm 0.10$ LogMAR binocularly. The average crowding intensity was $0.06 \pm 0.08$ LogMAR. Furthermore, linear regression analysis (Supplemental Table S2.1) showed the expected developmental effect of age$^{138,139}$ on the crowded DVA and the crowding intensity ($R^2=0.15$, $p<0.001$ and $R^2=0.17$, $p<0.001$ respectively): the older children showed better crowded visual acuities and lower crowding intensities than the younger children. Thus, all children had normal vision as inferred from their monocular and binocular acuities.

Speed-acuity test: Acuity

The average accuracy of the responses in the speed-acuity test are presented in Figure 2.2A for four different age groups. As expected, the accuracy improves as the optotype size increases in all age groups. The mean visual acuity estimated from the psychometric response functions was $-0.28 \pm 0.05$ LogMAR. This was significantly lower (paired t-test, $t(91)=-6.62$, $p<0.001$) than the mean acuity found with the FrACT. However, the average within-subject difference between the FrACT and the speed-acuity test was only $-0.04 \pm 0.06$ LogMAR. Linear regression analysis of these data indicated that there were no significant effects of age on the uncrowded visual acuity (Supplemental Table S2.1) as one would expect for children between 5 and 12.

Speed-acuity test: Reaction times

The reaction times of the children decreased as the optotype size increased (Figure 2.2B). The reaction time for the smallest optotype, which was below the children’s visual acuity threshold, was on average 0.94s longer than the reaction time for the largest optotype. For the second optotype size of $-0.25$ LogMAR, which was around the children’s visual acuity threshold, the difference was on average 0.45s. For large optotypes the chronometric functions approach an asymptote. Note, however, that
for the third, fourth and fifth optotype size (-0.12, -0.03 and 0.05 LogMAR), which were all above threshold, the asymptote is not yet reached; the mean reaction times were still 0.23, 0.15 and 0.11 s above it. Furthermore, older children were faster than younger children. For instance, the average reaction time for the five- and six-year-old children for the largest optotype were on average 0.33 s slower than the eleven-and twelve-year-olds. This difference was ~0.53 s around the visual acuity threshold.

This coarse inspection of the data thus shows a clear developmental dissociation between speed and accuracy; where the psychometric curves fall practically on top of one another (Figure 2.2A), the chronometric curves clearly differ between the age groups (Figure 2.2B). This is evidently different from the speed accuracy trade-off one can expect from an individual participant under different task conditions. For this reason, we analyzed the reaction time data independent of the acuity data.

![Figure 2.2](chart.png)

**Figure 2.2.** Results of the speed-acuity tests averaged across children in four different age groups. A. Psychometric functions which quantify the accuracy of the responses (in percentage correct answers) as a function of optotype size. B. Chronometric functions which quantify the reaction times (in seconds) as a function of optotype size. Error bars indicate ±1 standard error of the mean (SEM). The number of children per age group is listed in the legend.
As a first step in this analysis we quantified the reaction times as a function of optotype size by fitting the reaction time model (Equation 2.1) to the individual chronometric curves. Figure 2.3A shows the results of such an individual fit along with the measured reaction times. The average $R^2$ for the individual fits was $0.96 \pm 0.07$ (range $0.65 – 1.00$). The individual fit parameters along with the 95% prediction intervals for these parameters are presented in Figure 2.3B-D. Linear regressions applied to these fit parameters showed significant age effects on all parameters of the reaction time curves (Supplemental Table S2.2). The residual time decreased with age ($\beta=-0.046$, $t(176)=-10.65$, $p<0.001$), the log-transformed choice delay limit decreased with age ($\beta=-0.037$, $t(176)=-3.03$, $p=0.003$), and the log-transformed sensitivity parameter increased with age ($\beta=-0.056$, $t(176)=2.00$, $p=0.047$) which all points towards a significant increase in optotype discrimination speed.

Figure 2.3. Age dependence of parameters of the chronometric responses functions in the speed-acuity test. A: Example of a nonlinear fit to the reaction time data from a single subject. B-D: The 95% prediction intervals (shades) and the parameter values for the first (blue dots) and second test run (cyan dots) obtained in our subjects for the residual time $t_R$ (B), choice delay limit $A'^2$ (C), and sensitivity $k'$ (D).
Normative developmental data from cross-sectional studies such as ours are often stratified according to restricted age groups. However, this approach is not very efficient in terms of statistical power since more and smaller age groups are needed to characterize steeper age effects with sufficient resolution\textsuperscript{183,184}. In the present study, we therefore used a mixed nonlinear regression model to quantify the steep changes in reaction times as a continuous function of age (and optotype size), together with a numerical approach (bootstrapping) to estimate the range of reaction times that can be considered normal (Supplement 2.1). Collectively, the fixed- and random-effects in this model accounted for 86\% of the total variance in the reaction times across all optotype sizes and all subjects (conditional $R^2=0.86$), indicating that this model provided a very good description of our data. Furthermore, the three fixed-effects that captured the developmental improvements in visual discrimination speed were all statistically significant (Supplemental Table S2.3). This bolsters the conclusion from Figure 2.3 that the developmental enhancement of optotype recognition consist of an overall, size invariant decrease in reaction time, as indexed by the decrease in residual time, as well as size-specific reductions in discrimination time as indexed by the decrease in the choice delay limit and the increase in sensitivity.

The parameter values that this mixed regression model predicts for the reaction time curves for a child of a certain age can be calculated from the model’s fixed-effects in the following way:

$$A' = \exp(0.189 - 0.027 \cdot Age)$$
$$k' = \exp(1.558 + 0.076 \cdot Age)$$
$$t_n = 1.031 - 0.048 \cdot Age$$

By substituting these parameter values in Equation 2.1, one can compute the expected reaction time for this child for each optotype size, $x$, to obtain the norm (i.e., the marginal response).

The resulting norm curves and corresponding prediction intervals for the average reaction times of an individual subject with 10 trials per optotype size are presented for 8 different ages in Figures 2.4A-H. The age effects are clearly visible: i) the predicted reaction times for the 12 year olds are systematically shorter than for the 5 year olds, ii) the shape of the predicted curve depends on age (Figure 2.4I), and iii) the difference between the maximal and minimal reaction time is smaller for the older
children and their reaction time decreases faster with optotype size. Additionally, the width of the prediction intervals decreases with increasing optotype size and age. There is also overlap between the ranges of responses that can be considered normal for adjacent ages. This is most prominent for the smallest optotypes, which can be understood from the logarithmic scaling of the choice delay limit and sensitivity with age (Figures 2.3C and 2.3D). Because of the overlap of the prediction intervals, it is of interest to aggregate the reaction time data across the optotype sizes in a summary score. We have derived such a measure in Supplement 2.2.

Figure 2.4. A-H: Median of the predicted average reaction time and corresponding 95% prediction interval as a function of optotype size for 8 different ages. Predictions were computed for the center of each age category. The mean reaction times obtained from our subjects in the first (blue dots) and second (cyan dots) test run are superimposed. I: Shape of the predicted chronometric curve at 4 different ages. For each age the predicted reaction time on the largest optotype was subtracted from the predicted chronometric curve to reveal the age effect on the shape of the chronometric curve.
Development of symbol discrimination speed

Detection tasks

The children performed two detection tasks to investigate how much of the age-dependent reduction in reaction times that was found in the speed-acuity test might be due to faster visual discrimination compared with other factors that also influence the reaction time, such as stimulus detection and execution of the motor response. The results of the simple visual and auditory detection tasks are shown in Figure 2.5. Linear regression analysis of these data (Supplemental Table S2.4) indicated that age explained a significant proportion of the variability in the VDT ($R^2=0.56$, $F(1,92)=117$, $p<0.001$) and the ADT ($R^2=0.53$, $F(1,92)=101$, $p<0.001$). On both detection tests, the older children had shorter reaction times than the younger children (Figure 2.5A and 2.5B). Most children performed faster on the ADT than the VDT (average difference: $62 \pm 42$ ms, paired t-test, $t(90)=13.94$, $p<0.001$). The within-subject difference between the VDT and ADT is plotted against age in Figure 2.5C. Linear regression analysis (Supplemental Table S2.4) of these data revealed a significant age effect on

Figure 2.5. The mean reaction times in the visual (A) and auditory (B) detection task, as well as the difference between the reaction times on these two tasks (C) for all children plotted as a function of age. The solid lines are the results of the linear regression analysis. The shaded area represent the 95% prediction intervals for a new normally-sighted child. Note the factor 2 scaling difference between A, B and C.
this difference ($R^2=0.09, F(1,92)=8.71, p=0.004$); the average difference between the VDT and ADT decreased by ~50 ms over the 8 years inclusion range. Because this effect was small and the explained variance low, we also compared the slopes of the regression lines for the ADT and the VDT with a repeated measures ANCOVA. The results of this ANCOVA confirmed that the slope of the regression line was indeed significantly steeper for the VDT (slope ADT: -18 ± 2 ms/year, slope VDT: -24 ± 2 ms/year, $F(1,89)=9.30, p=0.003$). Thus, it appears that the speed of visual detection catches up on the speed of auditory detection.

In the reaction time model the residual time $t_r$ is thought to reflect the sum of efferent motor delays, sensory afferent delays, and other fixed delays that are unrelated to the discrimination process. If this assumption is true, the reaction times on the detection tasks should correspond to the residual time. To test whether this assumption holds true, the children’s reaction time for discriminating the easiest

![Figure 2.6. A. Mean reaction time for discriminating the easiest optotypes (>0.5 LogMAR) in the speed-acuity test plotted as a function of age. B. Age dependent difference between the mean reaction time for discriminating the easiest Landolt-Cs and the mean reaction time for detecting the large ‘O’ in the visual detection task (VDT). C. Age dependent difference between the mean reaction time for discriminating the easiest Landolt-Cs and the mean reaction time for detecting the sound stimulus in the auditory detection task (ADT). The solid lines present the results of the linear regression analyses. The shaded area present the 95% prediction intervals for individual observations.](image-url)
Development of symbol discrimination speed

Optotypes in the speed-acuity task were compared to their reaction time for detecting the salient stimuli in the VDT and the ADT. We first evaluated the effect of age on the children’s mean reaction time for the largest three optotypes (Figure 2.6A) as a model-invariant estimate of the residual time. In line with the results shown in Figure 2.3B, linear regression analysis of these data (Supplemental Table S2.5) showed a significant decrease in the reaction times with age ($t(90)=-11.82, p<0.001$), explaining 61% of the variance in reaction time for discriminating the easiest optotypes. Note that the decrease in reaction time with age is $61 \pm 5$ ms/year, resulting in a total decrease of nearly half a second over the 8 years inclusion range.

Figure 2.6B presents the difference between the reaction time for the easiest optotypes in the speed-acuity test and the VDT as a function of age. The children needed on average $0.36 \pm 0.11$ s (range 0.14-0.69) more to discriminate on which side the opening of the large Landolt-Cs was located during the speed-acuity test compared to detecting a large ‘O’ in the VDT. Regression analysis (Supplemental Table S2.5) revealed that this reaction time difference decreases $36\pm5$ ms/year, reducing the time needed for the discrimination response by $\sim290$ ms across the 8 years inclusion range. Figure 2.6C presents the reaction time difference with the ADT as a function of age. On average the children needed $0.42 \pm 0.12$ s (range 0.20-0.77s) more to discriminate the orientation of large optotypes compared to detecting the sound in the ADT. Furthermore, linear regression analysis (Supplemental Table S2.5) showed that the reaction time difference decreases more than $40 \pm 5$ ms/year. Taken together, these reaction time differences with the VDT and ADT show that the assumption that the residual time is unrelated to the discrimination process is not correct; a substantial fraction of the age-dependent decrease in reaction time in the speed-acuity task can be explained by improved discrimination of the visual stimuli.

**Test-retest reliability**

No significant differences were found between the visual acuity measured in the first (-0.28 ± 0.05 LogMAR) and second (-0.28 ± 0.04 LogMAR) test runs (paired t-test, $t(87)=0.51, p=0.61$). The difference between the visual acuity thresholds of the two test runs ranged from -0.25 to 0.14 LogMAR, and was less than 0.1 LogMAR for 94% of the participants. The absolute intraclass correlation between the first and second test was 0.78 (95%CI: 0.75-0.80) for the percentage correct answers and 0.74 (95%CI: 0.63-0.82) for the acuity estimates. Furthermore, the average within-subject difference between the mean reaction times per optotype size of the first and second
test run was 0.05 ± 0.13 s (paired t-test, \( t(87)=3.376, p=0.001 \)), indicating that the children were slightly faster on the second test. Even so, the absolute intraclass correlation between reaction times in the first and second test was 0.78 (95%CI: 0.75-0.81).

**Discussion**

The goal of this study was to investigate developmental effects on the speed of visual symbol discrimination in children between 5 and 12 years old, because a child’s ability to distinguish visual details fast is important for its participation in school and society. Towards that end, we used an optotype-discrimination reaction-time task which measures visual acuity and visual discrimination speed simultaneously. We found that it is feasible to use such a test for children in this age group and the results show that there are considerable developmental improvements in visual discrimination speed. This suggests that an important optimization of the visual discrimination process takes place in the developing visual system of 5-12 year-old children. The developmental dissociation between speed and accuracy that is revealed by our data implies that visual acuity alone cannot predict how long it takes for a child to see. The combination measurement of recognition acuity and recognition speed is therefore relevant for the assessment of a child’s visual development, and may be of aid in clinical diagnostics of visual impairment and rehabilitation indications.

**Development of visual discrimination speed**

The reaction times in the speed-acuity test depend strongly on the size of the Landolt-C. This is in line with a series of previous studies on the effect of stimulus strength on reaction time\(^{62-65}\). Additionally, the decrease in reaction times as optotype size increases is comparable to the effect of print size on reading speed\(^{67,68}\). Critical print size, i.e., the smallest print size at which participants approach the asymptote of reading speed, is on average about 0.2 LogMAR above visual acuity threshold. The current findings indicate that the smallest optotype size at which participants approach the asymptote on the speed-acuity test, is even larger – roughly 0.3-0.5 LogMAR above the acuity threshold (Figure 2.2B).

The reaction time model with a hyperbolic tangent function adequately described the effect of stimulus size on the reaction times. We are aware that a variety of alternative sensory decision-making models exists\(^{185-187}\). Given the relative simplicity of this particular model, and the goodness of fit for most participants, we think however that the applied model is a useful tool in the analysis of visual discrimination
speed. Indeed, by incorporating this quantitative description of reaction time as a function of optotype size in a mixed nonlinear regression analysis, we were able to identify distinct effects of age on the speed of visual processing in children with normal vision and normal development.

First, and foremost, we found that the residual time was much shorter in the oldest children compared with the youngest. This effect is clearly seen as a shift of the chronometric functions along the vertical axis (Figure 2.2B). Although the results from the auditory detection task indicate that a significant fraction (~140 ms of the ~400 ms total decline over 8 years) of this general decrease in reaction times with age cannot be attributed to developmental effects within the visual system (Figure 2.6B), comparison of the auditory and visual detection tasks (Figure 2.5C) shows that a part of the reaction time decrease is due to a general increase in the speed of the visual discrimination process (~50 ms over the course of 8 years). Interestingly, however, this general speed-up of visual processing only accounted for a limited fraction of the general decrease in reaction times on the speed-acuity task. As is revealed by subtracting the reaction times obtained in the visual detection task from the ones obtained for discriminating large optotypes (Figure 2.6C), there is an additional increase in discrimination performance of near 250 ms on average over the course of 8 years development. Where does this improvement come from? Does it reflect a developmental improvement in visual processing or is it related to non-visual factors? One possible explanation is that the additional processing required in the speed-acuity task over the detection task becomes more efficient with age. The difference in processing includes more complex cognitive judgment (discriminating between two alternative orientations of a symbol vs. registration of a suprathreshold sensory stimulus) and a more complex motor response (pressing one of two buttons vs. pressing a single button). Alternatively, there could be a general decrease in noise levels within the visual system which allows the older children to respond faster (by adopting lower decision bounds) while maintaining a similar accuracy.

On top of the large effect of age on the residual time, we also found that the shape of the chronometric curves changed with age. Since these shape-changes relate to the size of the visual stimulus and not to any other component of the speed-acuity task, they expose an optimization of the decision-making process which can only be attributed to developmental changes in visual processing. Indeed, the extra time that subjects needed to reach a decision regarding the orientation of the smallest Landolt-C compared with the largest C was larger in the younger children. Since
their acuities were not significantly different from the older children, this increase in the choice delay limit (Figure 2.3C, Figure 2.4) suggests that the younger children needed more evidence to reach a decision (i.e., they applied a higher decision bound) thereby sacrificing speed for accuracy to compensate for increased noise levels in their visual system. In addition, the decrease in reaction time as a function of optotype size was steeper in the older children, as is reflected in a compression of the chronometric curves along the horizontal axis (Figure 2.3D, Figure 2.4). Although this effect was also not as strong as the effect on the residual time, it shows that the sensitivity of extracting the relevant information from the stimulus was higher for the older children.

In children with visual impairment, one might expect significant effects on their reaction time curves because of their reduced sensitivity. Whether this prediction holds across clinical populations remains to be tested (see 188). Others have already demonstrated that differences in visual acuity resulted in differences on computerized neurobehavioral tests57, the useful field of view test55, cognitive tests189, and on neural markers of visual processing190. Similarly, simulated visual impairment influenced the outcome of cognitive and neuropsychological tests56,59 and visual acuity was a significant predictor of reaction time on a computer task in patients with macular degeneration191. However, none of the studies above tested visual processing speed and visual acuity simultaneously.

**Test features**

The speed-acuity test is an objective, easy-to-administer vision test which allows quick (~5 min), simultaneous assessment of visual acuity and visual discrimination speed. Visual discrimination speed, as measured with our speed-acuity test, can be conveniently summarized by a single delay index which provides an age-invariant comparative measure for the speed with which a subject is able to discriminate optotypes (Supplement 2.2). The difference between the visual acuity determined with the speed-acuity test and the FrACT was on average only 0.04 LogMAR. This difference is within acceptable limits according to the international standard EN ISO 8597192. A possible explanation for this difference in acuities could be that during the FrACT children responded verbally and during the speed-acuity test they responded with button presses. Some children refused to guess when the optotypes in the FrACT were small, which could have resulted in an underestimation of the acuity with the FrACT. In addition, the repeatability of the test was good193, with ICCs of 0.78 for
the accuracy curves, 0.74 for the visual acuity thresholds, and 0.78 for the reaction time curves. Moreover, no significant differences were found between the visual acuity thresholds of both test runs. Therefore, the speed-acuity test proves to be a valid test to measure visual discrimination speed and visual acuity simultaneously in children. The test-retest repeatability with longer intervals between tests needs to be addressed in further research.

Prediction intervals that were obtained by pooling the reaction-times from the two test runs into one average per optotype size (not shown) were very similar to the ones shown in Figures 2.3-2.4 for a single test consisting of 10 trials per optotype size. Only the upper bound of prediction interval for the sensitivity parameter ($k'$) was somewhat lower. Thus, as expected from the high ICCs, the benefit of adding more trials to the standard test seems to be of little practical importance. The reliability of the estimation of the visual acuity threshold and the shape parameters ($A'$ and $k'$) of the reaction-time curve might perhaps be improved further by adding optotype sizes around the visual acuity threshold. Note, however, that there are certain limitations to this that are imposed by the available space (distance from participant to screen) and display resolution.

**Conclusion**

The data which we obtained from 5-12 year-old school children with normal vision revealed large improvements in visual processing speed over the course of 8 years development that have not been documented before. This suggests that quantitative assessment of visual processing speed may be crucial for a better understanding of the developing visual system in general and better assessment of the impact of visual impairment in clinical populations. The current data provide the required normative data.
Supplement 2.1.

The accuracy data and reaction time data obtained in speed-acuity test were analyzed separately. To determine the visual acuity threshold, a cumulative Gaussian was fitted to the proportion correct scores that were obtained for the different optotype sizes \( x \) (in LogMAR):

\[
P_{\text{correct}}(x) = \gamma + (1 - \gamma - \lambda) \left[ \phi \left( C \frac{x - m}{w} \right) \right]
\]

where \( \phi \) is the cumulative standard normal distribution and \( \phi^{-1} \) its inverse; \( m \) is the inflection point of the curve; \( w \) is the width parameter reflecting the difference between stimulus levels at which \( \phi \) reaches 0.05 and 0.95; \( \gamma \) is the guess rate at chance level performance; and \( \lambda \) is the lapse rate of the subject. Fits were obtained with the psignifit toolbox for Matlab version 4.0. Because many children did not reach 100% accuracy for the largest optotypes, the lapse rate \( \lambda \) was allowed to vary between 0 and 0.1. \( \gamma \) was fixed to the 0.5 chance level performance for a 2AFC task. The acuity threshold was taken at 75% correct, which is halfway between the \( \gamma = 0.5 \) chance level performance for a 2AFC task and the 100% correct rate.

We used a well-documented model from the literature to quantify the average reaction times on the speed-acuity test as function of optotype size \( x \). This model describes chronometric response functions obtained in 2AFC sensory discrimination tasks with a hyperbolic tangent function:

\[
RT(x) = \begin{cases} 
A' \tanh \left[ A' k' (x - x_o) \right] + t_R & x > x_o \\
A'^2 + t_R & x \leq x_o
\end{cases}
\]

This model assumes that the brain accumulates noisy sensory evidence over time until the accumulated evidence scores reach a fixed decision bound (Supplemental Figure 2.1). At that time a decision is made. The height of the decision bound for the two alternatives is assumed to be symmetric around zero. The total reaction time, RT, is considered to be the sum of this stimulus-dependent decision time and an independent residual time. The residual time \( t_R \) is thought to reflect the sum of sensory afferent delays, efferent motor delays, and other fixed delays that are unrelated to the actual stimulus discrimination process.
To obtain reliable estimates of this lower reaction time limit, we included large optotypes in our stimulus series. The height of the decision bound \( A \) relative to the noise level \( \sigma \) is reflected in \( \sigma' = \frac{A}{\sigma} \). The parameter \( x_o \) is the optotype size at which the stimulus becomes too weak to bias the evidence scores towards either one of the two alternatives, i.e., the ‘critical optotype size’. At this level, the decision bounds are reached purely by chance after an average delay of \( A^2 \) seconds\(^{62} \), here referred to as the “choice delay limit”, and as a result, the chronometric function reaches its upper limit \( A^2 + t_r \). The parameter \( k' \) is a measure of the sensitivity of the subject’s visual system to the relevant stimulus features, where \( k' = \frac{k}{\sigma} \). Note that the signal-to-noise ratio \( \frac{k(x - x_o)}{\sigma} \) of the sensory signal determines the average “decision rate”, i.e., how fast the evidence scores tend to accumulate towards the decision bound. Fit parameters for these individualized fits were determined with a Levenberg-Marquardt nonlinear least squares algorithm (fitnlm, Matlab statistical...
In these fits, we fixed $x_o$ to the value of -0.43 LogMAR based on the observation that subjects approached chance-level performance at this smallest optotype size present in our stimulus set.

To assess the effect of age on the reaction times in the speed-acuity test, and to obtain prediction intervals for newly measured reaction times, we analyzed the chronometric functions with a mixed nonlinear regression model. The fixed-effect parameters in this model estimated the mean $A', k'$, and $t_r$ of the population (via parameters $\beta_{a_0}, \beta_{k_0}$, and $\beta_{t_{r0}}$) as well the average changes in these parameters with age (via parameters $\beta_{a_1}, \beta_{k_1}$, and $\beta_{t_{r1}}$). The random-effects ($b_{a,i}, b_{k,i}, b_{t_{r,i}}$) allowed $A', k'$, and $t_r$ to vary between individual participants. Since inspection of the $A'$s and $k'$s that were found for individualized curve fits (see Figure 2.4) indicated that the random effects did not follow a normal distribution, we fitted their log-transform instead. This resulted in the following definition of the parameters in Equation S2.3:

$$
A'_i = \log(\beta_{a_0} + \beta_{a_1} \cdot \text{Age}_i + b_{a,i}) \\
k'_i = \log(\beta_{k_0} + \beta_{k_1} \cdot \text{Age}_i + b_{k,i}) \\
t_{r,i} = \beta_{t_{r0}} + \beta_{t_{r1}} \cdot \text{Age}_i + b_{t_{r,i}}
$$

with $[b_{a,i}, b_{k,i}, b_{t_{r,i}}] \sim N(0, \Psi)$ a multivariate normal distribution with zero means and covariance matrix $\Psi$. Subscripts $i$ refer to individual participants. $\Psi$ was estimated from the data along with the fixed-effects and random-effects parameters (nlmeft, Matlab statistical toolbox). The continuous variable age was centered on the age of 9, the middle of the inclusion range.

To assess the effect of age on the reaction time curves (Supplemental Table S2.3), we pooled data from the two test sessions per subject. Confidence intervals for the fixed-effects factors were obtained by bootstrapping ($n=2000$). For each bootstrap trial, we resampled at the level of subjects and within subjects. The within-subject sampling generated a new set of reaction time means for each test block by sampling from the distribution of reaction times measured per optotype.

To obtain prediction intervals for the average reaction times of an individual subject on a single test block that contain 10 trials per optotype size (Figure 2.4), we again used a bootstrap procedure which resampled at the level of subjects and within subjects ($n=2000$). In this case, however, the resampled data from the 1st and 2nd test runs were fitted by two independent mixed models each bootstrap.
iteration to accommodate the fact that subjects were on average slightly faster in the second block. Individual predictions were then generated by adding randomly drawn residuals from both sets of bootstrapped models to their respective conditional responses (i.e., the model predictions which included contributions from both fixed- and random-effects predictors) such that the resulting predictions intervals span the complete range of possible observations. To ensure that any residual systematic deviation between the reaction time model and the data was accounted for in the prediction intervals, the drawing of residuals was conditioned on optotype size. The resulting reaction time predictions were also used to determine prediction intervals for the delay index (Supplement 2.2).

Supplement 2.2

A measure that quantifies whether a child is slow on the speed-acuity test compared with its peers should be age invariant and independent of the optotype sizes used in the test. In order to obtain such a measure, two effects had to be accounted for. First, the distribution of the observed reaction times ($RT_{\text{obs}}$) and the distribution of the predicted reaction times ($RT_{\text{mar}}$) are skewed towards longer reaction times (Figures 2.4A-H). This skewness causes an unequal weight of being slower or being faster than the norm. We therefore used the log-transform of the observed reaction times and the predicted reaction times to correct for the skewness of the reaction time distributions. The difference of these log-transformed reaction times for the $j^{\text{th}}$ optotype size ($\text{log}(RT_{\text{obs},j}) - \text{log}(RT_{\text{mar},j})$), which essentially denotes the measured reaction time in dB relative to the norm, is by definition age-invariant. However, the variance of the average reaction times decreases with optotype size (Figures 2.4A-H), which means that one may expect bigger differences between observed and predicted reaction times for smaller optotypes compared with larger optotypes. To account for this optotype-size dependence of the variance, we normalized the log-transformed difference scores for each optotype size $j$ to its expected standard deviation, $\sqrt{VAR_j}$.

The delay index ($DI$) that we computed for each child was thus defined in the following way:

$$DI = \frac{1}{n_{\text{opt}}} \sum_{j=1}^{n_{\text{opt}}} \frac{\text{log}(RT_{\text{obs},j}) - \text{log}(RT_{\text{mar},j})}{\sqrt{VAR_j}}$$  (S2.4)
where $RT_{obs,j}$ is the observed mean reaction time for the $j$th optotype size and $\hat{RT}_{mar,j}$ is the corresponding marginal response, i.e., the model prediction which includes only the fixed effects. $VAR_j$ is the variance of $\log(RT_{obs,j}) - \log(\hat{RT}_{mar,j})$. $VAR_j$ is age invariant (since $\log(RT_{obs,j}) - \log(\hat{RT}_{mar,j})$ is age invariant) and it turned out that it can be approximated quite well using the same hyperbolic tangent function as the chronometric curves (Equation 2.1), with parameters $A' = 0.37$, $k' = 37.27$ and $r_t = 0.0157$.

As a result, the delay index (Supplemental Figure 2.2, Equation S2.4), provides an age-invariant (linear regression: $t(176) = 0.55$, $p = 0.58$) comparative measure for a subject’s visual discrimination speed. Note that the prediction intervals are asymmetric around the mean, which is by definition at 0 (no delay). The 5th percentiles for different ages all fall around -0.96, while the 95th percentiles fall around 1.39. A value of 1 signifies that the reaction times, expressed in dB re. to the norm, are

![Supplemental Figure 2.2. Individual delay indices for the first (blue dots) and second (cyan dots) test run plotted against age. The median (red line) and 95% prediction intervals (shaded area) along with the 5th and 95th percentile (dashed red lines) were obtained from the mixed nonlinear model through bootstrapping (Supplement 1).](image-url)
on average 1 standard deviation above the norm for the respective optotype sizes. The cumulative percentile scores for the delay index are presented in Supplemental Figure 2.3. This graph allows one to infer percentile scores for newly tested children.

The intraclass correlation between the delay indices of the first and second test run was 0.78 (95%CI: 0.68-0.86). The delay index for the first test run was on average 0.09 ± 0.52 higher than on the second test run. This difference was not significant ($t(87)=1.58, p=0.11$).

Supplemental Figure 2.3. The cumulative percentile scores for the delay index.
**Supplemental Table S2.1.** Regression results for the effect of age on the visual acuity scores and crowding intensities measured with the Freiburg visual acuity test (FrACT).

<table>
<thead>
<tr>
<th></th>
<th>Fit</th>
<th>Intercept</th>
<th>Age (slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>F</td>
<td>β₀</td>
</tr>
<tr>
<td>Residual time, tR</td>
<td>0.39</td>
<td>113</td>
<td>1.03</td>
</tr>
<tr>
<td>Choice delay limit, log(A')</td>
<td>0.05</td>
<td>9.19</td>
<td>-0.09</td>
</tr>
<tr>
<td>sensitivity, log(k')</td>
<td>0.02</td>
<td>4</td>
<td>1.94</td>
</tr>
</tbody>
</table>

**Supplemental Table S2.2.** Regression results for the effect of age on the parameter values of the individualized nonlinear fits. Note that the choice delay limit and the sensitivity were log transformed.

<table>
<thead>
<tr>
<th></th>
<th>Fit</th>
<th>Intercept</th>
<th>Age (slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>F</td>
<td>β₀</td>
</tr>
<tr>
<td>Visual acuity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODS</td>
<td>0.03</td>
<td>2.74</td>
<td>-0.17</td>
</tr>
<tr>
<td>OD</td>
<td>&lt;0.01</td>
<td>0.003</td>
<td>-0.15</td>
</tr>
<tr>
<td>OS</td>
<td>0.03</td>
<td>2.93</td>
<td>-0.06</td>
</tr>
<tr>
<td>Crowded ODS</td>
<td>0.15</td>
<td>15.2</td>
<td>0.02</td>
</tr>
<tr>
<td>Crowding intensity</td>
<td>0.17</td>
<td>18.2</td>
<td>0.23</td>
</tr>
</tbody>
</table>

**Supplemental Table S2.3.** Fixed effects of the mixed nonlinear regression. The 95% confidence intervals of the parameter values were obtained through bootstrapping (n=2000). The goodness of fit of the model are described by the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Log Likelihood, and the proportion of variance explained by both the fixed and random factors (conditional R²). Note that for A' and k' the log-transform was fitted.

<table>
<thead>
<tr>
<th></th>
<th>Fixed factor</th>
<th>Age dependent factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Name</td>
<td>Estimate 95%CI</td>
</tr>
<tr>
<td>Parameters</td>
<td>Choice delay limit (A')</td>
<td>βₐ₀: 0.189</td>
</tr>
<tr>
<td></td>
<td>Sensitivity (k')</td>
<td>βₖ₀: 1.558</td>
</tr>
<tr>
<td></td>
<td>Residual time (tᵢ)</td>
<td>β₉₀: 1.031</td>
</tr>
</tbody>
</table>

**Goodness of fit**

<table>
<thead>
<tr>
<th></th>
<th>AIC</th>
<th>BIC</th>
<th>Log Likelihood</th>
<th>Conditional R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-547.2</td>
<td>-514.2</td>
<td>286.6</td>
<td>0.86</td>
</tr>
</tbody>
</table>
**Supplemental Table S2.4.** Regression results for the effect of age on the simple auditory (ADT) and visual (VDT) detection tests.

<table>
<thead>
<tr>
<th>Detection tasks</th>
<th>Fit</th>
<th>Intercept</th>
<th>Age (slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>F</td>
<td>$\beta_0$</td>
</tr>
<tr>
<td>ADT</td>
<td>0.46</td>
<td>77.7</td>
<td>434.5</td>
</tr>
<tr>
<td>VDT</td>
<td>0.57</td>
<td>120</td>
<td>556.5</td>
</tr>
<tr>
<td>VDT-ADT</td>
<td>0.09</td>
<td>9.3</td>
<td>125.9</td>
</tr>
</tbody>
</table>

**Supplemental Table S2.5.** Regression results for the effect of age on the reaction times for discriminating the easy optotypes in the speed-acuity test and for the reaction time difference between discriminating these easy Landolt-Cs and detecting the sound stimulus in the ADT, or detecting a large 'O' in the VDT.

<table>
<thead>
<tr>
<th>RT easy optotypes</th>
<th>Fit</th>
<th>Intercept</th>
<th>Age (slope)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>F</td>
<td>$\beta_0$</td>
</tr>
<tr>
<td>RT easy optotypes</td>
<td>0.61</td>
<td>140</td>
<td>1.27</td>
</tr>
<tr>
<td>RT easy optotypes - ADT</td>
<td>0.45</td>
<td>69.6</td>
<td>0.82</td>
</tr>
<tr>
<td>RT easy optotypes - VDT</td>
<td>0.35</td>
<td>48.2</td>
<td>0.70</td>
</tr>
</tbody>
</table>
Symbol discrimination speed in children with visual impairments

Introduction

According to the World Health Organization (WHO), visual impairment is characterized by low visual acuity and visual field deficits. Based on these criteria, worldwide, approximately 19 million children are considered visually impaired. Visual impairment in children is most often caused by ocular diseases and/or genetic factors affecting the eye, such as retinal dystrophies, albinism, optic atrophies, retinopathy of prematurity, and congenital cataract. The most prevalent cause of visual impairment in children in developed countries, however, is cerebral visual impairment. Cerebral visual impairment (CVI) results from diverse developmental disorders and brain injuries, such as genetic disorders, malformations of cortical development, preterm birth, closed head trauma, encephalitis, hypoxia and epilepsy. Because of its diverse etiology, a wide variety of visual problems may occur in CVI, ranging from deficits in lower visual functions such as visual acuity, visual field and contrast sensitivity, to higher visual functions such as visual attention, movement processing and visual search. Thus, the WHO definition of visual impairment may be too narrow for children with CVI, as well as for children with other visual impairments.

Presently, clinical tests of visual function do not assess visual processing speed. However, non-timed visual acuity tests may not reflect the demands of daily-life activities. Therefore, it has been argued that measurements which take into account speed and accuracy are key to a better quantitative assessment of visual impairment. This argument is strengthened by our finding that the basic ability to quickly discern visual details improves considerably in children with normal vision between 5 and 12 years of age due to faster, rather than more accurate, visual discrimination performance. This suggests that in addition to high-acuity vision, there is a need for fast visual processing to cope with daily-life demands. Reduced visual processing speed may explain some of the problems in higher visual functions in children with visual impairments. For example, studies have shown that children with CVI display orienting responses with longer latencies as well as longer search times. Also, children with other visual impairments (VIo) due to congenital ocular disorders and/or retinal abnormalities have longer search times and lower reading speeds compared to children with normal vision. Furthermore, in adults, visual acuity correlates with performance on visual processing speed tests and visual acuity is a significant predictor of reaction times in adults with macular degeneration. Although these studies correlated spatial and temporal aspects of
vision, none of the studies above tested visual processing speed and visual acuity simultaneously.

In the present study, we therefore employed an easy-to-administer vision test that allows quick, simultaneous assessment of visual acuity and visual discrimination speed, which we also used to investigate the developmental improvement of symbol recognition in school children\textsuperscript{196}. The aim was to test if children with visual impairments due to congenital ocular disorders and/or retinal abnormalities (VI\textsubscript{o}) and children with cerebral visual impairments (CVI) are indeed slower in discerning visual details than children with normal vision (NV), and to assess if such differences may be explained by their reduced acuity alone. In addition, we tested whether these children were slow in detecting and responding to large visual stimuli in the first place, as may be expected if they have, e.g., reduced contrast and/or luminance sensitivity. To further assess whether the outcomes were influenced by developmental delays and/or more general problems in sensorimotor processing, as may be expected, e.g., in the case of abnormal brain development, we also measured their reaction times on an auditory detection task. Finally, we explored whether abnormal reaction times on the two detection tasks could account for abnormal reaction times in the speed-acuity test, or whether the fine spatial discrimination required for the speed-acuity test takes extra time in children with visual impairment.

**Methods**

**Participants**

Forty-seven children participated in this study, thirty children (8.6 ± 2.3 years) with visual impairment due to congenital disorders of the eye and/or retinal abnormalities (VI\textsubscript{o}), and seventeen children (9.0 ± 1.5 years) with cerebral visual impairment (CVI). Control data were derived from 94 children (9.4± 2.0 years) with normal vision (NV) and normal development who were included in a separate study\textsuperscript{196}.

Children with VI\textsubscript{o} and CVI were recruited from Bartiméus, a Dutch institute for the rehabilitation of visually impaired people. The institution sent letters to the parents of children that met the inclusion criteria, and all children whose parents responded positively to this call were included (unless later testing showed that the child did not meet the inclusion criteria). Children with NV were recruited from primary schools in the neighborhood of this institute. The schools sent letters to the parents of all children, and those responding positively after reviewing the information brochure were included.
Inclusion criteria for the children with NV and VI were: age 5 to 12 years, normal birth weight (>2500 g), birth at term (>36 weeks), no perinatal complications, and normal development. For children with VI, additional criteria were an ophthalmological diagnosis and a crowded distant visual acuity (DVA) between 0.2 and 1.3 logMAR. Children with NV had to have a crowded DVA of 0.1 logMAR or better. Inclusion criteria for the children with CVI were age 5 to 12 years, being diagnosed of having CVI (by the Bartiméus institute), and having a crowded visual acuity of 1.3 LogMAR or better. The diagnosis of CVI was based on a thorough ophthalmological examination by a qualified pediatric ophthalmologist and orthoptist, including visual acuity and visual field tests, fundoscopy, and a detailed patient history, including a review of information available in medical records. Children who did not have the mental and motor skills to understand and execute the tasks were excluded, based on the judgment of their parents before the tests and on careful monitoring during testing by the examiner. The Freiburg visual acuity test (FrACT) was used to verify that the visual acuity of the children fell within the inclusion criteria (see section: Test procedures and equipment). Children with prescription glasses wore them during all tests.

Supplemental Table S3.1 presents the ophthalmological diagnosis, clinical characteristics (i.e., presence of nystagmus and strabismus), and binocular visual acuities of all participants with VI and CVI. The most prevalent diagnosis for the children with VI was albinism (n=8), followed by congenital stationary night blindness (n=6), infantile nystagmus (n=5) and hypermetropia (3). Most children with CVI were born prematurely or suffered from perinatal complications.

The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands) and conducted according to the principles of the Declaration of Helsinki and the Dutch code of conduct regarding the participation of minors. Informed consent was obtained in writing from all parents before the start of the measurements.

Test procedures and equipment

Test procedures and equipment were mostly the same as in the companion paper. Key properties of the stimuli and psychophysical methods are reproduced here for the reader’s convenience and supplemented where necessary. In short, crowded and uncrowded visual acuity was measured binocularly at 5 m (direct view) with the Freiburg visual acuity test (FrACT) software which employs a 24-trial staircase
procedure (best PEST\textsuperscript{174}) to determine the subject’s threshold\textsuperscript{173}. Stimuli were black Landolt-Cs (2.1 cd/m\textsuperscript{2}) with four possible orientations against a white background (235.6 cd/m\textsuperscript{2}) on a 23-inch LCD screen (Dell U2412M, 1920 x 1200 pixels, pixel pitch 0.27 mm). The children indicated the perceived orientation of the Landolt-C by pointing or giving a verbal response, and could take as long as they felt necessary to respond.

Subsequently, children performed the speed-acuity test binocularly at 5m using the same screen, luminance levels and contrast. In this two alternative forced choice (2AFC) reaction time task children had to indicate, as quickly and accurately as possible, where the opening of a high contrast (98.2% Michelson) black Landolt-C was located by pressing one of two mouse buttons (Figure 3.1A). Optotype sizes (n=9), presented in pseudorandom order, ranged from -0.3 LogMAR below to 1.2 LogMAR above a child’s uncrowded binocular visual acuity (as determined with the FrACT) with steps of \~0.1 LogMAR around visual acuity and \~0.2 LogMAR for the larger optotype sizes (method of constant stimuli, 10 trials per optotype size). In children with NV, the range was fixed (-0.43 to 1.09 LogMAR). We employed a 2AFC paradigm, instead of the FrACT’s 4AFC, to simplify the stimulus-response associations and movement components of the task in order to encourage and facilitate quick responses.

Two additional experiments were performed to examine whether the children’s slower responses represented a general deficit in sensorimotor processes, or

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.1.png}
\caption{A. The speed-acuity test in which the children had to indicated as fast and accurately as possible, on which side the opening of the Landolt-C was located (right or left) by pressing the corresponding mouse button. B. The detection tasks in which the children had to press a button as soon as they perceived a supra-threshold stimulus, either a big “O” in the visual detection task (VDT) or a loud sound burst in the auditory detection task (ADT). Reproduced from\textsuperscript{196}.}
\end{figure}
whether they were specific to visual tasks requiring fine spatial discrimination. First, we measured their reaction times on a visual detection task (VDT) in which they had to press a button as soon as they saw a large (1.3 LogMAR), high contrast (98.2% Michelson), black “O” appear on the screen (Figure 3.1B, left). And, to further probe for other, perhaps more general (developmental) problems in sensory processing and/or movement execution, we also tested them with an auditory detection task (ADT) in which they had to press a button as soon as they heard a loud (75dBA) sound stimulus (Figure 3.1B, right).

The custom Matlab software (version 2013b; MathWorks, Inc., Natick, MA, USA) for the speed-acuity test, the ADT, and the VDT recorded and stored stimulus timing and button presses at 1 ms precision using the Psychophysics Toolbox (version 3.0.12)\textsuperscript{175}.

Data analysis

Data analysis was performed in Matlab using the statistical toolbox (version 2013b). First, mean reaction times were calculated. After discarding reaction times <0.1 sec, further trials were excluded if the reaction time deviated by more than 3 times the median absolute deviation (MAD) from the median\textsuperscript{176} for each test and each optotype size. On average 7\% of the trials were excluded for the speed-acuity test, 6\% for the VDT and 8\% for the ADT. Most of the excluded trials were actually the first trial in a test because the children often did not realize that the test had started. Two children with VI\textsubscript{o} (aged 7 and 8 years) could not perform the ADT due to technical problems and for one 5-year-old child with VI\textsubscript{o} the results on the ADT were excluded because the large number of premature responses (reaction time <100 ms in ~40\% of the trials) indicated non-compliance with the task. For two children with CVI (aged 7 and 8 years) data from the speed-acuity test had to be excluded from further analysis because they appeared to have been guessing even for large optotypes well above their visual acuity (the median of their accuracy scores for the five largest optotypes was lower than 87.5\% correct). For two children with VI\textsubscript{o} (aged 6 and 8) the data from the speed-acuity test were excluded because they did not finish the test.

For the speed-acuity test, the reaction time data and accuracy data (percent correct) as a function of optotype size were analyzed separately. The visual acuity was determined from the cumulative Gaussian function fitted to the accuracy data (75\% correct threshold, see Figure 3.2 for illustration)\textsuperscript{178}. The guess rate was fixed to the 50\% chance level performance for a 2AFC task and the lapse rate was allowed
to vary between 0 and 10%. The reaction times were analyzed in relation to the chronometric response functions obtained in children with NV\textsuperscript{196}.

The chronometric response functions of children with NV are well described by a reaction time model\textsuperscript{62} which uses the following hyperbolic tangent function\textsuperscript{196}:

\[
RT(x) = \begin{cases} 
A' k'(x - x_o) \tanh(A' k'(x - x_o)) + t_R, & x > x_o \\
A'^2 + t_R, & x \leq x_o 
\end{cases} \tag{3.1}
\]

in which \(x\) is the size of the optotype (in LogMAR) to which the child responds, and \(x_o = -0.43\) LogMAR the size of the largest optotype at which children with NV perform at, or close to, chance level (referred to as the critical optotype size). The parameter \(t_R\) is the residual time which indicates the minimum time a child needs to respond. The reaction-time difference between the responses for optotype sizes at which a child performs at chance level and those for the largest optotype size is reflected in the choice delay limit (\(A'^2\)). Parameter \(k'\) is a measure for the sensitivity of the visual system and is a scaling factor for the decrease in reaction times as optotype sizes increase. Furthermore, the age-dependency of the average reaction times curves for children with NV between 5 and 12 of age are well described by the following set of equations for the three model parameters \(A'\), \(k'\) and \(t_R\)\textsuperscript{196}:

\[
\begin{align*}
A' &= \exp(0.189 - 0.027 \cdot \text{Age}) \\
k' &= \exp(1.558 + 0.076 \cdot \text{Age}) \\
t_R &= 1.031 - 0.048 \cdot \text{Age}
\end{align*} \tag{3.2}
\]

Thus, by combining Equations 3.1 and 3.2, we could obtain age-matched, normative reaction times as a function of optotype size for each child with VI\textsubscript{o} and CVI given his/her Age (in years). The measured reaction times of each child were then compared to the age-matched control data from children with NV using a summery score, the delay index, which aggregates the reaction time differences across all optotypes\textsuperscript{196}:

\[
DI = \frac{1}{n_{opt}} \sum_{i=1}^{n_{opt}} \log(RT_{obs,i}) - \log(\bar{RT}_{mar,i}) \sqrt{\text{VAR}_i} \tag{3.3}
\]
The DI was computed as the mean of the differences in response time between the child’s reaction time ($RT_{\text{obs},j}$) and the age-matched norm value ($\hat{RT}_{\text{norm},j}$, obtained using Equations 3.1 and 3.2) at each optotype size ($j$). These reaction times were log transformed to account for the fact that their distributions are skewed towards longer reaction times (see, e.g., the 95% prediction intervals of the chronometric curves in Figure 3.2, bottom panels). The variance term ($VAR_j$) was included in the denominator of the index to normalize the delay scores per optotype size, because the variance of $\log(RT_{\text{obs},j})-\log(\hat{RT}_{\text{norm},j})$ increases with decreasing optotype size. $VAR_j$ was estimated from Equation 3.1 with parameter values $x_o=-0.43$, $A'=0.37$, $k'=37.27$ and $t_R=0.0157$. In this way, the delay index provides an age- and optotype-size invariant measure of a child’s response delay. Only optotype sizes < 1.1 LogMAR were used to calculate the DI, because this was the largest optotype size used in the children with NV. 95% of the children with NV have a DI ≤ 1.39 SD units. A child with VI or CVI was therefore considered slow if its DI exceeded this 95th percentile score.

Note, that the standard DI from Equation 3.3 does not account for the reduced visual acuity of children with VI and CVI. Therefore, an acuity-adjusted DI was calculated as well. To compute this acuity-adjusted DI, the age-matched, normative reaction times curves in Equation 3.3 were shifted towards the right (i.e., towards larger optotype sizes) by modifying the value of the critical optotype size, $x_o$, in Equation 3.1 according to a child’s response accuracy. That is, the critical optotype size of a child was estimated from the slope of his/her psychometric response function at the 75% correct performance level by calculating the optotype size at which the tangent line at this inflection point intersects the 50% correct chance level (Figure 3.2, upper panels). As a result, the acuity-adjusted DI accounted for the child’s reduced visual acuity. Since the reaction time model of Equation 3.1 could adequately describe the reaction times of the children with visual impairments if it included $x_o$ as an additional free fit parameter (the average $R^2$ of individual fits was $0.82 \pm 0.18$; not shown), an acuity-adjusted DI of zero is expected if reduced visual acuity alone accounts for their response delays.

Statistical analysis

Our goal was to determine whether children with visual impairments are slower than children of the same age with NV. Therefore, we performed linear regression analyses with both age and group (NV, VI and CVI) as independent variables. For all
regression analyses age was centered on the age of 9, the middle of the inclusion range. In this way, the differences between the regression-line intercepts quantify the mean reaction-time differences at the age of 9. Unless stated otherwise, these vertical offsets adequately quantified the age-adjusted differences between the groups since there were no significant interactions between age and group (i.e., the slopes of the regression lines were not significantly different).

Type-I error was set at 0.05 for all statistical group comparisons. Values are reported as means ± 1 standard deviation (SD). An individual child was classified as slow on a given task if his or her score (reaction time or DI) exceeded the upper 95th percentile of the normative data.

Results

Speed-acuity test: Visual acuity

Figure 3.2 shows the psychometric and chronometric response curves for two visually-impaired children, as well as the normative chronometric curves for normally-sighted children of the same age (blue). The mean visual acuity estimated from the psychometric response functions was 0.30 ± 0.25 LogMAR for the children with VI<sub>o</sub> and 0.16 ± 0.29 LogMAR for the children with CVI. Despite the time pressure in the speed-acuity task, these averages were significantly lower than the mean uncrowded acuity found with a standard C-test, the FrACT (VI<sub>o</sub>: difference -0.07 ± 0.11 LogMAR, Wilcoxon signed rank test, Z=-2.94, p=0.003, CVI: difference -0.08 ± 0.12 LogMAR, Wilcoxon signed rank test, Z=-2.22, p=0.03). However, the absolute intraclass correlation between the two acuity measures of all children was 0.88 (95% CI: 0.70-0.95), which demonstrates high agreement. For 56% (24/43) of the children the difference between the two acuity measures was within 0.1 LogMAR, for 84% (36/43) of the children the difference was within 0.2 LogMAR and for 95% (41/43) of the children the difference was within 0.25 LogMAR. There was no significant effect of age on the uncrowded visual acuity in either group (VI<sub>o</sub>: t(28)=1.78, p=0.08; CVI: t(15)=1.71, p=0.11).

Speed-acuity test: Reaction times

The chronometric response curves for the two children with visual impairments in the lower panels of Figure 3.2 (red curves) show that both children were typically slower compared to 95% of the children with NV of the same age (blue). For the child in Figure 3.2A only the reaction times for the smaller optotype sizes were longer,
Discrimination speed in children with visual impairments

whereas for the largest optotype sizes this child performed within the normal range (shaded area). If the chronometric curve of this child is shifted along the horizontal axis to align its critical optotype size (dashed red line) with the critical optotype size of children with NV, the chronometric curve would fall within the normal range. Thus, this child had longer reaction times on the speed-acuity test compared to the children with NV, but if its reduced visual acuity is taken into account, these longer reaction times are to be expected. For the child in Figure 3.2B, on the other hand, the longer reaction times cannot be explained by his/her reduced visual acuity alone.

The reaction times on the largest optotypes are an indication of the minimum time children needed to discriminate the orientation of the Landolt-C. Therefore, we first analyzed the children’s mean reaction time for these easiest optotypes to determine whether the children with VIo and CVI were slower in discriminating
large optotypes (Figure 3.3). For children with VIo and CVI the reaction time on the easy optotypes was taken as the average of the reaction times for the two largest optotypes (> 0.5 LogMAR above their visual acuity). For the children with NV it was the average reaction time for optotypes >0.5 LogMAR. Linear regression analysis (Supplemental Table S3.2) indicated that as a group, the VIo and CVI children responded significantly later than the controls. Compared to the children with NV, the intercept of the regression line was 170 ± 28 ms higher for the children with VIo ($t(129)=6.16, p<0.001$), and 232 ± 36 ms higher for the children with CVI ($t(129)=6.49, p<0.001$). Of the children with VIo, 57% (16/28) had reaction times above the 95th percentile of children with NV and were therefore slower than normal, while 60% (9/15) of the children with CVI were slower than normal.
Discrimination speed in children with visual impairments

Delay index

The analysis shown in Figure 3.3 does not take into account the entire chronometric curve. The DI, on the other hand, provides a summary measure for the response delay by comparing the reaction times for all optotype sizes to the reaction times of children with NV in standard deviation units. A DI value of 0 signifies that a child responds as fast as an average child with NV of the same age, while a DI value of, e.g., 2 indicates that his/her reaction time curve is on average 2 standard deviations above the average performance of children with NV. As shown in Figure 3.4A, the average DI was 3.3 ± 1.9 for children with VI₀ and 3.0 ± 2.2 for children with CVI, which was significantly above the DI=0 score of children with NV (Mann-Whitney U-test, Z =7.7, *p*<0.001 and Z=5.6, *p*<0.001). The DI’s for the two individual subjects shown in Figures 3.2A and 3.2B were 2.94 and 6.94, respectively. Only one child
with VLo (4%, 1/28) and only two children with CVI (13%, 2/15) scored under the 95\textsuperscript{th} percentile of children with NV. This indicates that the vast majority of the children with VLo and CVI were overall slower on the speed-acuity test if we directly compare their reaction time curves to the reaction time curves of the children with NV of the same age. The DIs were not significantly different between the two clinical groups (Mann-Whitney U-test, $Z = 1.26, p=0.21$).

To test whether the longer reaction times may be explained by the children’s reduced visual acuity, acuity-adjusted DIs were calculated (Figure 3.4B). In this case the optotype sizes were redefined relative to the critical optotype size, i.e., the optotype size at which the children performed at chance level. As expected, the average acuity-adjusted DIs were lower than the standard DIs for the children with VLo ($1.1 ± 1.1$, Wilcoxon signed rank test, $Z = -4.6, p<0.001$) as well as for the children with CVI ($1.6 ± 1.6$, Wilcoxon signed rank test, $Z = -3.4, p<0.001$). For 60% of children the acuity-adjusted DIs fell within normal range (VI: 18/28 children, CVI: 8/15 children). However, a large proportion of the children had acuity-adjusted DIs that were still outside the normal limits: 36% (10/28) of those with VLo and 47% (7/15) of those with CVI had scores above the 95\textsuperscript{th} percentile of children with NV. This indicates that these children were significantly slower on the speed-acuity test than one might expect from their impaired visual acuity alone. In fact, the average acuity-adjusted DIs were still significantly higher than the standard DI of children with NV (Mann-Whitney U-test, VLo: $Z=4.3, p<0.001$ and CVI: $Z=3.6, p<0.001$), a feature that was not significantly different between the two groups of patients (Mann-Whitney U-test, $Z = -0.57, p=0.57$).

**Detection tasks**

The children also performed a visual detection task (VDT) to examine whether they had difficulties in just detecting the appearance of a salient visual stimulus. The regression analysis of these data revealed that both clinical groups were significantly slower than the controls (Figure 3.5A, Supplemental Table S3.3). In children with VLo reaction times were on average $79 ± 15$ ms longer than the children with NV ($t(132) = 5.19, p<0.001$), whereas children with CVI reacted on average $144 ± 19$ ms later than children with NV ($t(132)= 7.72, p<0.001$). In fact, half (15/30) of the children with VLo and 76% (13/17) of the children with CVI scored above the 95th percentile norm, indicating that these children were already late in detecting large, high-contrast visual stimuli.
To test if the significant delays in detecting the visual stimulus might have been due to other, more general (developmental) problems rather than visual impairments alone, we also tested the children in an auditory detection task (ADT). The regression analysis of these data (Figure 3.5B, Supplemental Table S3.3) revealed that the children with CVI responded on average 106 ± 14 ms later than normal on the ADT \((t(127)=7.51, p<0.001)\). In total, 59\% (10/17) of the children with CVI showed abnormally late reactions on the ADT (i.e., above the 95th percentile norm). To our surprise, 26\% (7/27) of the children with VIo also needed significantly more time to detect the sound stimulus than children with NV. Despite their normal hearing, the
group of children with VIo was significantly slower on the ADT than the children with NV (t(127)=2.52, p=0.01). The mean reaction-time difference was only 29 ± 12 ms, but this was still ~10% slower than normal.

In contrast to the speed-acuity test and the VDT, in which the two clinical groups showed comparable age-related improvements in reaction times than children with NV (Figure 3.3 and Figure 3.5A), the decrease in reaction time with age on the ADT seems to differ in the CVI group. More specifically, the slope of the regression line for children with CVI was significantly steeper than the slope of the regression line for children with NV (slope CVI: -55 ±10 ms/year, slope NV: -18 ± 3 ms/year, t(127)=-3.65, p<0.001). This effect is mainly attributable to the results of the younger children (<9 years) with CVI. For the children with VIo, the slope of the regression line was not significantly different from the control line (slope VIo: -15 ± 5 ms/year, t(127)=0.64, p=0.52).

The results from the detection tasks indicate that children with CVI tend to respond quite late on both the VDT and the ADT, while children with VIo only exhibit pronounced delays on the VDT. This suggests that in children with CVI the response delays might be caused by more general deficits, whereas in children with VI the lagging reactions may be primarily due to visual problems. If this is the case, than one would expect that the within-subject difference between the VDT and ADT in children with CVI does not exceed the reaction-time differences found in controls, whereas larger reaction-time differences are expected in children with VIo. Although the data are noisy, linear regression analysis confirmed these expectations (Figure 3.5C). The difference in reaction time between the VDT and the ADT was 28 ± 13 ms larger for the children with VIo compared to the controls (t(127)=2.18, p=0.03), while the regression line in the children with CVI was not statistically significant from the control line (difference at age 9: 18 ± 15 ms, t(127)=1.14, p=0.26).

**Discrimination versus detection**

Children with visual impairments had abnormally long reaction times on the speed-acuity test (Figure 3.3) and the two detection tests (Figure 3.5). The question is whether the longer reaction times on the speed-acuity test can be explained by a lag in stimulus detection alone, or whether the stimulus discrimination process takes longer too. Therefore, we compared the reaction times for the easy optotypes with the reaction times on the detection tasks (Figure 3.6, Supplemental Table S3.4). The vertical offset of the regression lines revealed that, on average, the difference
between the reaction time for the easy optotypes and the VDT was 92 ± 24 ms larger for the children with VIO than for the children with NV (t(127) = 3.79, p<0.001). For the children with CVI, the difference was on average 99 ± 31 ms larger (t(127) = 3.14, p=0.002). This indicates that children with VIO and CVI discriminated the symbols significantly later than one might expect from their increased reaction times in the VDT alone. In total, 25% (7/28) of the children with VIO and 40% (6/15) of the children with CVI needed more time than expected from their increased reaction times in the VDT alone. Thus, even after correcting for the time children needed to detect and respond to a visual stimulus, children with VIO and children with CVI needed significantly more time to discriminate and respond to the easy optotypes, which suggests that the sensory discrimination process is slower in these children.
compared to children with NV. Similar results were found for the difference between the reaction time on the easy optotypes and the reaction time on the ADT. The vertical offset of the regression lines revealed that the children with VIo reacted on average 116 ± 27 ms later than one might expect from their increased reaction times in the ADT alone \((t(124)=4.30, p<0.001)\) and in the children with CVI this extra delay in responding was on average 95 ± 34 ms \((t(124)=2.78, p=0.006)\). After correcting for (increases in) reaction times on the ADT, 38% (10/26) of the children with VIo and 33% (5/15) of the children with CVI were still abnormally late in discriminating symbols.

**Discussion**

Poor vision may affect not just the accuracy, but also the speed of visual judgments. In the present study we assessed both aspects simultaneously in 5-12 year-old children with VIo and CVI. As we expected, almost all children with VIo and CVI were slower on the speed-acuity test than children with NV for the same optotype sizes. This deficit was reflected in significant elevations of the delay index, a summary score that takes the entire reaction time curve into account (Figure 3.4A). For 60% of the children, this impairment in visual discrimination speed could be explained by their reduced acuity as indicated by their acuity-adjusted delay index (Figure 3.4B). However, the remaining 40% of the children still had longer reaction times than one would expect from their reduced visual acuity alone. This indicates that even if materials are magnified according to their visual acuity, a large number of children with VIo and CVI still need more time to discriminate visual stimuli. Thus, magnification may be insufficient to compensate for delays in visual discrimination. This conclusion is further supported by the fact that many of the children with VIo and CVI were abnormally slow in discriminating the largest symbols (Figure 3.3), even if one accounts for the possibility that this may have been due to slow stimulus detection and/or movement execution (Figure 3.6). Some of the children whose delay index fell out of range had normal reaction times on the easiest optotypes. This is an indication that these children are slow at discerning small optotypes but that they may be able to perform at normal speed if materials are large enough. However, for children with an elevated delay index and longer reaction times on the largest optotypes, magnification of materials alone cannot fully compensate for their deficit.

**Relation to previous research**

The observed delays in symbol discrimination are in line with a recent study showing
that children with visual impairments have lower maximal reading speed on the MNREAD test compared to age-matched controls. Our findings suggest that at least part of these reading deficits may be explained by a reduction in the basic ability to discern symbols in a timely manner. This deficit could also be one of the underlying causes of longer search times in children with visual impairments. Other studies have shown that children with visual impairments need more time to perform fine motor tasks. Although the tasks in our study did not require a complicated motor response, problems with fine motor skills could have resulted in longer reaction times for some of the children. Perhaps this might account for the unexpected finding that some of the children with VIo needed more time on the ADT. Conversely, it is quite likely that the reduced speed of visual processing, as demonstrated by our findings, contributes significantly to the extra time that children with visual impairments need to perform fine motor tasks.

**Study population**

We included children with a wide variety of ophthalmological diagnoses and underlying causes of CVI. This reflects the heterogeneity of visual impairments that is present in children in western Europe. It is therefore no surprise that the delay indices obtained with the speed-acuity test could not distinguish between the two clinical groups. Nevertheless, children with CVI were typically slower than children with VIo in both detection tasks and to some extent in discriminating large optotypes. Our finding that the group of children with CVI showed comparatively large delays in the VDT and similarly large delays in the ADT (Figure 3.5), suggests that general sensorimotor deficits and/or developmental delays played a more prominent role in this group, whereas the impairments in children with VIo seemed mostly due to slower visual processing alone. For individual children, however, the pattern could differ. Some children who responded late in the speed-acuity test had normal reaction times on the detection tasks, indicating that their response delays were not caused by delays in detection of the stimuli, or the ensuing motor response. For others, the increased reaction times on the easiest optotypes in the speed-acuity test were paralleled by increased reaction times on the visual or auditory detection tasks. For these children, the increase in reaction time can be explained by delays in the detection of stimuli and/or the increase in time that they need to generate the motor output. Thus, children with CVI and VIo showed individual differences, as might be expected with the diversity of causal factors and developmental differences.
Chapter 3

Limitations

Both clinical groups showed significant response delays on the speed-acuity test, even after correction for their impaired visual acuity, but the limited number of children that we could include with a particular diagnosis makes it difficult to draw conclusions about whether children with certain diagnoses or characteristics (e.g., retinal abnormalities, nystagmus or strabismus) have a higher risk of being slower. Furthermore, a large group of children with CVI, the ones with significant motor problems and/or significant cognitive developmental delays, could not be included in this study because they lack the skills needed to perform the tasks. Therefore, our results may not generalize to children with more severe CVI. Most likely, their speed of visual processing is even more severely affected. When children with significant cognitive delays are included in a study, it might be especially beneficial to use developmental age as an additional covariate in the analysis (although this is perhaps easier said than done, since, to our knowledge, there is no unequivocal measure of developmental age). In the current analyses, we only used the chronological age of the children as a covariate because data on their developmental age was not available. We believe, however, that the detection tasks already control for possible developmental delays, at least to some extent (see also\textsuperscript{157} for a similar argument). We therefore argue that developmental delays alone cannot account for the reduced symbol discrimination speeds in the speed-acuity test that we have found in the children with CVI.

Developmental differences

Children with normal vision show large developmental improvements in their reaction times, both on the detection tasks and on the speed-acuity test\textsuperscript{196}. Despite being slower than normal, children with VI\textsubscript{o} seem to show similar age-related improvements on all tasks as suggested by the fact that the slopes of the regression lines were not statistically different from the slopes of regression lines for children with NV (Supplemental Tables S3.2-S3.4). The reaction times of the children with CVI on the ADT, but not on the other tasks, changed with age in a manner different to that of the NV children. This might indicate that the children with CVI develop differently compared to the children with NV. However, this effect could also be the result of the relatively low sample size and the heterogeneity of the group. Longitudinal studies and larger groups are necessary to conclude whether or not children with VI\textsubscript{o} and CVI develop differently on the tasks we have used.
Conclusion

The speed-acuity test can be used in clinical pediatric populations and may provide important diagnostic insight for clinical and rehabilitation purposes by assessing visual recognition acuity and speed simultaneously. A significant number of children with Vlo and CVI were slower on this test compared to children with NV, even if their reduced visual acuities are taken into account. We therefore argue that visual acuity measures alone do not adequately capture the ability to discern foveal details quickly. Thus, measures which also assess the speed of visual processes should be employed, given that this aspect of visual processing may be crucial for a child’s normal participation in school and society. The observed deficits seem to be caused by (at least) two bottlenecks: one related to simply detecting and responding to visual stimuli, and the other one related to slower accumulation of sensory evidence to discriminate between visual stimuli.
### Supplemental Table S3.1.
Clinical characteristics of the children with their participant number (Pp), age, group, diagnosis, binocular crowded and uncrowded visual acuities at 5 m, and the presence of strabismus and nystagmus (+: manifest, +/-: latent, -: absent).

<table>
<thead>
<tr>
<th>Pp</th>
<th>Age</th>
<th>Group</th>
<th>Diagnosis</th>
<th>Crowded VA (LogMAR)</th>
<th>Uncrowded VA (LogMAR)</th>
<th>Strabismus</th>
<th>Nystagmus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>VI</td>
<td>Optic nerve hypoplasia</td>
<td>1.0</td>
<td>0.53</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>VI</td>
<td>Hypermetropia (refractive error &gt;4D)</td>
<td>0.49</td>
<td>0.18</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>VI</td>
<td>Incontinentia pigmenti</td>
<td>0.69</td>
<td>0.40</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.49</td>
<td>0.33</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.69</td>
<td>0.41</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>CVI</td>
<td>Noonan syndrome</td>
<td>0.48</td>
<td>0.46</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>CVI</td>
<td>Noonan syndrome</td>
<td>0.32</td>
<td>0.35</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.46</td>
<td>0.26</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>VI</td>
<td>Hypermetropia (refractive error &gt;4D)</td>
<td>0.39</td>
<td>0.04</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>VI</td>
<td>Albinism</td>
<td>-</td>
<td>0.29</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>VI</td>
<td>Macular hypoplasia</td>
<td>0.36</td>
<td>0.23</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.53</td>
<td>0.47</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>-</td>
<td>1.21</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.32</td>
<td>0.31</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>VI</td>
<td>Albinism</td>
<td>-</td>
<td>0.54</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.34</td>
<td>0.31</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>12</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.38</td>
<td>0.41</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>7</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.57</td>
<td>0.10</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>VI</td>
<td>Albinism</td>
<td>0.68</td>
<td>0.59</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>VI</td>
<td>Macular atrophy</td>
<td>0.64</td>
<td>0.24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>10</td>
<td>VI</td>
<td>Cone-rod dystrophy</td>
<td>1.13</td>
<td>0.82</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>11</td>
<td>CVI</td>
<td>Status after meningitis and cerebral infection</td>
<td>1.26</td>
<td>0.96</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>VI</td>
<td>Hypermetropia (refractive error &gt;4D)</td>
<td>0.38</td>
<td>0.05</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>8</td>
<td>VI</td>
<td>Albinism</td>
<td>0.28</td>
<td>0.11</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>VI</td>
<td>Albinism</td>
<td>0.66</td>
<td>0.35</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>26</td>
<td>7</td>
<td>CVI</td>
<td>Motor, cogn., and visual delay (35W, &gt;2500g)</td>
<td>0.41</td>
<td>0.33</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>11</td>
<td>VI</td>
<td>Albinism</td>
<td>0.70</td>
<td>0.48</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
### Discrimination speed in children with visual impairments

<table>
<thead>
<tr>
<th>Pp</th>
<th>Age</th>
<th>Group</th>
<th>Diagnosis</th>
<th>Crowded VA (LogMAR)</th>
<th>Uncrowded VA (LogMAR)</th>
<th>Strabismus</th>
<th>Nystagmus</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>10</td>
<td>VI</td>
<td>Myopia and retinal scarring</td>
<td>0.82</td>
<td>0.70</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>29</td>
<td>8</td>
<td>CVI</td>
<td>Optic nerve atrophy, microcephalus and bilateral occipital infarcts</td>
<td>0.43</td>
<td>0.35</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>30</td>
<td>7</td>
<td>CVI</td>
<td>Prematurity and dysmaturity (26W, 750g)</td>
<td>0.29</td>
<td>0.12</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>31</td>
<td>7</td>
<td>VI</td>
<td>Cone dysfunction (Bornholm)</td>
<td>0.54</td>
<td>0.44</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>32</td>
<td>6</td>
<td>VI</td>
<td>Coloboma of the iris and retina ODS, and optic nerve OS</td>
<td>0.39</td>
<td>0.19</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>VI</td>
<td>Albinism</td>
<td>0.48</td>
<td>0.38</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>34</td>
<td>8</td>
<td>CVI</td>
<td>Joubert syndrome</td>
<td>0.34</td>
<td>-0.01</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>35</td>
<td>9</td>
<td>CVI</td>
<td>Cerebral arteriovenous malformation, resulting in 2 strokes.</td>
<td>-0.15</td>
<td>-0.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>9</td>
<td>VI</td>
<td>Aniridia</td>
<td>0.64</td>
<td>0.60</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>37</td>
<td>12</td>
<td>VI</td>
<td>Albinism</td>
<td>0.66</td>
<td>0.48</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>38</td>
<td>8</td>
<td>CVI</td>
<td>Prematurity (25W, 990g)</td>
<td>0.13</td>
<td>-0.20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>39</td>
<td>11</td>
<td>CVI</td>
<td>White matter damage due to mitochondrial disease and internuclear ophthalmoplegia</td>
<td>0.49</td>
<td>0.31</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>11</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.27</td>
<td>0.27</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>41</td>
<td>8</td>
<td>CVI</td>
<td>Perinatal complications, Stickler syndrome</td>
<td>0.02</td>
<td>-0.01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>42</td>
<td>7</td>
<td>CVI</td>
<td>Prematurity (32W, 2180 g)</td>
<td>0.38</td>
<td>0.18</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>43</td>
<td>8</td>
<td>CVI</td>
<td>Prematurity (32W, 1900g) and perinatal complications</td>
<td>0.09</td>
<td>-0.09</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>44</td>
<td>10</td>
<td>CVI</td>
<td>Perinatal complications (hemiparesis)</td>
<td>0.27</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>45</td>
<td>7</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.80</td>
<td>0.66</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>46</td>
<td>8</td>
<td>CVI</td>
<td>Perinatal complications</td>
<td>0.49</td>
<td>0.14</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>7</td>
<td>CVI</td>
<td>Microcephalus, partial cataract and coloboma of the ins.</td>
<td>0.38</td>
<td>0.25</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Supplemental Table S3.2. Regression results for the effect of group (NV, VI, and CVI) and age on the reaction times on the large optotypes in the speed-acuity test.

<table>
<thead>
<tr>
<th>Estimated coefficients</th>
<th>β</th>
<th>SE</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>720</td>
<td>13</td>
<td>53.5</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>170</td>
<td>28</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>232</td>
<td>36</td>
<td>6.5</td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-61</td>
<td>7</td>
<td>-8.8</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>7</td>
<td>12</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>10</td>
<td>23</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Supplemental Table S3.3. Regression results for the effect of group (NV, VI, and CVI) and age on the reaction times on the detection tasks (VDT: visual detection task, ADT: auditory detection task).

<table>
<thead>
<tr>
<th>Estimated coefficients</th>
<th>β</th>
<th>SE</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>341</td>
<td>8</td>
<td>45.4</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>79</td>
<td>15</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>144</td>
<td>19</td>
<td>7.7</td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-24</td>
<td>4</td>
<td>-6.2</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>-0.8</td>
<td>7</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>-9</td>
<td>12</td>
<td>-0.8</td>
</tr>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>277</td>
<td>6</td>
<td>49.8</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>29</td>
<td>11</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>106</td>
<td>14</td>
<td>7.5</td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-18</td>
<td>3</td>
<td>-6.1</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>3</td>
<td>5</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>-37</td>
<td>10</td>
<td>-3.7</td>
</tr>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>65</td>
<td>6</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>28</td>
<td>13</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>18</td>
<td>15</td>
<td>1.1</td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-7</td>
<td>3</td>
<td>-2.2</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>3</td>
<td>6</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>0.8</td>
<td>11</td>
<td>0.1</td>
</tr>
</tbody>
</table>
**Supplemental Table S3.4.** Regression results for the effect of group (NV, VI, and CVI) and age on the difference between the reaction times on the large optotypes in the speed-acuity test and the reaction times on the detection tasks (VDT: visual detection task, ADT: auditory detection task).

<table>
<thead>
<tr>
<th></th>
<th>Estimated coefficients</th>
<th>β</th>
<th>SE</th>
<th>t-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Large optotypes - VDT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R²=0.35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>376</td>
<td>12</td>
<td>31.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>92</td>
<td>24</td>
<td>3.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>99</td>
<td>31</td>
<td>3.1</td>
<td>0.002</td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-36</td>
<td>6</td>
<td>-5.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>15</td>
<td>11</td>
<td>1.4</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>13</td>
<td>21</td>
<td>0.6</td>
<td>0.52</td>
</tr>
<tr>
<td>Intercept at 9 years (ms)</td>
<td>NV</td>
<td>440</td>
<td>13</td>
<td>33.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>116</td>
<td>27</td>
<td>4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Large optotypes - ADT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R²=0.40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (ms/year)</td>
<td>NV</td>
<td>-42</td>
<td>7</td>
<td>-6.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VI vs NV</td>
<td>22</td>
<td>12</td>
<td>1.8</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>CVI vs NV</td>
<td>-26</td>
<td>22</td>
<td>-1.2</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Optics of the human cornea influence the accuracy of stereo eye-tracking methods: a simulation study

Published as: Barsingerhorn AD, Boonstra FN & Goossens J. Optics of the human cornea influence the accuracy of stereo eye-tracking methods: a simulation study. *Biomedical Optics Express*. 2017; 8(2), 712-725
**Introduction**

Eye-tracking methods are used extensively in a wide range of research fields, such as studies on attention, visual search, reading and human-computer interaction\(^{200-202}\). Different types of eye-tracking systems exist (see e.g.\(^{100,112}\) for extensive overviews). The scleral search coil technique measures electromagnetic induction in a copper coil which is embedded in a contact lens\(^{104,105}\). Electro-oculography measures the electrical potential of the cornea-retinal dipole using electrodes placed on the skin\(^{107}\). However these methods are either relatively invasive and uncomfortable, or noisy and prone to drift, which makes them less suitable to use in clinical populations or in children. Therefore, there is a strong preference to use non-invasive eye-trackers. Currently, the majority of non-invasive eye-tracking systems are video-based. Images of the eye are captured by one or two video-cameras, and fast image-processing techniques calculate the gaze direction. Most systems estimate the point-of-gaze from images of the pupil, together with one or more corneal reflections, also called Purkinje images or glints, produced by external infrared (IR) light sources\(^{114,119,203,204}\). The different approaches to estimate the point-of-gaze, or the direction of gaze, can be categorized into 2D interpolation-based methods and 3D model-based techniques\(^{88}\). 2D interpolation methods estimate gaze by relating image features to 2D gaze coordinates on a screen by means of empirical mapping functions, whereas 3D model-based approaches estimate gaze direction from a geometrical model of the eye.

Traditionally, video-based eye-trackers require an in vivo calibration procedure to determine the mapping between image features and gaze in 2D systems, and to estimate parameters of the eye, such as the radius of corneal curvature, in 3D systems. This calibration procedure involves the fixation of several small visual targets at known locations by the test subject. However, certain participants, for example, infants and people with oculomotor problems, poor visual acuity, or cerebral visual impairment, are unable to perform such a fixation task reliably. To cope with this problem, researchers have developed alternative methods that make the calibration less dependent on such individual procedures. Several methods have been suggested to simplify the calibration process. Shih and colleagues were the first to propose a 3D method in which two cameras and two IR light sources are used to compute the optical axis of the eye from estimates of the location of the center of the pupil and the center of the corneal curvature\(^{205}\). Subsequently, Guestrin and Eizenman generalized the geometrical model to fully calibrated set-ups\(^{119}\). In these
set-ups the internal camera parameters such as the focal length are known, as well as the position and orientation of the cameras and light sources. Other researchers have proposed similar 3D methods to compute the optical axis of the eye from the pupil center and corneal reflections (e.g. 120, see 88 for an overview of the different approaches). Alternative methods have been proposed that estimate the optical axis from the shape of the limbus 206, ellipse descriptions of the pupil contour combined with conic algebra 118, or corneal reflections and pupil center combined with pupil-contour data 121. The main advantage of these stereo eye-tracking methods compared to other video-based eye-tracking methods is that, in principle, only one calibration point is needed to estimate the deviation between the optical and visual axes.

Important for all methods using pupil data is that the actual pupil is not observed, but a virtual image of the pupil (entrance pupil) due to the refractive power of the cornea. In general, it is assumed that the entrance pupil lies in front of the actual pupil, that the entrance pupil is slightly larger, and that the optical axis goes through the center of the actual pupil and the entrance pupil. Chen et al. investigated this assumption through ray-tracing for a stereo-camera set-up and found small deviations of the virtual pupil, with a resulting error in simulated gaze estimates of 0.08 degrees on average 120. However, they applied a simplified geometrical model of the eye, in which the cornea was modelled as a convex sphere with only one refractive surface. However, the actual cornea has two refractive surfaces: the anterior surface and the posterior surface, each having a different radius and center. As a result, light is refracted differently depending on the position on the cornea. Additionally, the anterior surface of the cornea is not a perfect sphere, but is slightly aspheric. Models that account for this aspect of the anatomy of the eye do exist, such as the Navarro eye model 207. Fedtke et al. used this model to estimate the location and shape of the entrance pupil as a function of viewing angle 208. The results revealed that the entrance pupil moves forward, tilts and curves towards the observer as the viewing angle increases, and the geometric mid-point of the entrance pupil departs from the optical axis. Moreover, as pupil size increases, the deviation from the optical axis increases 208. In addition, previous research has already revealed the influence of pupil size on the accuracy of gaze estimations in video-based eye-tracking 209 and the influence of gaze position on the estimates of pupil size 210.

However, the implications of these previous findings for the accuracy of stereo eye-trackers are unclear. As both cameras observe a different entrance pupil, the 3D reconstruction of the entrance pupil might not align with the actual pupil. Moreover,
the effect of more complicated and anatomically more accurate models of the eye on the reconstruction of gaze using stereo eye-trackers has not been investigated either. Therefore, the first aim of our present study was to estimate the effect of the anatomy of the eye on the accuracy of 3D methods to reconstruct the direction of gaze using video-based stereo eye-trackers. The entrance pupil for each camera also depends on the viewing angle and pupil diameter. As a result, head translations, gaze angles as well as changes in pupil diameter may all influence the accuracy of stereo eye-trackers. The second aim is to quantify the effect of head translation and pupil diameter on gaze estimations.

**Methods**

**Ray-tracing model**

The Navarro schematic eye model (Table 4.1) was used to simulate the virtual pupil. Optical modelling was done by ray-tracing in Matlab (version R2013b, The MathWorks Inc, Natick, MA, USA). To model a stereo eye-tracker set-up, two virtual cameras with $2048 \times 1088$ pixels and a focal length of 3000 pixels (16.5 mm with a pixel size of 0.0055 mm) were used in a right handed coordinate system with the origin between the nodal points of the cameras. Figure 4.1 shows a top view of this set-up.

Table 4.1. Parameters of the Navarro schematic eye model. The surface of the anterior cornea is described by the formula $x^2 + y^2 + (1 + Q)z^2 - 2Rz = 0$, where $Q$ is the conic constant and $R$ is the radius of curvature.

<table>
<thead>
<tr>
<th></th>
<th>Radius of curvature (mm)</th>
<th>Asphericity (Q)</th>
<th>Center of corneal curvature (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anterior cornea</strong></td>
<td>7.72</td>
<td>-0.26</td>
<td>7.72</td>
</tr>
<tr>
<td><strong>Posterior cornea</strong></td>
<td>6.5</td>
<td>0</td>
<td>7.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Thickness (mm)</th>
<th>Refractive index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cornea</strong></td>
<td>0.55</td>
<td>1.367</td>
</tr>
<tr>
<td><strong>Aqueous</strong></td>
<td>3.05</td>
<td>1.3374</td>
</tr>
</tbody>
</table>
The eye model was placed at nine different 3D positions to simulate head translations in the fronto-parallel plane ($X \in [0, 30, 60] \text{ mm}$, $Y \in [0, 30, 60] \text{ mm}$, $Z = 400 \text{ mm}$), and 81 different gaze directions (horizontal and vertical angles of $-15$, $-10$, $-5$, $-2.5$, $0$, $2.5$, $5$, $10$ and $15$ degrees). In addition, six pupil diameters ranging from $1.0$ to $6.0$ mm in $1.0$ mm steps were analyzed. To isolate the effects of pupil translation and pupil rotation, the actual pupil was rotated around the center of the pupil instead of the center of rotation of the eye. This prevented translations of the pupil due to the rotation of the eye.

First, we created a general ray-tracing model of the pupil (Figure 4.2). The model contained 32 equally spaced points on the pupil boundary to calculate the position and orientation of the entrance pupil. For each point on the pupil boundary, 125751 rays were aimed from that pupil point to the posterior surface of the cornea, and subsequently the refracted rays were computed using Snell’s law. This process was repeated for the anterior cornea. To calculate the corresponding point on the entrance pupil, it had to be determined which of these refracted rays would intersect the nodal point of each camera. Towards that end, the rays were first rotated according to the gaze direction. Subsequently, for each camera the minimal distance between the refracted rays and the nodal point of the camera was calculated and the ray with the smallest distance to the nodal point was used in an optimization procedure to find the optimal refracted ray. The optimal refracted rays for both cameras were triangulated to obtain the image point of the virtual pupil (Figure 4.2). After calculating all the virtual pupil points in 3D, we determined the location and orientation of the entrance pupil by fitting a plane through the 32 virtual image points and taking the orientation of the normal vector of this plane.
Gaze reconstruction

To test the accuracy of gaze reconstructions when using an anatomically accurate model of the human cornea in stereo eye-tracking, we compared two different stereo eye-tracking methods. For both methods, the 3D points of the virtual pupil were first projected onto a 2D image for each camera using the Camera Calibration toolbox of Bouguet\textsuperscript{211}. This procedure corresponds with the normal setting in which stereo eye-trackers are used to reconstruct the direction of gaze from two 2D images of the eye. The 2D images of the pupil were not pixelated in order to mimic the optimal situation without image noise. This ensures that a potential difference between the reconstructed gaze direction and the actual gaze direction is caused only by the optics of the eye, and is not confounded by additional image noise. A least-squares ellipse fit\textsuperscript{212} was then applied to the pupil boundary points in the respective 2D images.

The optical axis of the eye passes through both the center of the actual pupil and the center of curvature of the cornea (Figure 4.3, grey dashed line). Previously, it has been assumed that one can calculate the center of corneal curvature (CC) from the glint locations on the cornea, and that the center of the virtual pupil (VP) lies on the optical axis. If true, the VP-CC vector can be used to determine the orientation of the optical axis of the eye. However, this assumption is based on a simplified eye model.
Chapter 4

The first gaze reconstruction method uses the VP-CC vector (Figure 4.3, light blue arrow), to estimate the gaze direction (see, e.g.\textsuperscript{119,120}).

In the Navarro eye model, the center of the anterior corneal curvature is located 4.12 mm behind the center of the actual pupil on the optical axis (Figure 4.2). If a spherical model of the anterior cornea is used, the center of corneal curvature can be estimated using the reflections of two IR light sources (e.g.\textsuperscript{119}). The optics behind this is displayed in Figure 4.3A. Incident rays perpendicular to the surface, and thus coinciding with the normal vector at the surface, are reflected back along their own paths. For all other rays the angle between the normal vector and incident ray is equal to the angle between the normal vector and refracted ray (Figure 4.3A). If the cornea had a perfect spherical surface, then all the normal vectors at the cornea surface intersect at the center of corneal curvature (Figure 4.3B). However, if the corneal surface is aspherical, the normal vectors at the corneal surface will not intersect at one point (Figure 4.3C). In fact, it depends entirely on the point of reflection of the light sources on the cornea whether the corresponding normal vectors intersect on or off the optical axis of the eye, or whether they intersect at all. Therefore, algorithms which approximate the center of corneal curvature by converging into a single point, may give values that are not on the optical axis of the eye.

To determine the size of the estimation error, an algorithm with two simulated light sources was used to determine the location of the center of corneal curvature.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4_3.png}
\caption{Estimation of the center of corneal curvature (CC). A. The ray-trace model to estimate the CC based on a spherical anterior cornea, L1 and L2 are the light sources, C1 and C2 are the cameras. The incident ray is reflected at the anterior corneal surface. The angle of incidence $\theta_i$ is equal to the angle of reflection $\theta_r$. The normal vectors intersect at the center of corneal curvature. B. The intersecting normal vectors in case of a spherical corneal curvature. C. The intersecting normal vectors at the surface of an aspherical cornea.}
\end{figure}
Optics of the human cornea and the accuracy of stereo eye-tracking

The light sources were located 50 mm left from the left camera and 50 mm right from the right camera and 50 mm above the cameras. Figure 4.3 illustrates the reflection of light source L1 on camera C1 and the reflection of light source L2 on camera C2. In reality, there are four points of reflection, which include also the reflection of light source L1 on camera C2 and the reflection of light source L2 on camera C1. We used ray-tracing to determine the point of reflection of each light source for both cameras. The location at the cornea where the difference between the angle of incidence and the angle of reflection was minimal was estimated using an optimization procedure. Subsequently, the reflections were projected onto the 2D image plane of the cameras. These 2D image coordinates were then used to calculate the apparent location of the center of corneal curvature following the method proposed by [9]. This method assumes that the anterior cornea acts as a spherical mirror and that the center of curvature belongs to the plane defined by the camera C, the light source L and the corresponding image point U of the corneal reflection (Figure 4.3A). Because three coplanar vectors satisfy the constraint $v_1 \times v_2 \cdot v_3 = 0$, the center of corneal curvature can be estimated by finding the least-squares fit of Equation 4.1 for the four reflection points.

$$(L - C) \times (U - C) \cdot (CC - C) = 0$$ \hfill (4.1)

To obtain the center of the virtual pupil, the centers of the fitted ellipses to the pupil data from each camera were triangulated using the same internal and external camera parameters that were used to project the entrance pupils on the 2D image plane. Next, the horizontal (or pan) angle ($\phi$) and the vertical (or tilt) angle ($\theta$) were calculated from Equation 4.2, with $p$ the estimated center of the virtual pupil, and $c$ the location of the estimated center of corneal curvature.

$$\frac{p - c}{\|p - c\|} = \begin{bmatrix} \cos \phi \sin \theta \\ \sin \phi \\ -\cos \phi \cos \theta \end{bmatrix}$$ \hfill (4.2)

These results were compared to a hypothetical situation, in which the center of corneal curvature is known (Figure 4.3B). This fixed center of corneal curvature (fCC) was positioned 4.12 mm behind the center of the actual pupil on the optical axis, as defined by the Navarro eye model (Table 4.1). This simulation isolates the
effect of deviations of the center of the virtual pupil from the eye’s optical axis. The comparison of the results of the VP-CC vector method and the VP-fCC vector method shows how errors in estimating the center of corneal curvature influence the overall accuracy of the VP-CC method.

The second gaze reconstruction method applies conic algebra and a parametric description of the pupil boundary to estimate the gaze direction (Figure 4.4). Although it has been demonstrated that the image of the pupil is not a perfect ellipse, this method uses ellipse fits to define a cone through the nodal point of the camera and the virtual pupil image. By intersecting both cones the orientation and location of the virtual pupil was estimated as previously described by. The orientation of the virtual pupil is the normal vector \( n \) of the virtual pupil surface, and Equation 4.3 was used to calculate the pan and tilt angles.

\[
\begin{bmatrix}
  x_n \\
  y_n \\
  z_n
\end{bmatrix} =
\begin{bmatrix}
  \cos \phi \sin \theta \\
  \sin \phi \\
  -\cos \phi \cos \theta
\end{bmatrix}
\]  

(4.3)

The results of the VP-CC method and the conic algebra method were compared. Moreover, to test the assumption that the distance between the entrance pupil and the estimated center of corneal curvature remains stable, we calculated this distance for the different locations and orientations of the eye as well as for each pupil diameter.

Figure 4.4. Gaze reconstruction using conic algebra. For each camera a cone can be projected through the projection of the virtual pupil at the image plane. The virtual pupil lies at the intersection of these cones.
Results

Figure 4.5 illustrates the 3D shape, orientation and location of the virtual pupil derived from the ray-tracing model for two different gaze directions. As expected from previous work, the virtual pupil (blue) lies in front of the actual pupil (black), and is larger than the actual pupil. Additionally, the 3D shape of the virtual pupil is slightly distorted compared with the actual pupil, its orientation (red arrow) is tilted compared to the actual pupil and its geometric center (VP) deviates slightly from the optical axis of the eye (dashed gray line). These effects are larger for the horizontal rotation (Figure 4.5A) than for the vertical rotation (Figure 4.5B) due to the geometry of the simulated set-up.

Gaze reconstructions obtained with the VP-CC vector method are illustrated as well. The accuracy of this method is affected by the concentric gradient in refractive power of the cornea, which influences the estimated location of the center of the virtual pupil, as well as the aspherical shape of the anterior cornea, which influences the estimated location of the center of corneal curvature (Methods). In this example both points lie almost on the optical axis, resulting in a small gaze error. However, in general, both points can deviate significantly from the optical axis, resulting in larger gaze errors.

![Figure 4.5. Location, orientation and shape of the virtual pupil obtained through ray tracing compared to the actual pupil for two different cases. A. The actual pupil is rotated 15 degrees to the right. Top view of the simulation. B. The actual pupil is rotated 15 degrees up. Side view of the simulation. The optical axis (OA) always goes through the fixed center of corneal curvature (fCC) and the center of the actual pupil (AP). The 3D shape of the virtual pupil boundary varied between conditions. A plane was fit through the boundary points of the virtual pupil (VPplane). The normal vector of this plane (VPnorm) was used to describe the orientation of the virtual pupil. The VP-CC vector is defined by the estimated center of corneal curvature (CC) and the center of the virtual pupil (VP).](image-url)
Figure 4.6 shows the orientation of the virtual pupil as obtained through ray-tracing (black crosses) and the results of the different gaze reconstruction methods (cyan, blue and red dots) for two different head positions, and for a series of different orientations of the actual pupil (open circles) with a diameter of 4 mm. Note, that the orientation of the virtual pupil, which was determined from a plane fitted through the virtual pupil boundary points (Figure 4.5), can deviate significantly from the orientation of the actual pupil. This difference depends systematically on the orientation of the actual pupil as well as on the translation of the head. If the pupil is located at the central position between both cameras (X = 0, Y = 0, Z = 400 mm, Figure 4.6A), a symmetrical pattern is observed with increasing differences as the gaze angles increase in the horizontal and vertical direction. The discrepancies range up to 2.88 degrees for horizontal gaze angles of 15 degrees (~19%), and up to 0.96 degrees (~6.5%) for vertical gaze angles of 15 degrees. When the actual pupil is translated 6 cm to the right and 6 cm down (Figure 4.6B) the pattern changes, with the smallest errors now occurring for a horizontal gaze angle of ~10 degrees to the right, and a vertical gaze angle of ~10 degrees down. The maximum differences increase to 4.47 degrees horizontally and 1.61 degrees vertically.

Figure 4.6. Reconstruction of gaze direction with stereo eye-tracking methods compared to the orientation of the actual and virtual pupil. A. The actual pupil was positioned at the central location between both cameras (X = 0mm, Y = 0mm, Z = 400mm). B. The actual pupil was shifted 6 cm to the left and 6 cm downwards, simulating a translation of the head.
Gaze reconstructions obtained with the conic algebra method correspond closely to the orientation of the virtual pupil (Figure 4.6, red circles). Within the range of simulated gaze angles and head translations for a pupil diameter of 4 mm, the maximum difference between the orientation of the virtual pupil and the conic algebra method was only 0.12 degrees. The VP-fCC method resulted in significantly smaller errors (Figure 4.6, cyan circles). For this method, the maximum horizontal and vertical errors were 0.08 and 0.09 degrees, respectively, when the pupil was located at the central position, and 0.12 and 0.18 degrees when the head was translated 6 cm to the left and 6 cm downwards. Note that these errors were only due to shifts in the center of the virtual pupil away from the eye’s optical axis, because the VP-fCC method assumes that one would be able to obtain error-free estimates of the center of corneal curvature (Methods). In practice, however, the center of corneal curvature must be estimated from the corneal reflections of the IR light sources. The VP-CC method includes this estimation procedure (Methods). Errors obtained with the VP-CC method (Figure 4.6, dark blue circles) increased compared with the theoretical VP-fCC method, but were still small compared with the conic algebra method. In the central position the VP-CC method resulted in maximum errors of 0.22 degrees horizontally and 0.47 degrees vertically. When the actual pupil was translated 6 cm up and to the right these errors increased, ranging up to 1.16 degrees horizontally and 1.3 degrees vertically for 15 degrees pupil rotations.

Figure 4.7 quantifies the systematic effect of pupil orientation and (head) translation on the gaze reconstruction errors for the different reconstruction methods. Each curve shows the results obtained with the eye model placed at a given location in the fronto-parallel plane. Left-hand panels plot the horizontal gaze errors as a function of horizontal pupil orientation, with data averaged across the nine different vertical pupil orientations. Error bars denote the range of horizontal gaze errors across those vertical pupil orientations. Right-hand panels plot vertical gaze errors as a function of vertical pupil orientation, with data averaged across the nine different horizontal pupil orientations. Error bars denote the range of vertical gaze errors across those horizontal pupil orientations. Colors identify the magnitude of the horizontal or vertical translation from the central position. Note that a 3 cm horizontal translation from the median plane corresponds closely to the natural location of the human eye. For symmetry reasons, we only translated the eye model in 1 quadrant of the fronto-parallel plane (down and to the left).
For each of the nine simulated positions, gaze errors depended systematically on the orientation of the actual pupil. In general, the average horizontal and vertical gaze errors increased monotonically with increasing rotation of the actual pupil, but note that cross-talk between the horizontal and vertical components was quite limited. The latter can be inferred from the size of the error bars. Horizontal and vertical translations also had a significant influence on the estimated gaze direction, but note that the errors in the horizontal and vertical direction of gaze were remarkably invariant to vertical and horizontal head translations, respectively. Especially for the conic algebra method, this type of cross-talk is quite small, which is why the three curves of a given color overlap. In fact, gaze errors observed for the conical algebra method show a remarkably linear pattern; component errors increase at a fixed gain with the corresponding gaze angle and head translations add a fixed, position-dependent bias to these errors. For the simulated set-up, horizontal errors were approximately 3 times larger than vertical errors.

For the VP-fCC method the gaze reconstruction errors also showed a near-linear dependency on translations and rotations of the actual pupil, but the resulting errors were in the opposite direction and much smaller. More specifically, the influence of horizontal pupil orientation was nearly 80 times smaller, and the effect of horizontal pupil translation was about 25 times smaller compared to the conic algebra method. This difference was smaller for errors in the vertical component, but still significant. There is also a slight position-dependent cross-talk between the horizontal and vertical components (see also Figure 4.5).

For the VP-CC method the relationship between translation and rotation of the actual pupil and the error in estimated gaze direction clearly deviated from the other two methods. It showed smooth nonlinearities causing the average component error to increase with an increasing gain as a function of the actual pupil orientation. Cross-talk also increased with increasing pupil rotation and both effects were amplified by translations of the pupil away from the central position. Compared to the VP-fCC method the errors were on average between five and ten times larger. This shows that the aspherical shape of the anterior cornea has a significant impact on mislocalization of the center of corneal curvature. The largest errors (up to 1.5 degrees) occurred for 15 degrees rotations combined with a 6 cm translation.
Figure 4.7. Accuracy of the different gaze reconstruction methods for a fixed pupil diameter of 4 mm. A. the conic algebra method. B. the VP-fCC method. C. the VP-CC method. Each plot shows the horizontal/vertical gaze reconstruction errors as a function of the horizontal/vertical orientation of the actual pupil for each of the nine different pupil/head translations. Data are averaged either across the nine different vertical pupil orientations (left-hand plots) or the nine different horizontal pupil orientations (right-hand plots). Colors identify the magnitude of the horizontal (left-hand plots) or vertical (right-hand plots) translation. Note the scaling differences between plots A, B and C.
The effect of pupil diameter on the different gaze reconstruction methods is presented in Figure 4.8. The results reveal a strong effect of pupil size on the conic algebra method. The gaze error decreases as pupil size increases for all head positions. The maximum difference in gaze error for a pupil with a diameter of one millimeter and

![Graphs showing the effect of pupil size on stereo eye-tracking methods. A. the conic algebra method. B. the VP-fCC method. C. the VP-CC method. The mean and range of the gaze errors is plotted for all pupil diameters at two head positions. The dashed lines indicate the results for a pupil diameter of 4 mm. For clarity of the graph, the actual pupil was only translated horizontally (left-hand plots) or vertically (right-hand plots). Note the scaling differences between the horizontal and vertical errors.](image-url)
a pupil with a diameter of six millimeter reaches 1.3 degrees. The effect of pupil size on the VP-fCC and VP-CC methods is smaller with maximum differences between the gaze error of the smallest and largest pupil diameter of 0.14 and 0.18 degrees for the two methods respectively. Note that in contrast to the conic algebra method, the gaze error increases as pupil size increases for the VP-fCC and VP-CC method.

In Figure 4.9 we show the results of calculations that tested how the distance between the center of corneal curvature and the center of the virtual pupil depended on pupil diameter, head position, and orientation of the virtual pupil. The distance increased for larger pupil sizes and larger gaze angles if the actual pupil is located at the central location (Figure 4.9A). Both the horizontal and vertical orientation of the actual pupil have an influence, although the horizontal angle had a stronger effect. Additionally, the position of the actual pupil changed this pattern (Figure 4.9B) into an asymmetrical function.

Figure 4.9. The influence of size, orientation and position of the actual pupil on the distance between the estimation of the center of corneal curvature and the center of the virtual pupil. The color coding in each square represents the distance between VP and CC in mm for one simulated gaze orientation. The left panels show the results for a pupil size of 1 mm and the right panels the results for a pupil size of 6 mm. A. The actual pupil was located at the central position. B. The actual pupil was shifted 6 cm to the left, and 6 cm down.
Discussion

The results of the ray tracing model demonstrate that the reconstructed entrance pupil based on the images from two (synchronized) cameras does not have the same orientation as the actual pupil. The position of the eye and the gaze angle both influence the difference in orientation. In general, larger gaze angles result in bigger differences between the orientation of actual pupil and the virtual pupil. This is in line with previous research on the orientation of the entrance pupil\(^{208}\) and demonstrates that the assumption that the virtual pupil is positioned in parallel with the actual pupil is not correct. Additionally, in contrast to previous arguments that stereo eye-tracking methods are invariant to head movements (cf.\(^{119}\)), we found significant effects of head translations on the accuracy of these methods.

Furthermore, the results of gaze reconstructions using two typical stereo eye-tracking methods revealed that the results from both methods are affected by using a more accurate eye model. However, the magnitude of the errors differed substantially between the two methods. Reconstructions based on conic algebra, which use the entire pupil boundary, resulted in relatively large errors. By comparing the conic algebra method with the orientation of the virtual pupil, we demonstrated that the former method essentially reconstructs the orientation of the entrance pupil, rather than the actual pupil. This is a logical consequence of using the images of the entire pupil: both cameras observe a different virtual pupil, and conic algebra assumes that both cameras are looking at the same object. In addition, because the image of the virtual pupil is not a perfect ellipse\(^{208}\), the use of fitted ellipses might have contributed to larger gaze errors too. The horizontal errors could exceed 4 degrees. The vertical errors were smaller, which can be explained by the location and orientation of the simulated cameras. They were positioned at the same height, while only their horizontal positions differed. As a result, the virtual pupil observed by either camera differed especially in the horizontal direction, and less in the vertical direction. The large errors of the conic algebra method indicate that this method cannot be readily applied for stereo eye-trackers. However, as the errors appeared to be systematically related to the direction of gaze and to head position, it might be possible to correct for these errors. Note, however, that such corrections will be complicated by the influence of pupil diameter. Further research is necessary to investigate the efficacy of possible methods to correct for these errors.

On the other hand, the gaze errors resulting from the VP-CC method were smaller, remaining within about 1.3 degrees. The largest gaze errors were found for
large translations and large rotations (15 degrees) of the actual pupil. This effect is mainly caused by the error in estimating the center of corneal curvature. The gaze errors of the VP-fCC method, where only the center of the virtual pupil is estimated, were on average between five and ten times smaller, with a maximum of only 0.2 degrees within the applied range of eye locations, gaze angles and pupil sizes. The difference between the VP-CC method and the VP-fCC method are caused by the error in estimation of the center of corneal curvature. Additionally, the gaze error is in the opposite direction for the VP-fCC method. With this method the gaze direction is overestimated, whereas the other methods underestimate the gaze direction. The error in the estimation of the center of corneal curvature is caused by the aspherical anterior cornea. When the eye is translated and rotated, the rays from the light sources are reflected further away from the apex of the cornea. Especially the normal vectors at these reflection points will then intersect further away from the center of the cornea. This leads to larger errors in the estimation of the center of curvature and therefore to larger gaze errors. The gaze errors found in our simulations are similar to the errors found in different prototypes that have been developed. Most systems have an average accuracy of ~1 degree\textsuperscript{117,120,123}.

Previous studies\textsuperscript{119,120} assumed that the distance between the center of corneal curvature and the center of the virtual pupil remains constant. This was based on simple eye models with only one spherical refractive surface of the cornea. We used a more realistic eye model and demonstrate that this distance depends on the orientation, position and diameter of the actual pupil (Figure 4.9). As described above, errors in the estimation of the center of curvature and in the estimation of the virtual pupil resulted from the assumptions regarding the shape of the cornea. As a result of different optical geometries, the distance between these points varies. Chen et al. proposed to reduce the noise during stereo eye-tracking by correcting the distance between the center of corneal curvature and the center of the virtual pupil to a fixed distance\textsuperscript{120}. However, we demonstrated that this assumption is not valid. The effect of such a correction on the accuracy of these methods has to be determined. It might result in additional errors when the pupil size changes. Furthermore, the effect of pupil diameter on the gaze reconstructions and the distance between the center of curvature and the center of the virtual pupil is in line with findings in eye-tracking methods that use one camera, where pupil size systematically influences the outcome of eye-tracking\textsuperscript{209}. Changes in pupil size are common during experiments, for example due to differences in cognitive load, in
arousal or to differences in luminance on the screen.

The set-up used in this article is a realistic set-up for remote stereo eye-trackers. However, as the spatial location of the eye influenced the results, the results may differ in set-ups where the distance between the eye and the cameras is different. Similarly, the distance between the cameras and orientation of the cameras relative to each other may influence the results, as well as the location of the IR light sources. The different results for the nine spatial locations of the eye indicate that the orientation of the eye relative to the cameras, i.e., a combination of the gaze orientation and the angle between the optical axis of the camera and the position of the eye, cause differences in the orientation of the virtual pupil. Additionally, the position of the IR light sources influences the estimation of the center of corneal curvature. Further research is necessary to investigate whether the set-up can be optimized.

Conclusion

We demonstrated that the shape of the cornea has a significant influence on the accuracy of stereo eye-tracking methods. Pupil size, gaze direction and head position all influence the accuracy of eye-tracking methods. The gaze reconstruction that uses the center of the pupil and reflections of IR light sources is more accurate than the conic algebra method, which uses the entire pupil. However, the gaze errors of the conic algebra method appear to be systematic, and therefore a correction could result in more accurate gaze reconstructions. In conclusion, stereo eye-tracking methods that assume a spherical cornea with one refractive surface can be an option in situations where reliable calibration is not possible. However, more accurate measurements require the use of a more elaborate model of the eye geometry in which the optics of the cornea are better taken into account.
Optics of the human cornea and the accuracy of stereo eye-tracking
Development and validation of a high-speed stereoscopic eye tracker

Published as: Barsingerhorn AD, Boonstra FN & Goossens J. Development and validation of a high-speed stereoscopic eye tracker. *Behavior Research Methods*. 2018; Epub ahead of print.
Introduction

Standard video-based eye trackers rely on an individual calibration procedure in which participants are asked to fixate multiple small targets at known locations. This calibration is necessary to convert the image features of the eyes into estimates of the point of gaze (POG) on the screen. Most healthy adults are able to perform such a calibration procedure effortlessly. However, if participants are unable to reliably fixate small stimuli, for example in the case of oculomotor deficits, low vision, or reduced attention, accurate calibration is not possible. As a result, no reliable gaze data can be obtained with standard video-based eye trackers in these participants.

The scleral search coil method can be an alternative for some of these participants. This technique measures electromagnetic induction in a copper coil which is embedded in a contact lens. Each search coil can be accurately calibrated before use so that only one in vivo calibration point is needed to determine the orientation of the coil with respect to the visual axis of the subject’s eye. This method has very high temporal and spatial resolution allowing even the smaller types of eye movements to be studied. However, a major disadvantage is its somewhat invasive and uncomfortable nature. The cornea needs to be anesthetized and the recording time is typically limited to about 30-45 min (but see also). These disadvantages makes it an unsuitable technique for most clinical settings and for eye movement recordings in (young) children. Therefore, researchers have developed non-invasive video-based eye tracking techniques with two cameras that rely on a simplified calibration procedure. With these stereo eye-tracking methods one can estimate the position of the eye and the orientation of its optical axis from the stereo images and the known geometry of the setup. Only the deviation between the optical and visual axis needs to be determined, which can be done through a one point calibration procedure. Note that even without this in vivo calibration, the changes in orientation of the optical axis still accurately reflect changes in eye orientation; they only display a fixed offset with respect to the gaze angles. The spatial accuracy of the proposed prototypes of stereo eye trackers is typically around 1 degree, or even below 1 degree. This is sufficient for a range of eye-tracking applications. However, the sampling rate of those systems range between 20 and 30 Hz, which is not enough to analyze the kinematics of rapid eye movements. Since the temporal resolution of eye trackers affects estimates of saccade peak velocities and other kinematic parameters, it has been recommended to use sampling rates of at least 200 Hz, 250 Hz, or even 300 Hz. Therefore, our aim was to develop a stereo
eye tracker with a sampling rate of at least 250 Hz.

The accuracy of eye trackers is typically evaluated by measuring the deviation between the location of the visual target and the reconstructed point of gaze. However, it is known that participants do not always look exactly at the center of the target. In addition, individual characteristics of participants, such as iris color and the physiology and anatomy of the eye can lower the accuracy of video-based eye trackers. Therefore, the second aim of the current study was to validate the accuracy of our stereoscopic eye tracker against the accuracy of an established high-speed remote eye tracking system, the Eyelink 1000 plus, by recording simultaneously with the two systems.

Previous studies have shown that the accuracy of standard remote eye trackers deteriorate if the head moves. In addition, recent simulations have revealed that head movements could also reduce the accuracy of stereo eye trackers. Therefore, we assessed the robustness against head movements for the two systems by testing the accuracy of the Eyelink 1000 plus and the stereoscopic eye tracker at nine different head positions.

Methods

Stereo eye-tracker hardware

The stereo eye-tracker is shown in Figure 5.1. It consisted of two USB 3.0 cameras (LumeneraLt225 NIR, Lumenera Corporation, Ottawa, Canada, pixel size 5.5 × 5.5 µm) and two 850 nm infrared illuminators (Abus TV6700, ABUS KG, Wetter, Germany) mounted on an optic rail. The cameras were positioned ~12 cm apart with their optical axes directed towards the location of the subjects’ eyes. The first illuminator was placed ~6 cm to the right of camera 1, the right one in Figure 5.1, and the second illuminator was placed ~6 cm to the left of camera 2.

The temporal resolution of the cameras depended on their spatial resolution settings due to the limited amount of data that can be transferred over a USB 3.0 connection. At full spatial resolution (2044 × 1088 pixels) the cameras could film at ~180 fps while the cameras reached ~380 fps at 1048 x 480 pixels. For the present eye tracking application we selected the latter option. The lenses with manual focus and diaphragm had a focal length of 16mm (Navitar NMV-16M23, Navitar Inc, Rochester, NY, USA). Infrared-passing filters (UV/Vis-Cut R-72; Edmund Optics Inc, Barrington, NJ, USA) which pass wavelengths > 720 nm were added on the lenses to block light in the visible spectrum. The eye tracking software was executed on a
laptop (Dell M3800; Dell Inc., Round Rock, TX, USA) equipped with eight 2.3Ghz central processing units (Intel core i7-472HQ, Santa Clara, CA, USA), an OpenGL graphics card (Nvidia Quadro K1100M; Santa Clara, CA, USA), and a 64 bit Windows 7 Professional operating system (Service Pack 1, Microsoft Corporation, Redmond, WA, USA). Each camera was connected to a separate USB 3.0 bus to achieve high camera frame rates.

**System calibration**

The system was calibrated in two steps. First, the internal and external camera parameters were obtained through a stereo camera calibration procedure in which images of a calibrated checkerboard pattern were taken while it was moved around the cameras at various angles (Matlab computer vision toolbox, Matlab R2016b, MathWorks, Inc., Natick, MA, USA). This calibration procedure allowed us to correct for lens distortions (using the function undistortPoints from the MATLAB computer vision toolbox) and express the image coordinates in both cameras in a common world-centered frame of reference using a right-handed Cartesian coordinate system with its origin located at the nodal point of camera 1. The x- and y-axis of the tracker’s coordinate system were parallel to the image plane of camera 1, and the positive z-axis pointed away from it (towards the subject). In the second step, the position of the illuminators and the position and orientation of the screen were determined.
The illuminators and the screen were not directly observed by the cameras, as they were located behind the camera system. Therefore, a planar mirror with a dot stimulus pattern attached to its surface was used to observe the virtual images of the illuminators and the screen, as suggested in several studies\footnote{120,122,123,219}. To reconstruct the position and orientation of the mirror, we determined the 3D location of the markers from their image coordinates in each of the two cameras by means of triangulation. Both cameras also observed virtual images of the illuminators and the computer screen in the mirror. Thus, also the 3D locations of the virtual images of the illuminators and the screen could be obtained through triangulation. The 3D positions of the illuminators were then determined from the 3D locations of their virtual images behind the mirror and the 3D position and orientation of the mirror itself. In a similar way the 3D position and orientation of the screen were determined from the virtual image of a dot stimulus pattern that was displayed on the screen during the procedure. We performed all triangulations with the function \texttt{triangulate} from the computer vision toolbox of Matlab, as it automatically converts the image coordinates into world-centered system coordinates using the parameters from the stereo camera calibration.

**Stereo eye tracker software**

We developed the eye tracking software in Visual Studio 2012 with C# as the programming language. The LuCam SDK V6.3 (Lumenera Corporation, Ottawa, Canada) camera driver was used to set the acquisition parameters (shutter time, resolution, etc) and capture the images from the two cameras. The EmguCV library (OpenCV\footnote{220} wrappers for C#) was used for online image processing. The stereo eye tracking software, as well as the offline gaze reconstruction algorithms described below, are available at \url{https://github.com/Donders-Institute/Stereo-gaze-tracking}.

The eye tracking program supports tracking of the pupil and the corneal reflections (glints) of both eyes for two cameras. It can run simultaneously with (custom) stimulus presentation software on the same computer, which eliminates the need for a separate eye tracking computer. The code for detecting, segmenting and tracking the pupil and glints was adapted from the open source software of the ITU Gaze Tracker\footnote{221} which was designed for single-camera setups. The acquisition and analysis of the images from the two cameras run in separate threads. A simplified flow chart of the processing within each thread is shown in Figure 5.2. First, the eyes are detected by a Haar cascade classifier\footnote{222} which is part of the EmguCV library. Once the eyes
are found, the pixel coordinates of the pupil centers and the glints are estimated. To extract the pupil, the image of the eye is segmented by thresholding the image. The center of the pupil is then estimated at subpixel resolution by fitting an ellipse (least squares method) to the pixel coordinates of the pupil contour. Subsequently, the glints are detected with a different intensity threshold. The centroids of the glints provide an estimation of the center of the glints at subpixel resolution. The center of the pupil is used to update the eye positions to allow for robust tracking in the presence of head movements. If the software fails to detect a pupil or glints, the eye positions might have changed due to head movements, or the participant could have blinked. Here, we had to balance between robustness and speed of the eye tracking. Detecting the eyes is relatively slow (~16 ms) because of the computational load. Therefore, redetecting the eyes too quickly after a failure to detect a pupil or glints is inefficient. The participant could have blinked, in which case the tracking could continue immediately after the blink without redetection necessary. However,

![Flow chart of the image processing for each camera. First the eyes are detected, after which the pixel coordinates of the pupil and glints are extracted. A reset counter is used to force the software to redetect the eyes only if no pupil or glints are detected in 30 consecutive frames. If only one eye is found, the software continues tracking that eye for 30 frames. After 30 frames the eyes are redetected to continue binocular tracking.](image.png)
Chapter 5

if the software waits too long to redetect the eyes, data might be lost if the eyes truly shifted due to head movements. Therefore, the software redetects the eyes only if no pupil or glints are found in 30 consecutive frames. The timestamps and the pixel coordinates of the pupils and glints are saved in a data file.

Figure 5.3 shows a screenshot of the online display. Because both eyes are filmed by two cameras there are four eye images visible. The result of image thresholding for the pupil detection is shown in blue, and in red for the glint detection. The estimated centers of the pupils are shown with white crosshairs. The estimated centers of the glints are shown with yellow crosshairs. The thresholds for the image segmentation can be adjusted separately for each eye in each camera, while the upper and lower bounds for pupil size and glint size are set to common values. For most participants (~2/3) the default values shown in Figure 5.3 yield accurate results. Storage of the tracking data can be paused to prevent unnecessarily large data files. In addition,
online feedback is given by showing the horizontal and vertical components of the pupil-glint vectors as a function of time for both cameras over a period of ~3 s. The pupil-glint vectors are used in most standard eye trackers to calculate the point of gaze using a mapping obtained through the calibration procedure\textsuperscript{88}. In our online display it does not reflect the exact point of gaze or the exact gaze direction, but it gives a clear indication about the quality and noise level of the data. For example, the occurrence of a number of larger and smaller saccades is clearly visible in these uncalibrated signals.

In addition, a histogram equalizer (EmguCV) can be used to improve the image segmentation and a Gaussian filter can be applied to reduce the effect of pixel noise on the estimation of the pixel coordinates of the pupils and glints. For some applications monocular tracking may be preferred. Therefore, the stereo tracker can be set in a monocular tracking modus. Furthermore, it is possible to adjust which part of the camera sensor is used to make sure the participant is well within the selected field of view.

We have included two videos of the online display in Supplement 5.1 and 5.2 to illustrate the tracking stability of the system (videos are available at: https://doi.org/10.3758/s13428-018-1026-7). Both subjects (first and last author) were head free and looked at the online display window, which was placed on the stimulus screen for illustration purposes. The images that are visible on the laptop screen are the unprocessed images from the two cameras.

**Gaze reconstruction**

The reconstruction of gaze was performed offline in Matlab R2016b. Multiple researchers have demonstrated that with two cameras and two light sources at known locations, the 3D location of the eye as well as the 2D orientation of the optical axis of the eye can be approximated by estimating the center of the virtual pupil ($p_v$) and the center of corneal curvature ($c$)\textsuperscript{119,120,122,123}. Because image acquisition in the two cameras ran asynchronous, for both cameras the pixel coordinates of the pupil and glints were first interpolated at the timestamps at which the other camera captured an image of the eyes.

Due to refraction of light rays at the corneal surface the cameras do no observe the actual pupil (Figure 5.4A). They only see a virtual image of the pupil. Different approaches have been used to estimate the center of the virtual pupil. Some of these methods correct for refraction at the corneal surface\textsuperscript{117,121} assuming a spherical
shape of the cornea, while others do not correct for the refraction\textsuperscript{119,120,122}. We decided to not correct for refraction at the corneal surface, because it has been shown that, even if a more realistic model of the cornea is used, the location of the center of the virtual pupil remains within 0.2 deg of the optical axis\textsuperscript{218}. Therefore, the center of the virtual pupil was triangulated from the image points $v_1$ and $v_2$ of the virtual pupil in the two cameras and the nodal points $o_1$ and $o_2$ of the cameras (Figure 5.4A) by computing $p_v$ from the least squares solution of:

$$\begin{cases} p_v = o_1 + \lambda_1 (v_1 - o_1) \\ p_v = o_2 + \lambda_2 (v_2 - o_2) \end{cases}$$  \hspace{1cm} (5.1)$$

Different approaches have been used to obtain the center of corneal curvature\textsuperscript{119,120,122,123}, all assuming that the cornea acts as a spherical mirror during the process of glint formation. We used the approach suggested by Zhu and Ji\textsuperscript{122}. Based on the reflection law of convex mirrors, this method assumes that each camera observes a virtual image of the IR illuminator at the same location regardless of the camera’s location. The position of the virtual image of the illuminator is then only determined by the actual position of the illuminator and by the location of the eye.

Figure 5.4. Ray-tracing diagrams (not to scale) showing a schematic top view of the two cameras, the two infrared illuminators, and one eye. A. The center of the pupil projections onto each camera ($v_1$ and $v_2$) and the nodal points of these cameras ($o_1$ and $o_2$) were used to triangulate the center of the virtual pupil $p_v$. B. For each illuminator $L_j$, causing glint $u_{1j}$ in camera 1 and glint $u_{2j}$ in camera 2, its virtual image $L'_j$ was obtained through triangulation (yellow lines). Subsequently, the center of corneal curvature $c$ was estimated by intersecting the line through illuminator $L_1$ and its virtual image $L'_1$ with the line through illuminator $L_2$ and its virtual image $L'_2$. 

110
With two cameras the 3D location of the virtual image of an illuminator can be triangulated from the corresponding glint coordinates in the two cameras (Figure 5.4B). I.e., for each illuminator \( L_j \), causing glint \( u_{j1} \) in camera 1 and glint \( u_{j2} \) in camera 2, its virtual image \( L'_j \) was obtained through:

\[
\begin{align*}
L'_j &= o_1 + \lambda_1 (u_{j1} - o_1) \\
L'_j &= o_2 + \lambda_2 (u_{j2} - o_2)
\end{align*}
\] (5.2)

Subsequently, the center of corneal curvature \( c \) was estimated by intersecting the line through illuminator \( L_1 \) and its virtual image \( L'_1 \) with the line through illuminator \( L_2 \) and its virtual image \( L'_2 \):

\[
\begin{align*}
c &= L_1 + \lambda_1 (L_1 - L'_1) \\
c &= L_2 + \lambda_2 (L_2 - L'_2)
\end{align*}
\] (5.3)

In practice, the 3D reconstructions using triangulation are not perfect. Due to image noise and an aspherical cornea\(^{120,124,218} \), often the lines do not intersect. Therefore, the position at which the distance between the lines is minimal is taken as the 3D position. This introduces errors in the estimation of \( c \) and \( p_v \). Chen et al. proposed a method to reduce the variability resulting from image noise by assuming that for a given subject the distance \( K \) between \( c \) and \( p_v \) remains constant\(^{120} \). Because image noise mainly affects the localization along the z-axis of the tracker’s coordinate system (i.e., in the cameras viewing direction), the x and y coordinates of the virtual pupil are kept fixed, while its z-coordinate is changed to satisfy the constraint that \( K \) is fixed. Given the estimated center of corneal curvature \( c = (x_c, y_c, z_c) \) and the virtual pupil center \( p_v = (x_{pv}, y_{pv}, z_{pv}) \), the z-coordinate of \( p'_{pv} \) is computed as:

\[
z'_{pv} = z_c - \sqrt{K^2 - (x_c - x_{pv})^2 - (y_c - y_{pv})^2}
\] (5.4)

where \( K \) is a subject/eye-specific value.
In the next step, the $x$, $y$, and $z$ coordinates of $p'_v$ and $c$ in the tracker’s coordinate system were mapped onto $X$, $Y$ and $Z$ coordinates of a right-handed Cartesian coordinate system whose XY-plane was coincident with the orientation of the screen and with its origin located at the center of the screen. In this stimulus coordinate system the X-axis was horizontal, the Y-axis vertical and the positive Z-axis came out of the screen (towards the subject). This coordinate transformation involved translations, $T$, and rotations, $R$, determined from the system calibration procedure (see above):

$$
\begin{align*}
\mathbf{P}'_v &= T(R \mathbf{p}'_v)
\mathbf{C} &= T(R \mathbf{c})
\end{align*}
$$

(5.5)

Following Guestrin and Eizenman\textsuperscript{119}, the orientation of the optical axis of the eye was then described by the horizontal (pan) angle $\theta_{\text{eye}}$ and the vertical (tilt) angle $\phi_{\text{eye}}$ where the origin of the coordinate system is translated to the center of corneal curvature. The angles $\theta_{\text{eye}}$ and $\phi_{\text{eye}}$ were obtained from $\mathbf{C}$ and $\mathbf{P}'_v$ as follows:

$$
\begin{align*}
\mathbf{P}'_v - \mathbf{C} &= \begin{bmatrix}
\cos \phi_{\text{eye}} \sin \theta_{\text{eye}} \\
\sin \phi_{\text{eye}} \\
-\cos \phi_{\text{eye}} \cos \theta_{\text{eye}}
\end{bmatrix}
\end{align*}
$$

(5.6)

The orientation of the visual axis, which defines the direction of gaze, was then estimated from the orientation of the optical axis and the deviation between the optical axis and visual axis using:

$$
\hat{g} = \begin{bmatrix}
\cos(\phi_{\text{eye}} + \beta_{\text{eye}}) \sin(\theta_{\text{eye}} + \alpha_{\text{eye}}) \\
\sin(\phi_{\text{eye}} + \beta_{\text{eye}}) \\
-\cos(\phi_{\text{eye}} + \beta_{\text{eye}}) \cos(\theta_{\text{eye}} + \alpha_{\text{eye}})
\end{bmatrix}
$$

(5.7)

were $\alpha_{\text{eye}}$ and $\beta_{\text{eye}}$ are the horizontal and vertical angles between the visual axis and the optical axis, respectively. These angles $\alpha_{\text{eye}}$ and $\beta_{\text{eye}}$ were estimated from the horizontal and vertical angles of the optical axis with respect to the line of sight during a single point calibration procedure which involved fixation of a small target at the center of the screen (see below). The median length of the $\mathbf{c} - \mathbf{p}_v$ vector
measured during this procedure provided the subject-specific value of $K$ used in Equation 5.4.

Since the visual axis goes through the center of corneal curvature $C$, the point of gaze ($POG$) on the screen can be estimated from the following parametric equation:

$$POG = C + \lambda_g \hat{g}$$

where $\lambda_g$ is the distance from $C$ at which the line of sight intersects the screen. Given the estimated center of corneal curvature $C = (X_c, Y_c, Z_c)$ and because the screen is a planar scene at $Z = 0$, the value of $\lambda_g$ is given by:

$$\lambda_g = -\frac{Z_c}{-\cos(\phi_{eye} + \beta_{eye})\cos(\theta_{eye} + \alpha_{eye})}$$

This reconstruction of $POG$ does not account for the kinematics of the eyeball, but see\textsuperscript{117,223} for an alternative procedure which incorporates Listing’s law under the assumption that Listing’s plane is parallel to the XY-plane.

The correction for the effect of image noise on the z-component of the 3D reconstructions (Equation 5.4) worked well to reduce the noise level up to ~50%. However, it turned out that it introduced systematic errors in the accuracy of the stereo tracker by systematically under- or overestimating the gaze angles up to ~15% if the actual length of the $c-p_v$ vector differed substantially from $K$. This is in line with recent simulations of stereo eye tracking methods, which indicated that the length of the $c-p_v$ vector varies systematically as a function of head position, pupil size and gaze angles due to asphericity of the cornea\textsuperscript{218}. We first attempted to account for these variations in the length the $c-p_v$ vector by replacing $K$ in Equation 5.4 with a low-pass filtered measure of the actual length of the $c-p_v$ vector, $K_{actual}$. While this approach did reduce the systematic errors to some extent, it could not adequately account for translations of the head. From this we inferred that the systematic under- or overestimation of the gaze angles is not only due to variations in length of the $c-p_v$ vector. Indeed, the consequences of the aspherical properties of the cornea are complex. Simulations with an aspherical model of the cornea
suggest that the systematic variation in length of the $\mathbf{c} - \mathbf{p}_v$ vector is caused by shifts of the virtual pupil and a mislocalization of the “center of corneal curvature” (through triangulation or other proposed methods) because an aspherical cornea does not have a unique center of curvature\(^{218}\). The consequence of the shifts of the virtual pupil is probably small because they are primarily along the optical axis of the eye, but the mislocalization of the “center of corneal curvature” introduces more severe errors because it can put the estimate of $\mathbf{c}$ off the optical axis in a way that varies systematically with the position and orientation of the eye with respect to the cameras. From this, we inferred that corrections for systematic errors in the reconstruction of the POG would have to include the 3D position of the eye as well. Empirically, we found that the reconstruction errors of the POG can be attenuated significantly with a variable gain factor that reflects the difference between $K$ and the actual length of the $\mathbf{c} - \mathbf{p}_v$ vector and that acts on both the eye position term and the gaze orientation term of Equation 5.8 in the following way:

$$\hat{\mathbf{g}}_{\text{corrected}} = \frac{K}{K_{\text{actual}}} \left( \mathbf{C} + \lambda_\mathbf{g} \hat{\mathbf{g}} \right)$$ \hspace{1cm} (5.10)

where $K_{\text{actual}}$ is a filtered measure of the actual length of the $\mathbf{c} - \mathbf{p}_v$ vector. A median filter with a width of 20 samples (using the function medfilt1, Matlab 2016b) was used. From the corrected POG, we obtained the corrected direction of gaze:

$$\hat{\mathbf{g}}_{\text{corrected}} = \frac{\mathbf{POG}_{\text{corrected}} - \mathbf{C}}{\mathbf{POG}_{\text{corrected}} - \mathbf{C}}$$ \hspace{1cm} (5.11)

and subsequently, the corrected gaze angles $\theta_{\text{eye corrected}}$ and $\phi_{\text{eye corrected}}$ were obtained through:

$$\hat{\mathbf{g}}_{\text{corrected}} = \begin{bmatrix} \cos \phi_{\text{eye corrected}} \\ \sin \phi_{\text{eye corrected}} \\ -\cos \theta_{\text{eye corrected}} \cos \phi_{\text{eye corrected}} \\ \sin \theta_{\text{eye corrected}} \end{bmatrix}$$ \hspace{1cm} (5.12)

The proposed corrections also worked when the head was translated (see results).
Note that Eqs. 8-11 perform the corrections in stimulus coordinates. It may be possible to apply the correction in the tracker’s coordinate system. However, because $K_{\text{actual}}$ varies as a function of head translations, and because $K$ is estimated from a one-point calibration in which the participant looks at the center of the screen (see below), we think it is essential to use this reference point as the origin for the corrections. Otherwise the correction could introduce additional errors, especially if the position of the eye changes due to head motion.

**Filtering**

The coordinates of $P'$ and $C$ in the stimulus coordinate system were filtered with a median filter with a width of 20 samples (using the function medfilt1, Matlab 2016b), the same filter that was applied to $K_{\text{actual}}$. We decided to use a median filter, because this type of nonlinear digital filter smooths signals by attenuating noise whilst preserving edges in signals. For moderate to small levels of noise they prove to be efficient in removing small noise peaks with almost no impact on the dynamics of saccades.\(^{224}\)

**Eye-tracker validation study**

To validate the stereo eye tracker, we performed a study in which we recorded eye movements simultaneously with the stereo tracker and an Eyelink 1000 plus in remote tracking mode (which relies on tracking the pupil-glint vector). Because the Eyelink also has an infrared illuminator (890 nm), we removed one of the infrared illuminators of the stereo eye tracker and placed the stereo tracker on top of the Eyelink (Figure 5.5). The Eyelink camera (2048 x 2048 pixels) was equipped with a 16mm C-mount lens supplied by the manufacturer, and placed underneath the stimulus screen using its desktop mount. The camera screw was aligned with center of the monitor and the top of the illuminator was at approximately at the height of, and parallel with, the lower edge of the monitor. The eye-to-camera distance was ~55 cm, which is within the recommended range (between 40- 70 cm). The camera-to-screen distance was ~10 cm. Eyelink data were recorded at 500Hz with the sample filtering level set to ‘Standard’. This is a heuristic filter with a width of 3 samples.\(^{225}\)

Nine healthy participants (25 ± 4 years) with normal visual acuity were included in the validation study. None of the subjects wore glasses or contact-lenses. All participants gave informed consent before the start of the experiment. A bite-board was used to stabilize the head at ~65cm from the screen. The participants performed a visually-guided saccade task at nine different head positions. The participants started in the
central head position, in which they were positioned in front of the center of the screen by placing their “cyclopean eye” on the Z-axis of the stimulus coordinate system at a distance of 65 cm. In the other conditions the biteboard was translated ±5 cm horizontally, and ±3 cm vertically from the central position.

Before the start of the experiment a small target sticker was placed on the participant’s forehead, just above the eyebrows, and a monocular calibration procedure was performed at the central head position. During calibration of the right eye, the left eye was occluded, and during calibration of the left eye the right eye was occluded. This forced the subjects to fixate the calibration targets with the eye being calibrated. We made sure that the Eyelink software did not confuse the glint caused by its own IR illuminator and the (significantly) weaker one from the IR illuminator of the stereo tracker by raising the glint thresholds to sufficiently high levels (without losing the glints while the participant looked at the edges of the stimulus display). For the Eyelink we performed a 13-point calibration procedure, after which we performed a 13-point validation to check the quality of the calibration. The calibration was accepted only if the Eyelink software indicated that the quality was good, otherwise the calibration was repeated. For the stereo tracker we used a one-point calibration. Fixation of the central target during the Eyelink calibration was used to estimate the deviation between the optical and visual axes of the eye and the subject-specific
value of $K$. No recalibration was performed at the other head positions. In addition, we did not apply any drift corrections.

In the experimental task subjects had to make saccadic eye movements to visual targets at various locations on the screen (Dell U2412M, 1920 x 1200 pixels, pixel pitch 0.27 mm). In each trial a central fixation dot (0.5 deg in diameter) was presented at the center of the screen for a random duration of 1000-1600ms. Then the fixation spot was extinguished and the peripheral target (0.5 deg diameter) was presented at a pseudo random location for 1000 ms. The targets were presented at five different eccentricities (3, 6, 9, 12 and 15 deg). For eccentricities up to twelve degrees, targets were presented in twelve different directions (0:30:330 deg), and the targets at 15 degrees eccentricity were presented in six different directions (0, 45, 135, 180, 225 and 315 deg). This resulted in a total of 54 trials at each head position. Participants were instructed to fixate the center of the targets as accurately as possible. Stimulus presentation was done with custom Matlab software using the Psychophysics Toolbox\textsuperscript{226,227}. This stimulus software ran on the same laptop as the stereo eye tracking program. The EyeLink toolbox for Matlab\textsuperscript{228} was used for communication with the Eyelink computer.

Offline analysis of the data was done in Matlab. The sampling rate of the cameras of the stereo eye tracker is not fixed, primarily because occasional redetecting of the eyes takes time (Figure 5.2). If the eye positions are known, each camera can track stably at $\sim$350 Hz. In the present experiments, the average sampling rate for a given camera was 299 ± 29 Hz (range 212-336Hz). Because the cameras run asynchronously, the raw data of each camera were interpolated at the timestamps at which the other camera captured an image of the eyes. This resulted in a final gaze signal with an average refresh rate of 510 ± 92 Hz (range 349-660Hz) across the different conditions.

For one participant we could not collect data for one of the nine head position due to technical problems. In addition, for one participant we could only track one eye with the Eyelink for one of the head positions, and for another participant we could only track one eye with the stereo tracker for five head positions.

**Fixation analysis**

Because at each head position different gaze angles are needed to fixate a target at the same screen location, we used the POG estimates for the fixation analysis. This facilitates comparison of the results at different head positions. The POG estimates in
mm on the screen (Equation 5.10) were converted into POG estimations in degrees. These POG estimations in degrees reflect the orientation of an imaginary eye placed on the Z-axis of the stimulus coordinate system at 65 cm from the screen which is looking at the same point on the screen as the measured eye.

For each target the mean fixation location for the Eyelink and the stereo tracker was calculated by taking the average point of gaze during an 80 ms fixation window which started 20 ms after the end of the saccade. If a corrective saccade was made, we took an 80 ms fixation window which started 20 ms after the end of the corrective saccade. Saccades were detected on the basis of the calibrated Eyelink data with custom software. The detection of the saccade onsets and offsets was based on an eye velocity threshold criterion of 45 deg/s. All saccade markings were visually checked to exclude saccades in which blinks or other artefacts were present. Only trials without missing samples or artefacts during the fixation window for both eye trackers were included in the analysis. At the central head position, an average of 8 ± 7 percent of the trials had to be excluded per participant due to artifacts or missing samples from either the stereo tracker and/or the Eyelink. Overall, 13 ± 18 percent of the trials had to be excluded for each block of 54 trials per head position.

Subsequently, for each eye tracking system and for each head position and each measured eye, the mean absolute error (MAE) between the targets and the fixation positions was calculated for the horizontal POG and the vertical POG separately to assess the accuracy of both systems. To evaluate the precision of the eye trackers we adopted the two most commonly used measures of precision for eye tracking systems\textsuperscript{112}: the standard deviation (SD) and the root mean squared angular displacement (RMS[s2s]) of the samples in the 80 ms fixation window.

**Saccade analysis**

We analyzed the kinematics of the evoked saccades at the central head position to assess the dynamic properties of the two eye-tracking systems. For this analysis we used the estimated gaze angles instead of the point on gaze (POG) on the screen. For the Eyelink we converted the head referenced eye position estimates to eye rotation angles in degrees and for the stereo eye tracker we used the corrected gaze angles (Equation 5.12). The data from the stereo eye tracker were resampled at the timestamps of the Eyelink samples using linear interpolation to obtain a fixed sample rate for subsequent filtering. Following the recommendations of Mack et al.\textsuperscript{125}, we used a low-pass Butterworth filter (8\textsuperscript{th} order, 40Hz cut-off). Zero-phase filtering was
applied to avoid phase-distortion (using the function filtfilt, Matlab 2016b). Both the corrected gaze angles from the stereo tracker and the eye rotation angles from the Eyelink were passed through this filter. Subsequently, the filtered horizontal and vertical components were differentiated with respect to time by calculating the inter-sample difference in gaze angle (using the function gradient, Matlab 2016b) and dividing this difference by the inter-sample interval (2 ms). From this the vectorial eye velocity was computed using the Pythagorean equation.

The amplitude and the peak velocity were determined for the two systems for each first saccade in a trial. Subsequently, we determined the relation between saccade amplitude and peak velocity, i.e., the main sequence. We fitted an exponential function (see, e.g.) through the amplitude-peak velocity relation for the two systems for each participant:

$$v_{\text{peak}} = v_0 \left(1 - \exp \left(-\frac{\text{Amp}}{\text{Amp}_0} \right)\right)$$

(5.13)

where $v_{\text{peak}}$ is the peak velocity (in deg/s), $\text{Amp}$ is the saccade amplitude (in deg), $v_0$ is the saturation level (in deg/s) and $\text{Amp}_0$ is the shape parameter.
Results

The result of simultaneous recorded horizontal and vertical POG estimations in degrees (see methods) by both eye tracking systems for one trial are presented in Figure 5.6A and 5.6B. Figure 5.6C shows the corresponding 2D trajectories of the POG estimations. In this example the accuracy of both systems appeared to be comparable, while the precision of the Eyelink appeared to be slightly better. Furthermore, the vectorial eye velocity profiles, computed after applying a low-pass Butterworth filter (8th order, 40Hz cut-off) to the position data, were very similar for the two systems (Figure 5.6C). Additional examples at a larger scale are presented in Supplement 3.

Figure 5.7 shows the results of the fixation analysis for both systems for one participant in the central head position. In general, the POG estimations obtained with the Eyelink and the stereo eye tracker were close to the target (Figure 5.7A). As can be seen in Figure 5.7B and 5.7C, throughout the measured range there was a good correspondence between the location of the targets and the estimated POG for both systems.

![Figure 5.6 Simultaneous records of both eye tracking systems during one trial. A. The horizontal POG estimations as a function of time. B. The vertical POG estimations as a function of time. C. The corresponding vectorial eye velocity traces, calculated after applying a Butterworth filter (order 8, cut-off 40 Hz) to the position data. D. The 2D trajectories of the POG estimations. The results for both eyes are shown. The target positions are plotted as black circles.](image-url)
The accuracy of the Eyelink and the stereo tracker is displayed for all participants and for all head positions in Figure 5.8. At the central head position the accuracy of both systems was typically better than one degree. The average accuracy of the stereo eye tracker at the central head position was $0.69 \pm 0.21$ deg horizontally and $0.73 \pm 0.24$ deg vertically.

Figure 5.7. Fixation data from one participant in the central head position (both eyes). A. The 2D target locations and corresponding 2D POG estimations from the Eyelink and stereo tracker. B. The horizontal POG estimates plotted against the horizontal target location. C. The vertical POG estimates plotted against the vertical target location.
deg vertically, compared to 0.56 ± 0.18 deg horizontally and 0.73 ± 0.37 deg vertically for the Eyelink.

When the head was translated ±5 cm horizontally with respect to the central head positions the accuracy of both systems remained similar to the accuracy at the central head position. The average accuracy for the three central head positions was 0.73 ± 0.22 deg horizontally and 0.75 ± 0.26 deg vertically for the stereo tracker, and 0.63 ± 0.25 deg horizontally and 0.80 ± 0.40 deg vertically for the Eyelink.

However, as can be seen in Figure 5.8, vertical head translations affected the accuracy of both eye tracking systems. When the head was translated 3 cm upwards from the central head position, vertical gaze estimations of the Eyelink became less accurate. The average accuracy for the three upper head positions was 0.78 ± 0.36 deg horizontally and 1.53 ± 0.87 deg vertically for the Eyelink. For the stereo eye tracker the upward head translation did not reduce the accuracy, with an average accuracy of 0.71 ± 0.23 deg horizontally and 0.72 ± 0.27 deg vertically. The opposite

![Figure 5.8. Box plots showing the accuracy of the Eyelink and the stereo eye tracker for the nine different head locations. The mean absolute error (MAE) for all individual eyes are superimposed (dots).](image-url)
effect was observed when the head was translated 3 cm downwards. In that case the accuracy of the Eyelink remained stable with average accuracies of $0.56 \pm 0.20$ deg horizontally and $0.74 \pm 0.28$ deg vertically, while the vertical estimates of the stereo eye tracker became less accurate with average accuracies of $1.62 \pm 1.19$ deg vertically and $0.80 \pm 0.29$ deg horizontally. Although these translations of the head resulted in average errors of >3 degrees for some participants, this did not occur for all participants. Two participants had reduced accuracies after head translations in both systems, while one participant had reduced accuracies for the Eyelink and one participant had reduced accuracies for the stereo tracker.

To evaluate the precision of the eye trackers we determined the RMS[$s2s$], which is an indication of the inter-sample noise. The results of the RMS[$s2s$] for the nine different head positions are presented in Figure 5.9. The average RMS[$s2s$] across all head positions was $0.04 \pm 0.007$ deg horizontally and $0.03 \pm 0.008$ deg vertically for the stereo tracker and $0.03 \pm 0.009$ deg horizontally and vertically for the Eyelink.

Figure 5.9. Box plots showing the inter-sample precision (RMS[$s2s$]) of the eye trackers for the nine different head locations. The inter-sample precision (RMS[$s2s$]) for all individual eyes are superimposed (dots).
In addition, we calculated the standard deviation (SD) of the POG estimations during the 80 ms fixation window. This precision measure indicates how dispersed the samples were from the mean fixation position. The results for the SD at the nine different head positions are presented in Figure 5.10. The average SD across all head positions was $0.14 \pm 0.02$ deg horizontally and $0.10 \pm 0.02$ deg vertically for the stereo tracker and $0.06 \pm 0.01$ deg horizontally and $0.06 \pm 0.01$ deg vertically for the Eyelink.

Figure 5.10. Box plots showing the standard deviation (SD) of the data samples during fixation as a measure of the precision of the two eye trackers for the nine different head locations. The SD for all individual eyes are superimposed (dots).
The main sequence relationship between saccade amplitude and peak velocity (Equation 5.13) was determined independently for both systems. The main sequence relationship for each of the nine participants at the central head positions is presented in Figure 5.11. For both systems, the relationships were in line with those reported previously for visually-guided saccades. The fits of the main sequence relationship between amplitude and peak velocity were not significantly different for the two systems.

Figure 5.11. The relationship between the amplitude of the saccades and their peak velocity as measured with the two systems. Each panel presents the results for the left eye of one participant at the central head position. Each point represents one saccade. The lines represent the fits of the main sequence (Equation 5.13).
To analyze the correspondence between the two eye tracking systems, we constructed Bland-Altman plots for the saccade amplitude and the saccade peak velocity (Figure 5.12). In these plots, data from the central head position were pooled across eyes and participants. The average trial-to-trial difference between the two methods was small for both the amplitude (-0.07 deg) and the peak velocity (-16.6 deg/s), indicating that there was no significant bias in either measure. Because both eye trackers have an accuracy of ~0.7 deg in the horizontal and vertical direction, one may expect a standard deviation of the difference in amplitude of $\sqrt{(0.7^2 + 0.7^2)} + (0.7^2 + 0.7^2) \approx 1.4$ deg, which predicts a 95% CI of -2.8 - 2.8 deg.

Figure 5.12. Bland-Altman plots showing the trial-to-trial differences between the stereo tracker and the Eyelink and the 95% confidence intervals for: A. the saccade amplitudes, and B. the peak velocity of the saccades. Data are from the central head position, pooled across eyes and participants.
However, the 95% CI of the measured difference in amplitude was smaller (-2.02 – 1.88 deg), suggesting that the accuracy of the trackers as shown in Figure 5.7 is underestimated due to inaccurate fixations of the subjects.

Discussion
We successfully developed a high-speed stereoscopic eye tracker. The average sampling rate for each camera in the present experiments was ~300 Hz. The interpolated gaze signal from the two cameras combined had an average refresh rate of ~500 Hz. The sampling rate of the two cameras is variable, because occasional redetecting of the eyes takes time. Furthermore, the two cameras track the eyes asynchronously, which causes variation in the refresh rate of the stereo tracker.

To validate the stereo eye tracker we compared the accuracy of the point of gaze (POG) estimations obtained with the offline gaze reconstruction algorithms with the accuracy of an Eyelink 1000 plus in remote tracking mode. The results revealed comparable accuracies (<1 deg) at the central head positions. The Eyelink was less accurate than the specified 0.5 deg\textsuperscript{232}. One of the potential causes of the reduced accuracy we encountered could be fixation disparity (disagreement between the alignment of the left and right eye). We used a monocular calibration, but measured binocularly during the experiment. Under binocular viewing conditions, the maximum amount of disparity which still allows to fuse the input of both eyes into a single percept is about one-third of a degree\textsuperscript{233}. We did not analyze the fixation disparity of our participants during the experiment. However, inspection of the data showed that the POG estimations of the left and the right eye were not always consistent, which indicates that fixation disparity was present. This is in line with the results from the Bland-Altman plot analysis, which suggests that part of the measured fixation errors are due to inaccurate fixations of the subjects, rather than inaccuracies of the recording systems.

Vertical head displacements affected the accuracy of both eye tracking systems. The accuracy of the Eyelink was reduced for upward translations of 3 cm, while the accuracy of the stereo eye tracker was reduced for downward translations of 3 cm. This difference is most likely caused by the change in 3D eye position with respect to the camera(s) and IR light source(s). Although the translations of the head resulted in average errors of >3 degrees for some participants, this did not occur for all participants. This difference in reduced accuracy of the stereo eye tracker after vertical head translations may have been caused by differences in the anatomy of
the eye. Our stereo eye tracking method assumes a spherical shape of the cornea. However, in reality the cornea is slightly aspheric\textsuperscript{207}. This asphericity results in biased estimates of the eye’s optical axis\textsuperscript{218}. Due to the head translations, the asymmetry of the glints’ locations around the optic axis increased, causing a stronger effect of the asphericity. The corrections proposed in Equations 5.8-5.12 were quite successful in compensating for these effects, but we are the first to acknowledge that it is not immediately obvious how this is accomplished. To our knowledge, there are no stereo gaze reconstruction methods available at present that actually model the asphericity of the cornea. How commercially-available eye tracking systems, such as the Eyelink or the Tobii spectrum, may or may not compensate for head translations in remote tracking mode is currently undisclosed. Implementing an aspherical eye model in the stereo gaze reconstruction could, in principle, result in more accurate gaze estimations, reduced noise levels, and higher robustness against head movements. This would provide additional benefits in testing clinical populations or (young) children. In line with Guestrin and Eizenman\textsuperscript{223}, we found that the accuracy of the POG estimates did not improve significantly if Listing’s law was included in the POG estimation (not shown). This could be due to the simplifying assumption that Listing’s plane is parallel to the vertical XY-plane, but note that for the range of eye movements that we studied (eccentricities up to 15 deg) relatively small changes in eye torsion are expected in the first place. For a larger range of eye movements, independent measurement of eye torsion (e.g, by tracking the iris pattern) would be required to better account for the effects of eye torsion.

Previous studies on stereo eye tracking only described the accuracy of the prototypes, but did not report the precision of the systems. We used two precision measures as an indication of the noise level in the gaze estimations of both systems. The inter-sample noise of the two systems, as indicated by the RMS[s2s], was only 0.03 deg. For the Eyelink this is in line with the technical specifications provided by the manufacturer (<0.05 deg). Compared to other available eye trackers the RMS[s2s] level of our stereo eye tracker is comparable, or even lower (for overviews of the RMS[s2s] of different eye trackers, see\textsuperscript{217,234}). The second precision measure, the standard deviation of the samples, is an indication of the dispersion of the gaze estimates. The average SD of the stereo tracker was approximately twice the SD of the Eyelink (0.12 vs. 0.06 deg). Thus, the gaze estimates of the stereo tracker were more dispersed around the mean fixation position. This could be caused by an increased number of sources of noise due to segmentation errors and pixel noise.
The Eyelink uses only the center of the pupil and the center of one glint to estimate the gaze, thus there is noise in four degrees of freedom. The stereo tracker needs two pupil centers and four glint centers to estimate the gaze, therefore there is noise in 12 degrees of freedom. Another potential source of the increased noise level in the stereo eye tracking signals is the asynchronous sampling. If one of the cameras captured the eyes at a specific timestamp, we interpolated the raw data for the other camera. If those interpolated pixel coordinates do not provide an accurate representation of the actual 3D position and orientation of the eye, the resulting gaze estimation could deviate from the actual gaze position. This could have added variable errors to the gaze estimation of the stereo eye tracker. The results suggest that the impact of potential noise sources on the precision of the stereo eye tracker was higher for the SD compared to the RMS. Apparently, the noise reduction in Equation 5.4 and the application of the median filter on the 3D position data were more successful in removing sample-to-sample noise than low frequent variability. The exact source of this low frequent variability is unclear. A potential source for low frequent noise or drift in video-based eye trackers are pupil size variations. However, the small wobbles seen in the eye position signals from the stereo tracker during fixation epochs did not correlate with the variations in pupil size (see Supplement 5.3).

It may be possible to optimize the filtering of the stereo eye tracking data to increase its precision. However, it is important to carefully select the filters, as they can introduce artefacts and can impact the peak velocity of saccades\textsuperscript{125,126}. Another solution to decrease the level of noise could rely on an indirect mapping approach. Because the average accuracy of the stereo tracker is good, average gaze data from an arbitrary set of fixations throughout the field of interest can in principle be used as if they were fixations at known target locations to create a polynomial mapping function between the pupil-glint vectors and gaze for each camera, or one could train a neural network to convert the raw pupil and glint coordinates from each camera into gaze estimates (see e.g.\textsuperscript{235,236}). Such an indirect mapping approach would reduce the number of noise sources in the reconstructions and independent gaze estimates from the two cameras can be averaged while still avoiding an elaborate calibration procedure. Furthermore, the noise level could decrease by increasing the spatial resolution of the eye images. This can be achieved by moving the cameras closer to the eyes, by using different lenses, or choosing a higher resolution setting on the camera. In the latter case, however, the temporal resolution of the eye tracker will
decrease while the other two options will restrict head motion.

Finally, the analysis of the saccade kinematics revealed the relationship between amplitude and peak velocity of the saccades, i.e., the main sequence\textsuperscript{229}, were similar for the two systems. Moreover, a direct trial-by-trial comparison of the amplitude and peak velocity measurements obtained the two systems showed good agreement (Figure 5.12). Therefore, the stereo eye tracker may be used to determine overall differences in the main sequence between different conditions and/or different participant groups.

Our participants did not wear glasses or contact-lenses. We recommend assessing the accuracy and precision of the stereo tracker in participants with glasses and contact lenses. The gaze reconstruction of the stereo eye tracker is based on an eye model. Therefore, the optics of glasses, especially in case of astigmatism, could significantly impact the gaze reconstruction. Furthermore, detection of the eyes could become problematic in case of glasses. The software uses a simple classifier to detect the eyes, the same classifier as was implemented in the ITU gaze tracker\textsuperscript{221}. This classifier might not be the optimal solution for the eye detection. It takes relatively long to detect the eyes and it occasionally detects the nose instead of an eye. The eye tracking software could be adapted to select the location of the eye manually if a bite-board or chin-rest can be used to stabilize the head. This would result in higher and more constant sampling rates. Moreover, other options to detect the eyes exist (for an overview, see\textsuperscript{237}). However, it is likely that these other methods are still relatively slow. The easiest option would be to use a marker placed under, or above the eyes, for example a calibrated black square on a white sticker. Not only is such a marker easy to detect, it also provides additional information about the 3D position and orientation (i.e., yaw, roll and pitch) of the head.

In conclusion, we successfully developed a high-speed (>350Hz) stereoscopic eye-tracker. The validation study shows comparable accuracy (<1 deg) for the Eyelink 1000 plus and our stereo system. The noise level of the stereo tracker is slightly higher. Application of the stereo eye tracker could be particularly helpful when calibration is not possible, or when experimental time is limited. In addition, it could facilitate testing of children and clinical populations (see Supplement 5.4 for proof of principle). Finally, it could be beneficial to use the stereo eye tracker in (the training of) naïve experimental animals such as macaque monkeys, or in test situations in which the relative gaze angles provide sufficient information (e.g., to quantify the amplitude and frequency of a nystagmus).
Development of a high-speed stereoscopic eyetracker
Supplement 5.3.

Four examples of simultaneous records of both eye tracking systems during one trial for three different participants (PP 4, 8 and 6). A. The horizontal point of gaze (POG) estimations as a function of time. POG data are expressed in degrees (See Methods). B. The vertical POG estimations as a function of time. C. The corresponding vectorial eye velocity traces (in deg/s), calculated after applying a Butterworth filter (order 8, cut-off 40 Hz) to the position data. D. Pupil size as a function of time. The pupil size data from the Eyelink are in arbitrary units, the pupil size data of the stereo tracker are in pixels. A median filter with a width of 20 samples (using the function medfilt1, Matlab 2016b) was used to filter the pupil data of the stereo tracker. Black traces indicate the location of the visual target.
Development of a high-speed stereoscopic eyetracker

A  B  C  D

Stereo tracker left  Stereo tracker right  Eyelink left  Eyelink right

Horizontal POG (deg)

Vertical POG (deg)

Velocity (deg/s)

Time (ms)

Pupil size

PP 4
Chapter 5

Stereo tracker left
Stereo tracker right
Eyelink left
Eyelink right

Pupil size

Horizontal POG (deg)
Vertical POG (deg)
Velocity (deg/s)

Time (ms)

PP 6 A B C D

0 5 10 15 20 25 30

0 1 2 3 4 5 6

0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

136
Supplement 5.4.

To illustrate the feasibility of applying stereo eye tracking together with a one-point calibration procedure in the target groups, we present data from two children in a head-free condition, an 11 year old with normal vision and a 10 year old with glasses and nystagmus. A. The horizontal point of gaze (POG) estimations as a function of time. POG data are expressed in degrees (See Methods) B. The vertical POG estimations as a function of time. C. The corresponding vectorial eye velocity traces (in deg/s), calculated after applying a Butterworth filter (order 8, cut-off 40 Hz) to the position data. D. Pupil size as a function of time. The pupil size data of the stereo tracker are in pixels and were filtered with a median filter with a width of 20 samples (using the function medfilt1, Matlab 2016b). Black traces indicate the location of the visual target.
Chapter 5

11 yo with normal vision (headfree)

Horizontal POG (deg)

Vertical POG (deg)

Velocity(deg/s)

Time (ms)

Pupil size

Stereo tracker left

Stereo tracker right
Development of a high-speed stereoscopic eyetracker

10 yo with glasses and nystagmus (headfree)

- A: Horizontal POG (deg)
- B: Vertical POG (deg)
- C: Velocity (deg/s)
- D: Pupil size

Time (ms)

Stereo tracker left
Stereo tracker right
Saccade latencies during a preferential looking task and objective scoring of grating detection in children with and without visual impairments

Submitted as: Barsingerhorn AD, Boonstra FN & Goossens J. Saccade latencies during a preferential looking task and objective scoring of grating detection in children with and without visual impairments
Introduction

The most commonly used clinical method to assess visual acuity in infants, toddlers, and non-verbal children is the preferential looking technique\textsuperscript{15}. This method is based on the phenomenon that infants, when simultaneously presented with a patterned target and a blank target, have a greater tendency to look at the pattern\textsuperscript{16}. The grating with the finest stripes that yields a consistent orienting response provides an estimate of the child's visual acuity. The standard diagnostic tool, which uses the preferential looking technique to assess visual acuity, is the Teller acuity card test (TAC)\textsuperscript{17,18}. The advantage of the TAC is that it provides a fast and fairly reliable estimate of grating visual acuity\textsuperscript{238}. However, the outcome of the TAC relies on a subjective assessment of whether the patient can see the grating\textsuperscript{238}. This assessment is not only based on visual judgment of the eye and head movements; other factors such as verbal responses, facial expressions, or pointing also influence the examiner's judgment. Thus, the outcome of the TAC depends on the experience of the clinician and on the cooperation and attention of the child\textsuperscript{15}. Teller suggested already in 1983 that the analysis of the infant’s head and eye movements could provide valuable and more objective information\textsuperscript{239}. Only recently, several researchers have attempted to create more objective preferential looking paradigms by combining eye tracking with computerized preferential looking tasks\textsuperscript{240–242}. These studies assessed visual acuity based on the eye tracking data and compared the results with the outcome of traditional preferential looking tests, such as the TAC and the Keeler infant acuity cards\textsuperscript{240–242}. Different measures have been adopted to quantify the visual scanning behavior during these tasks and to assess whether the subject resolves the grating. These measures are all based on the assumption that the grating is resolved if the target pattern is fixated for a prolonged period. Two studies used a relative fixation-time criterion by measuring the percentage of the time that a participant fixates the patterned target field during a trial\textsuperscript{240,241}. A third study used an absolute fixation-time criterion (at least 167 ms within the target area) to assess whether the grating was resolved\textsuperscript{242}. The visual acuity estimates based on prolonged viewing of the target corresponded well with the outcome of the traditional preferential looking tests in adults\textsuperscript{240,241} and in infants\textsuperscript{241,242}. In conclusion, computerized tests which combine preferential looking paradigms with eye tracking provide a rapid, automated, and more objective measure of grating acuity. In addition, it has been shown that eye tracking can provide additional information about other visual functions, such as visual field size, contrast sensitivity and color perception\textsuperscript{243,244}. 


Previous studies that have used eye tracking to evaluate preferential looking behavior in computerized TAC tests have only assessed its potential to estimate the visual acuity of the participants. However, as Teller stated: “... the quality and intensity of the infant’s staring behavior on each trial contains more information than one gets out of the single left-right judgement imposed by the forced choice method.” Video-based eye tracking techniques now allow for quantitative assessment of this behavior. One of the important variables that can be assessed is the saccade latency. Saccades are the fast eye movements that change the line of sight from one point of fixation to another. Saccade latency is the interval between stimulus presentation and the onset of a saccade and this latency reflects visual processing, target selection and motor programming. Furthermore, saccade latencies are abnormal in a range of disorders in which cortical areas associated with vision and eye movements are affected. Therefore, quantifying the latencies and the accuracy of saccades evoked during a preferential looking task could provide valuable insight into the development and integrity of the oculomotor and the visual system.

Saccade latencies can be influenced by a wide range of factors, such as the contrast and luminance of the target, the amplitude and direction of the saccade and the nature of the task (for an extensive overview, see). Furthermore, the spatial frequency of the target influences saccade latency as well. This is relevant when assessing the latencies of orienting responses evoked during preferential looking tests, as the spatial frequency of the gratings is systematically varied in these tests. In addition, several studies in participants with normal vision have demonstrated that saccade latencies are longer in children than in adults, and that the latencies decrease with age through childhood, until adult levels are reached at approximately 10 to 12 years of age. Therefore, reference values for saccade latencies during preferential looking paradigms have to be age- and target specific.

Although preferential looking tests are often used in children with visual impairment, previous studies which combined computerized preferential looking tests with eye tracking only tested adults and infants with normal vision but have not addressed ophthalmological abnormalities. Video-based eye tracking in participants with ophthalmological problems can be challenging because certain eyes are difficult to track, for instance due to the presence of nystagmus, or abnormal anatomical properties of the eyes. Nonetheless, saccade latencies in participants with visual impairments have been assessed in other tasks. Recently, it has been shown that children and adults with infantile nystagmus have longer saccade
latencies, and that children with cerebral visual impairment (CVI) have delayed orienting responses towards cartoons. However, in these studies only one spatial frequency or one stimulus size was presented. Assessment of saccade latencies and saccade metrics in preferential looking test provides measures of both speed and accuracy of the visual system for different spatial frequencies. It has been argued that such measures are key to better quantify visual impairment.

The aim of the present study is to determine the latencies of orienting responses during a preferential looking task in children with normal vision and in children with visual impairments between 6 and 12 years old, and to assess the feasibility of scoring grating detection in these populations with video-based eye tracking. Towards that end, we compared different detection scoring methods and identified factors which could reduce the chance of successful eye tracking. We also compared the latencies of stimulus-evoked primary saccades for the two populations to determine whether the onset of the orienting responses was delayed in the children with visual impairments.

Methods

Participants

The current study was part of a larger project in which we assessed visual processing speed in children with normal vision and in children with visual impairments. The children who participated in the present study also participated in those previous studies. 88 children (9.6 ± 1.8 years) with normal vision (NV), 15 children (9.0 ± 1.6 years) with cerebral visual impairment (CVI) and 19 children (9.0 ± 2.4 years) with visual impairment due to congenital or acquired disorders of the eye without additional impairments (mental or neurological) (VI_o) participated. The children with VI_o and CVI were recruited through Bartiméus, a specialized Dutch institute for visually impaired people. The children with NV were recruited from primary schools in the surrounding region. For the children with NV and VI_o the following inclusion criteria were applied: age 6 to 12 years old, birth at term, normal birth weight, no perinatal complications and normal development. In addition, children with NV had to have a crowded visual acuity of 0.1 logMAR or better. Additional criteria for children with VI_o were: a crowded distant visual acuity (DVA) between 0.2 and 1.3 logMAR and a congenital or acquired ocular abnormality without mental or neurological impairment. For the children with CVI inclusion criteria were: being diagnosed with CVI by experienced pediatric ophthalmologists at the institute for the
visually impaired based on a thorough ophthalmological examination and detailed patient history, age 6 to 12 years, and having a crowded distance visual acuity of 1.3 LogMAR or better. All children performed the Freiburg visual acuity test (FrACT\textsuperscript{173}) to verify if they met the acuity inclusion criterion. Children who did not have the mental and motor skills to understand and execute the tasks were excluded. Clinical characteristics and visual acuities of the children with VI\textsubscript{o} and CVI are presented in Supplemental Table S6.1.

The study was conducted according to the principles of the Declaration of Helsinki and was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands). All parents or legal guardians provided informed consent in writing before the start of the study.

**Procedure**

The children were seated unrestrained at approximately 65 cm from a 23-inch LCD screen (Dell U2412M, 1920 x 1200 pixels, pixel pitch 0.27 mm), and were instructed to keep their back against the back of the chair to keep the distance to the screen constant during the experiment. Similar to Sturm et al. (2011), each stimulus consisted of a 2 × 2 grid of target fields on a black background. One field contained a black-and-white square wave grating while the other three were uniform grey fields (Figure 6.1). The grating was presented randomly at one of the four locations. Each field subtended 9.6 × 9.6 deg and the center of each field was positioned 10.2 deg from the center of the screen. Three different spatial frequencies were used: 1.05, 2.11 and 7.02 cyc/deg. This corresponds to 1.45, 1.15 and 0.63 LogMAR. The luminance of the background and the black stripes was 2 cd/m\textsuperscript{2} and the luminance of the white stripes was 236 cd/m\textsuperscript{2}, as measured with a luminance meter (Minolta LS-100; Minolta Co. Ltd., Osaka, Japan). The mean luminance of the grey fields (~118 cd/m\textsuperscript{2}) was matched to the space-average luminance of the grating to prevent that the participants could detect the position of the grating based on differences in luminance. In contrast to the previous studies that have used eye tracking during a preferential looking task\textsuperscript{240–242}, we added a high contrast fixation dot (98.2% Michelson) at the center of the screen before each trial. The fixation dot was presented for a random duration of 320-640 ms, after which the fixation dot disappeared and the stimulus was presented for 3 s (Figure 6.1). The three spatial frequencies were presented in pseudo random order with 10 trials per frequency, resulting in a total of 30 trials. The children were instructed to fixate the fixation dot
Saccade latencies during a preferential looking task at the start of each trial and to look at the target as soon as the fixation dot disappeared.

Custom Matlab software (version 2013b; MathWorks, Inc., Natick, MA, USA) with the Psychophysics Toolbox (version 3.0.12) was used to generate the stimuli. The software was executed on a laptop (Dell M3800; Dell Inc., Round Rock, TX, USA) equipped with an OpenGL graphics card (Nvidia Quadro K1100M; Santa Clara, CA, USA).

Figure 6.1. Stimuli of the preferential looking task. Each trial started with a central fixation dot which was presented for a random duration between 320-640 ms. As soon as the fixation spot disappeared, the stimulus array appeared for a duration of 3000 ms. The children were instructed to fixate the fixation dot at the start of each trial and to look at the square wave grating target as soon as the fixation dot disappeared. The target (three different spatial frequencies) appeared randomly at one of the four positions, the other fields were grey.
Eye tracker

We used a stereoscopic eye tracking system with two USB 3.0 cameras and two infrared light sources in a fully calibrated configuration with respect to the stimulus screen\textsuperscript{248}. In this way we could obtain calibrated measures of the two-dimensional (2D) orientation the optical axes of the two eyes and their three-dimensional (3D) location in space. During data collection, each camera tracked both eyes with an average framerate of \textasciitilde300 Hz. The image coordinate of the pupils and the image coordinates of the reflections of the infrared light sources were extracted online and stored on disk for off-line reconstruction of gaze\textsuperscript{119,120,203,248}. After combining the asynchronously sampled data from the two cameras, the final gaze position signals had an average refresh rate of \textasciitilde500 Hz. The spatial accuracy of these signals was \textasciitilde0.7 degrees in both directions\textsuperscript{248}. The advantage of this stereoscopic system is that a one-point calibration procedure is enough to obtain calibrated measures of gaze under head-free conditions. Only the subject-specific angles between the optical and visual axes must be determined for a given participant. In the current study we used the gaze position data when participants where fixating the central fixation dot during the preferential looking task as the one-point calibration. Fixation periods were identified by the experimenter by using a mouse tool which marked the beginning and end of stable fixation of the fixation dot. Therefore, there was no separate calibration necessary before the start of the test.

Data processing

The data were analyzed in Matlab. The sampling rate of the gaze position signals was variable because the two cameras of the stereo eye tracker ran asynchronously\textsuperscript{248}. Therefore, the data of the stereo eye tracker were resampled to a fixed sampling rate of 500 Hz using linear interpolation to facilitate the saccade detection based on velocity and acceleration thresholds. Saccades were detected with a velocity threshold criterion of 25°/s and an acceleration threshold criterion of 3000°/s\textsuperscript{2} for saccade onsets and offsets. All saccade markings were visually checked and corrected if necessary. Subsequently, saccade latency was determined as the difference between stimulus onset and the onset of the saccade. Primary saccades had to start within 80 to 900 ms after stimulus onset. Participants were excluded from the analysis if no saccades were found in more than one third of the trials.

The median latency of the primary saccades was determined for each participant and for each spatial frequency. Subsequently, primary saccades were categorized
into correct, goal-directed saccades and incorrect saccades. Only trials in which the starting position of the primary saccade fell within a square window of $2.75 \times 2.75$ deg centered at the fixation dot were included in this analysis. This excluded trials in which children already fixated the location of the target pattern by chance before the stimulus onset. Correct saccades were those primary saccades that had an amplitude of $>1.5$ deg and landed on the target, or within 1.4 deg from the target boundary. Incorrect saccades were saccades of $>1.5$ deg which landed outside this target window. We determined the median saccade latencies for the correct and the incorrect saccades separately.

The accuracy of the responses was assessed with different scoring methods to account for the variance in response patterns observed during the experiment. Figure 6.2 illustrates the variability in response patterns by showing four trials of a child with normal vision (Figure 6.2A) and four trials of a child with visual impairment due to albinism and with a nystagmus (Figure 6.2B). We compared three methods to establish which method is most reliable in discriminating whether the children could resolve the gratings. In the first method, we scored the accuracy based on the endpoint of the primary saccades. For each participant and for each spatial frequency we calculated the percentage of correct primary saccades, i.e., trials in which the primary saccade landed on target or within 1.4 deg from the target boundary (Figure 6.2A-1 & 6.2B-1). This analysis provides insight in the accuracy of the first orienting response. However, if the primary saccade was in the wrong direction, this did not necessarily mean that the participant could not resolve the grating. In a large number of these trials, the participants seemed to have guessed the potential location of the grating but corrected their initial error by making a second, goal directed saccade to fixate the grating (Figure 6.2A-2). In addition, the children did not always fixate the fixation dot at the time the stimulus appeared, but their first saccade was directed toward the grating (Figure 6.2A-4 & 6.2B-2-4). It also happened that the participant’s gaze was already at the location of the grating at the time the stimulus appeared, after which the subject continued to fixate this target. To account for these behaviors, we used a second method to assess whether participants successfully located the target position. In this second method, accuracy scores are based on prolonged viewing of the target $^{240,241}$. Only trials in which the gaze position was available for at least 2 out of 3 seconds were included in the analysis. The gaze position was considered to be on target if it fell within the target area (see Figure 6.2), i.e., on target or within 1.4 degrees from the target boundary.
Figure 6.2. Illustration of observed response patterns and scoring criteria. A. Four trials of a child with normal vision. B. Four trials of a child with visual impairment due to albinism and with a nystagmus. The target window is indicated with the blue dashed lines, the central fixation window is indicated with the yellow dashed lines, the gaze coordinates are plotted as red lines and the primary saccade is plotted in green.
We calculated the percentage of presentation time in which the gaze position fell within the target area. To limit the possibility of false positives, we considered that the participant could resolve the grating if the relative fixation time (RTF) exceeded 50%, i.e., if the gaze of the participant was on target during 50% of its presentation time. Subsequently, for each participant and each spatial frequency, we calculated the proportion of trials in which the participant resolved the grating according to this RTF criterion. The examples presented in Figure 6.2A are all responses in which the child correctly located the grating. However, the trials in Figure 6.2A-2 and Figure 6.2A-4 would be considered inaccurate with the first method and accurate with the second method, while the trial in Figure 6.2A-3 (goal-directed saccade after which the gaze returned to the center of the screen) would be considered inaccurate with the second method and accurate with the first method. Therefore, we combined the two previous methods for the third method. In this case, a response was considered correct if 1) the primary saccade was goal-directed, or if 2) the RTF exceeded 50%. As a result, all examples in Figure 6.2A would be correctly scored as accurate responses with the third method. Since the gratings were randomly presented at four different positions on the screen, we considered that the participant could resolve the grating if he or she correctly looked at the grating in more than 62.5% of the trials (i.e., halfway between 25% chance level performance and 100% correct performance).

**Statistical analysis.**

We assessed whether the latencies of the primary saccades change with age, whether the latencies depend on the spatial frequency of the grating, and whether the developmental effects are equal for the three spatial frequencies for children with normal vision. A repeated measures ANOVA with age and spatial frequency as independent variables and the saccade latency of the primary saccade as dependent variable was performed. We also applied the repeated measures ANOVA separately for the saccade latencies of the correct and incorrect primary saccades. For all repeated measures ANOVAs age was centered on the age of 9, the middle of the inclusion range. Children were only included in the repeated measures ANOVAs if latencies were available for all three grating frequencies. Subsequently, we used the multiple linear regression models from the repeated measures ANOVAs to determine the upper 95\textsuperscript{th} percentile in the data from the children with normal vision. The onset of the saccades of an individual child with visual impairments was considered to be delayed if the median latency exceeded the upper 95\textsuperscript{th} percentile of the normative data. Alpha (type 1 error) was set on 0.05 for all statistical group comparisons.
Results

The flow chart in Figure 6.3 shows the number of children who participated, the number of children from whom we could collect eye tracking data, and the number of children in whom the quality of this data was deemed sufficient for analysis. For participants in whom eye tracking was not possible, we reported the main reason why the eye tracking failed. For 71/88 children with normal vision and 15/34 children with visual impairments (nine children with VI and six children with CVI) we collected eye tracking recordings. For 56/88 children with normal vision and 13/34 of the children with visual impairments (seven children with VI and six children with CVI) the quality of the eye movement data was sufficient for the analyses (Figure 6.3).

In Supplemental Table S6.1, we indicated for all children with visual impairments whether we were able to record their eye movements. The main reason why eye tracking failed was the presence of eyeglasses. This was particularly problematic in the children with visual impairments, as most of these children wore prescription glasses. However, the presence of eyeglasses did not necessarily mean that eye tracking was impossible. Eight of the thirteen children with visual impairments in whom we were able to collect valid eye tracking data wore glasses. In addition, in five children with normal vision and prescription glasses we could collect eye tracking data as well. Factors that influenced the success of eye tracking in children with eyeglasses were related to the size and color of the frame, and thickness of the glasses. Smaller frames, black frames, and thick glasses made detection of the eyes and/or extraction of the relevant image features difficult.

Figure 6.3. Flow chart showing the total number participants, the number of children in which we could collect eye tracking data, and the number of children in which the quality of the eye-tracking data was deemed sufficient for analysis for A. children with normal vision, and B. children with visual impairments. In case eye tracking was not possible, the main reason why the eye tracking failed is listed.
Accuracy

The grating-detection accuracy of the children with normal vision is presented in Figure 6.4A for each of the three different scoring methods. Because the spatial frequencies of the gratings were at least 0.6 LogMAR above their visual acuity (as measured with the Freiburg acuity test), the children with normal vision should have been able to resolve gratings of all three spatial frequencies. Therefore, children with normal vision were expected to look correctly at the grating in more than 62.5% of the trials. The accuracy of the responses was determined with three different methods, to establish which method is most reliable in discriminating whether the children could resolve the gratings. To that end, we determined the number of children who did not reach the 62.5% threshold.

First, we scored the grating-detection accuracy as the percentage of correct primary saccades, i.e., trials in which the primary saccade landed on target or within 1.4 deg from the target boundary. With this method 10/56 participants with normal vision did not reach 62.5% correct for the finest grating, 3/56 did not reach 62.5% correct for the middle grating, and 1/56 did not reach 62.5% correct on the coarsest grating. The median percent correct was 85%, 89% and 89%, respectively. With the second scoring method, we considered that the participant could resolve the grating if the relative fixation time (RTF) exceeded 50%, i.e., if the gaze of the participant was on target during 50% of its presentation time. This method resulted in a total of 6/56 children on the finest grating, 3/56 children on the middle grating, and 4/56 children on the coarsest grating who did not reach beyond 62.5% correct. The median accuracy with this method was 100% for all three spatial frequencies. The third scoring method combines the two other methods, i.e., a response was considered correct if the primary saccade was goal-directed, or if the RTF exceeded 50%. This third method was most reliable to determine whether children with normal vision could resolve the gratings; only one child did not reach beyond the 62.5% threshold on the middle grating. Furthermore, the median accuracy was 100% for all spatial frequencies as well.

The accuracy of the responses of the children with visual impairments is presented in Figure 6.4B. For only one of these children, the finest grating and the middle grating fell below its visual acuity (see Supplement 6.1). The coarsest grating was above visual acuity for all children and therefore, they should have been able to resolve these gratings. However, for the finest and middle grating, 10/13 children with visual impairments did not reach the 62.5% correct threshold if the response
accuracy was determined from the endpoints of the primary saccades alone. For the coarsest grating, the accuracy of the primary saccade fell below 62.5% correct in 9/13 children. The median accuracy of the first orienting response was only 33%, 28%, and 50%, respectively. Mann-Whitney U-tests revealed that the accuracy of the children with visual impairments was significantly lower than the accuracy of the children with normal vision for all three grating frequencies as scored by the endpoint of the primary saccades ($Z=3.75$, $p<0.001$, $Z=4.40$, $p<0.001$, and $Z=4.27$, $p<0.001$ respectively). Similar to the results in the children with normal vision, the accuracy scores were higher if the scoring was based on sustained viewing of the target pattern (RTF). In this case, the number of children who did not score above the 62.5% correct threshold was 4/13 for the finest grating, 3/13 for the middle grating, and 5/13 for the coarsest grating. The median accuracy based on prolonged viewing of the target pattern was 67%, 85%, and 67%, respectively. Even though the accuracies of the children with visual impairments were higher with this method compared to the first method, the accuracies as scored with the second method

Figure 6.4. Boxplots of the accuracy of the responses determined with three different scoring methods: 1. Primary saccade ends on the target area (PrimSac), 2. Relative fixation time (RTF) > 50%, i.e., if the gaze of the participant was on target during 50% of its presentation time, and 3. Primary saccade ends on the grating) or RTF exceeds 50% (Combined). A. grating-detection accuracy of the children with normal vision, and B. grating-detection accuracy of the children with visual impairments. The colours indicate the spatial frequency of the grating.
were significantly lower than the accuracies of the children with normal vision ($Z=3.44$, $p<0.001$, $Z=2.62$, $p=0.008$, and $Z=3.10$, $p=0.002$ respectively). As for the children with normal vision, the third scoring method yielded the best accuracies, with median scores of 78% for the finest, 87% for the middle grating, and 83% for the coarsest grating. However, still 5/13 children with visual impairment did not reach the 62.5% correct threshold for the finest grating, and this number was 3/13 on the middle and coarsest grating. Furthermore, even with this method, the accuracies of the children with visual impairments were significantly lower than the accuracies of the children with normal vision ($Z=4.02$, $p<0.001$, $Z=3.82$, $p<0.001$, and $Z=3.66$, $p<0.001$ respectively).

Saccade latencies

The saccade latencies of children with normal vision are presented in Figure 6.5. The data are stratified by spatial frequency (colors) and plotted as a function of age. First, we analyzed the latencies of all primary saccades, independent of whether they were directed towards the grating or not (Figure 6.5A). A repeated measures ANOVA ($n=56$) revealed that the saccade latencies significantly decreased with age (main effect: $F(1,54)=9.59$, $p=0.003$) and that this developmental effect did not differ significantly between the three spatial frequencies (interaction spatial frequency $\times$ age: $F(2,108)=0.81$, $p=0.45$). However, a significant difference between the three spatial frequencies was found (main effect: $F(2,108)=16.85$, $p<0.001$). Post-hoc t-tests showed that the saccade latencies were on average longer for the finest grating compared to the middle (difference $30 \pm 6$ ms, $p<0.001$) and the coarsest grating (difference $31 \pm 7$ ms, $p<0.001$). The average difference between the latencies for the middle and coarsest grating was not statistically significant (difference: $0.4 \pm 5$ ms, $p=0.99$).

The results of the ANOVA on the reaction times of only the goal-directed primary saccades (Figure 6.5B) were very similar. Seven children had to be excluded from this ANOVA because they did not make correct saccades towards all spatial frequencies. As a result, 49 out 56 children with normal vision were included in this analysis. The latencies of the correct saccades decreased significantly with age (main effect: $F(1,47)=5.06$, $p=0.03$) and this developmental effect did not differ significantly between the gratings (interaction: $F(2,94)=0.17$, $p=0.85$). Furthermore, a significant main effect of spatial frequencies was found ($F(2,94)=13.56$, $p<0.001$): saccadic latencies were on average longer for the finest grating compared to the
middle (difference 28 ± 6 ms, \( p < 0.001 \)) and the coarsest grating (difference 35 ± 8 ms, \( p < 0.001 \)). The difference between the middle and coarsest grating was not statistically significant (difference: 7 ± 5 ms, \( p = 0.39 \)).

The saccade latencies of the incorrect primary saccades are presented in Figure 6.5C. Only 18 children made saccades in the incorrect direction for all the spatial frequencies. Due to this low number of children, especially the low number of older children (two 10-year-olds, one 11-year-old and one 12-year-old), interpretation of age effects and effects of grating acuity could be spurious. Therefore, we did not perform a repeated measures ANOVA on this data. Instead, we only compared the latencies of correct saccades to the latencies of all primary saccades. This repeated measures ANOVA with the type of latency measure as an additional within-subject factor showed that the latencies of correct saccades were on average 4 ± 2 ms faster than the latencies of all primary saccades (\( F(1,47) = 5.7, p = 0.02 \)), while the effects of age and spatial frequency were not significantly different between these two measures (\( p \)-values > 0.2).

The children with visual impairments had on average more trials in which they did not correctly fixate the fixation dot. Therefore, we only analyzed the latencies of
all primary saccades in the children with visual impairments. The saccade latencies of the children with visual impairments are presented in Figure 6.6. In total, 8 out of 13 children with visual impairments scored ≥ 95th percentile of the saccade latencies of the children with normal vision for at least one of the spatial frequencies. Three children had longer latencies (≥ 95th percentile) on all spatial frequencies, one child on the finest and the middle grating, one child only on the finest grating, and three children only on the coarsest grating. To test whether on average the children with visual impairments had longer saccade latencies, we performed a repeated measures ANOVA with age, spatial frequency and group (normal vision vs. visual impairment) as the independent variables and the latency of the primary saccade as the dependent variable. The results revealed a significant effect of group \((F(1,65)=16.26, p<0.001)\): the saccade latencies of the children with visual impairment were on average 62 ± 15 ms longer.

Figure 6.6. Median saccade latencies for the children with visual impairments (numbers correspond with individual participants, a circle indicates a child with CVI, Supplemental Table S6.1) for A. gratings of 0.63 LogMAR, B. gratings of 1.15 LogMAR, and C. 1.45 LogMAR. The solid lines are the result of the repeated measures ANOVA and the dashed black lines indicate the 95th percentile in children with normal vision. Latency data were pooled across correct and incorrect trials.
Discussion

The aim of the present study was to determine the latencies of saccadic eye movements evoked during a preferential looking task in children with normal vision and in children with visual impairments between 6 and 12 years, and to assess the feasibility of scoring grating acuity in these populations with video based eye tracking. In line with previous research\textsuperscript{161–164}, the saccade latencies of the children with normal vision decreased with age. This developmental effect was similar for all spatial frequencies. Furthermore, both the latencies of all primary saccades and the latencies of the correct saccades were longer for the finest grating of 7 cyc/deg, even though this grating was still well above the children’s visual acuity (at least 0.6 LogMAR). A similar effect of spatial frequency on saccade latencies has been found in adults; in adults latencies increased for higher spatial frequencies\textsuperscript{245}.

To determine whether the onset of the saccades of children with visual impairments was delayed, we compared their saccade latencies with the data from controls. In 8/13 children with visual impairments, the saccadic latencies were abnormally long (> 95\textsuperscript{th} percentile of children with normal vision) for at least one of the spatial frequencies and on average the children with visual impairments were 62 ± 15 ms slower than children with normal vision. This corresponds well to previous studies, which revealed longer saccade latencies in children and adults with infantile nystagmus\textsuperscript{246,247}, and in children with cerebral visual impairment\textsuperscript{69}. Apart from longer saccade latencies, children with visual impairments also have longer search times\textsuperscript{36,37,72}, lower reading speeds\textsuperscript{73,74}, and they need more time for the discrimination of optotypes\textsuperscript{188} compared to children with normal vision.

Previously, it has been argued that eye tracking in participants with ophthalmological problems can be challenging, because their eyes can be more difficult to track \textsuperscript{19}. This is in line with our experience. For 19/34 of children with visual impairments we could not obtain reliable eye tracking data. The limited number of children with a certain diagnosis makes it difficult to draw conclusions on whether certain diagnoses or clinical characteristics (e.g., nystagmus, astigmatism, and high hyperopia) result in difficulties with eye tracking. The problems with eye tracking in clinical populations could largely differ between different eye trackers. Commercially available eye trackers might use different algorithms to detect the eyes, which might be more robust against anatomical abnormalities and/or eyeglasses. However, in a recent study on oculomotor behavior of children with infantile nystagmus, eye tracking with an Eyelink 1000 plus was only successful in 47%-72% of the participants\textsuperscript{246}. Thus, also
commercial eye trackers appear to have problems with tracking the eyes of clinical populations.

Because our goal was to assess the feasibility of including saccade latencies as an outcome measure of a preferential looking task in children with visual impairments, we only used three spatial frequencies. As a result, we could not assess grating acuity of the children and we could not assess whether the saccade latencies would increase even more for finer gratings towards the discrimination threshold. Recent studies on the time children need to discriminate optotypes revealed that the button-press reaction times increased as the optotypes approached the discrimination threshold in children with normal vision and in children with visual impairments. In addition, because in these studies the optotypes ranged from below visual acuity to well above threshold, the reaction times could be corrected for reduced the visual acuity. The results revealed that 40% of the children with visual impairments needed more time to discern optotypes than one might expect from their reduced visual acuity alone.

The accuracy of the grating-discrimination responses depended on the method used to assess whether the children could resolve the grating. Because the gratings were all well above threshold in the children with normal vision, we expected that these children could easily resolve the grating. However, it turned out that the primary saccade was not always in the correct direction. For about 18% of the children with normal vision, their visual acuity would have been profoundly underestimated if this criterion would have been used to determine the visual acuity. This is not a consequence of using eye tracking during the preferential looking task. Similar results have been found in traditional preferential looking tests. If only the direction of infants’ first fixation was used to determine response accuracy, their visual acuity appeared consistently lower compared to scores based on prolonged target fixation. In the current study, the accuracy scores based on prolonged target fixation also appeared to be more reliable to detect whether a child did successfully locate the target grating. However, some children appeared to make goal-directed saccades, after which they redirected their gaze towards the center of the screen to prepare for the next trial, fixating only briefly on the grating. In those cases, the relative fixation time on target does not adequately score the subject’s visual acuity. The most representative accuracy scores were obtained with a novel scoring method, which combines the two previous scoring criteria. For the children with visual impairments, the highest accuracies were also obtained with this combined...
method. However, their accuracies were significantly lower than the accuracies of the children with normal vision, even though the gratings were in general above visual acuity for both groups. This could indicate that in children with visual impairments other criteria are needed to reliably estimate their visual acuity based on eye tracking data in computerized preferential looking tasks. An alternative could be to use a less strict criterion for detection than 62.5% correct, or to add an expert human observer for these children. Similar recommendations have been made to assess grating detection in infants. Given their low guess rates and high lapse rates the ideal threshold is often <50% correct\textsuperscript{250}. Studies with more children with visual impairments and a wider range of grating frequencies are necessary to estimate the best criteria for this population.

The low success rates of eye tracking in children with visual impairments and the low accuracies of their orienting responses have implications for the use of eye tracking during preferential looking task in clinical settings. However, although it might not be possible to solely rely on eye tracking to assess visual acuity with a preferential looking tasks for all participants, if eye tracking is possible, it does provide valuable additional information, such as saccade latencies. The present study revealed that most children with visual impairments had longer saccade latencies than children with normal vision.
Supplemental Table S6.1. Clinical characteristics of the children with their participant number (Pp), age, group, diagnosis, crowded visual acuity (VA) measured with the Freiburg acuity test, refractive error, the presence of strabismus and nystagmus (+: manifest, +/-: latent, -: absent) and whether eye tracking was possible (+: possible, otherwise the reason why eye tracking failed is given).

<table>
<thead>
<tr>
<th>Pp</th>
<th>Age</th>
<th>Group</th>
<th>Diagnosis</th>
<th>Crowded VA (LogMAR)</th>
<th>Refraction</th>
<th>Strabismus</th>
<th>Nystagmus</th>
<th>Eye Tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>12</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.32</td>
<td>OD: S-4.50 = C-3.25 x 115</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S-4.50 = C-2.50 x 170</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>VI</td>
<td>Albinism</td>
<td>0.54 (uncrowded)</td>
<td>OD: S-0.75 = C-1.50 x 3</td>
<td>+/-</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S-1.75 = C-1.25 x 23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.34</td>
<td>OD: S-8.00 = C-1.0 x 20</td>
<td>-</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S-7.75 = C-0.75 x 142</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>12</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.38</td>
<td>OD: S-0.75 = C-0.75 x 10</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S-0.75 = C-0.75 x 180</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>7</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.57</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>VI</td>
<td>Albinism</td>
<td>0.68</td>
<td>OD: S+1.50 = C-2.00 x 174</td>
<td>+/-</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S+1.50 = C-1.50 x 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>VI</td>
<td>Macular atrophy</td>
<td>0.64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>21</td>
<td>10</td>
<td>VI</td>
<td>Cone-rod dystrophy</td>
<td>1.13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>22</td>
<td>11</td>
<td>CVI</td>
<td>Status after meningitis and cerebritis</td>
<td>1.26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>VI</td>
<td>Hypermetropia</td>
<td>0.38</td>
<td>OD: S+7.75 = C-0.75 x 177</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S+6.50 = C-0.50 x 156</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>8</td>
<td>VI</td>
<td>Albinism</td>
<td>0.28</td>
<td>OD: S+5.00 = C-1.0 x 120</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S+7.25 = C-0.5 x 45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>VI</td>
<td>Albinism</td>
<td>0.66</td>
<td>OD: S+2.50 = C-0.5 x 132</td>
<td>+</td>
<td>+/-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S+30 = C-0.5 x 83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>7</td>
<td>CVI</td>
<td>Motor, cogn. and visual delay (35W, &gt;2500g)</td>
<td>0.41</td>
<td>OD: S+2.75 = C-0.75 x 175</td>
<td>+</td>
<td>-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OS: S+2.50 = C-0.75 x 14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pp</td>
<td>Age</td>
<td>Group</td>
<td>Diagnosis</td>
<td>Crowded VA (LogMAR)</td>
<td>Refraction</td>
<td>Strabismus</td>
<td>Nystagmus</td>
<td>Eye Tracking</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>----------------------------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>27</td>
<td>11</td>
<td>VI</td>
<td>Albinism</td>
<td>0.70</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Technical issues</td>
</tr>
<tr>
<td>28</td>
<td>10</td>
<td>VI</td>
<td>Myopia and retinal scarring</td>
<td>0.82</td>
<td>OD: S-5.00 = C-5.00 x 173 OS: S-6.00 = C-0.25 x 140</td>
<td>+/-</td>
<td>+/-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>29</td>
<td>8</td>
<td>CVI</td>
<td>Optic nerve atrophy, microphthalmus and bilateral occipital infarcts</td>
<td>0.43</td>
<td>OD: S-14.0 OS:-</td>
<td>+</td>
<td>+/-</td>
<td>Long dark eyelashes</td>
</tr>
<tr>
<td>30</td>
<td>7</td>
<td>CVI</td>
<td>Premature and dysmature (26W, 750 g)</td>
<td>0.29</td>
<td>OD: S+4.50 = C-1.50 x 152 OS: S+3.75 = C-0.75 x 36</td>
<td>+</td>
<td>-</td>
<td>Technical issues</td>
</tr>
<tr>
<td>31</td>
<td>7</td>
<td>VI</td>
<td>Cone dysfunction (Bornholm)</td>
<td>0.54</td>
<td>OD: S-7.50 = C-2.00 x 7 OS: S-7.50 = C-2.25 x 166</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>32</td>
<td>6</td>
<td>VI</td>
<td>Coloboma of the iris and retina ODS, and optic nerve OS</td>
<td>0.39</td>
<td>OD: S+1.25 = C-1.50 x 2 OS: S+3.00 = C-1.00 x 10</td>
<td>+</td>
<td>-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>VI</td>
<td>Albinism</td>
<td>0.48</td>
<td>OD: S+1.50 = C-1.75 x 173</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>34</td>
<td>8</td>
<td>CVI</td>
<td>Joubert syndrome</td>
<td>0.34</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>35</td>
<td>9</td>
<td>CVI</td>
<td>Cerebral arteriovenous malformation, resulting in 2 strokes.</td>
<td>-0.15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Long dark eyelashes</td>
</tr>
<tr>
<td>36</td>
<td>9</td>
<td>VI</td>
<td>Aniridia</td>
<td>0.64</td>
<td>OD: S+3.75 = C-1.50 x 2 OS: S+3.25 = C-2.50 x 173</td>
<td>+</td>
<td>+</td>
<td>Aniridia</td>
</tr>
<tr>
<td>37</td>
<td>12</td>
<td>VI</td>
<td>Albinism</td>
<td>0.66</td>
<td>OD: S+2.50 = C-2.50 x 177 OS: S+1.50 = C-2.50 x 176</td>
<td>+</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>38</td>
<td>8</td>
<td>CVI</td>
<td>Premature (25W, 990 g)</td>
<td>0.13</td>
<td>OD: S+1.50 = C-0.50 x 74 OS: S+1.25 = C-0.25 x 27</td>
<td>-</td>
<td>-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>39</td>
<td>11</td>
<td>CVI</td>
<td>White matter damage due to mitochondrial disease and internuclear ophthalmoplegia</td>
<td>0.49</td>
<td>OD: S-0.50 = C-1.00 x 100 OS: S-0.50 = C-1.00 x 84</td>
<td>+</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>40</td>
<td>11</td>
<td>VI</td>
<td>Infantile nystagmus syndrome</td>
<td>0.27</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pp</td>
<td>Age</td>
<td>Group</td>
<td>Diagnosis</td>
<td>Crowded VA (LogMAR)</td>
<td>Refraction</td>
<td>Strabismus</td>
<td>Nystagmus</td>
<td>Eye Tracking</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>------------</td>
<td>------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>41</td>
<td>8</td>
<td>CVI</td>
<td>Perinatal complications, Stickler syndrome</td>
<td>0.02</td>
<td>OD: S-6.75 = C-1.5 x 73 OS: S-6.75 = C-1.5 x 74</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>42</td>
<td>7</td>
<td>CVI</td>
<td>Premature (32W, 2180 g)</td>
<td>0.38</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Could not sit still enough</td>
</tr>
<tr>
<td>43</td>
<td>8</td>
<td>CVI</td>
<td>Premature (32W, 1900g) and perinatal complications</td>
<td>0.09</td>
<td>OD: S+4.0 = C-0.5 x 180 OS: S+3.5</td>
<td>+</td>
<td>-</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>44</td>
<td>10</td>
<td>CVI</td>
<td>Perinatal complications, hemiparesis</td>
<td>0.27</td>
<td>OD: S+3.0 = C-2.25 x 8 OS: S+2.5 = C-1.0 x 25</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>45</td>
<td>7</td>
<td>VI</td>
<td>Congenital stationary night blindness (CSNB)</td>
<td>0.80</td>
<td>OD: S-4.50 = C-1.50 x 5 OS: S-4.25 = C-1.75 x 164</td>
<td>+</td>
<td>+</td>
<td>Problematic glasses</td>
</tr>
<tr>
<td>46</td>
<td>8</td>
<td>CVI</td>
<td>Perinatal complications (36W, 3100 g)</td>
<td>0.49</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>7</td>
<td>CVI</td>
<td>Microcephalus, partial cataract and coloboma of the iris</td>
<td>0.38</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
Summary and General Discussion
The aim of this thesis was, i) to quantify the development of visual processing speed in children with normal vision between five and twelve years old, and, ii) to determine whether children with visual impairment are slower in discerning visual details than children with normal vision. For that purpose, new methods were developed to assess the speed and accuracy of visual processes simultaneously and to quantify oculomotor behavior in these children.

7.1 Summary

Part 1: Symbol discrimination speed in children with and without visual impairments

Many visually-guided tasks require rapid perception of visual details, but how fast children can discern foveal stimuli and how this ability improves with age is still unknown. In Chapter 2, we investigated developmental effects on the speed of visual symbol discrimination in children between 5 and 12 years old with normal vision with a combined Landolt-C-discrimination reaction-time test. Children (n=94) had to indicate, as fast and accurately as possible, the orientation of a Landolt-C symbol (90 trials). Task difficulty was manipulated by varying symbol size (-0.43 to 1.09 LogMAR at 5m). The resulting reaction times were analyzed with a drift-diffusion model. Reaction times on a visual and auditory detection task were measured to assess the contribution of other factors, such as delays in stimulus detection and execution of motor response. In the visual detection task (VDT) the children had to press the mouse button as soon as they saw the visual stimulus (a large “O”), in the auditory detection task (ADT) the children had to press the mouse button as soon as they heard a loud white noise burst. Detection and discrimination were significantly faster in older children. 5-year-olds needed ~440ms for visual detection and ~980ms for discrimination of the largest symbols while 12-year-olds only needed ~250ms and ~500ms for the same tasks. The extra time needed for discrimination compared with detection decreased with age. The decrease in reaction time with increasing optotype size was also age-dependent and indicated an increase in sensitivity with age. Despite the time pressure, acuity thresholds were normal (within the EN ISO-8597 standard). Our data revealed substantial developmental improvements in visual discrimination speed, which suggests that an important optimization takes place in the developing visual system of 5-12 year-old children. Since the speed-acuity test allows for quick and reliable assessment of visual recognition acuity and speed, it may be useful in clinical testing too.
The developmental dissociation between speed and accuracy that is revealed by our data in Chapter 2 implies that visual acuity alone cannot predict how long it takes for a child to see. The combined measurement of recognition acuity and recognition speed is therefore relevant for the assessment of a child’s visual development, and may also be of aid in clinical diagnostics of visual impairment and rehabilitation indications. That is what we assessed in Chapter 3, where we determined whether children with visual impairments are slower than children with normal vision in discerning visual details and assessed if such differences may be explained by their reduced acuity alone. 5-12 year-old children with visual impairments due to retinal/ocular dysfunction (VIo; n=30) or cerebral visual impairment (CVI; n=17) performed the speed-acuity test in which they indicated the orientation of Landolt-C symbols as fast and accurately as possible. The reaction times for symbols ranging between -0.3 and 1.2 LogMAR relative to visual acuity were compared to normative data (Chapter 2). To test whether children were slow in detecting symbols to begin with, we also compared their reaction times on a visual detection task (VDT) to normative data. An auditory detection task (ADT) was used to probe for other, more general deficits. The results revealed that 88% of the children had abnormally long reaction times in the speed-acuity test. This delay was partly explained by their reduced acuity but ~40% of the children still needed more time to discriminate acuity-matched optotypes. Children responded late in the VDT too, especially those with CVI, but this impairment could not fully account for their delayed symbol discrimination. In children with CVI, reaction times in the ADT were affected as much as those in the VDT, suggesting more general sensorimotor problems in CVI. The results from Chapter 3 thus show that children with VIO and CVI are abnormally slow in discerning foveal details. Our findings imply that magnification of materials is often insufficient to compensate for this deficit, partly because detection is already hampered. The speed-acuity test offers valuable insight in the effect of a child’s visual impairment.

Part 2: Development of a high speed stereo eye-tracker

Oculomotor control and visual processing are tightly linked. Therefore, eye movements could have played a role in the reaction times in the speed-acuity test in Chapter 2, and the delays in visual processing in children with visual impairments in Chapter 3. In addition, eye movement recordings can provide valuable diagnostic insights. Therefore, we were interested in the oculomotor behavior of the children with visual impairments. However, commercially available eye trackers were unsuitable, due to the need for extensive individual calibration procedures. Therefore, we decided
to develop our own high-speed stereoscopic eye tracker that relies on a simplified calibration procedure.

Current stereo eye-tracking methods model the cornea as a sphere. However, the human cornea is slightly aspheric. In Chapter 4 we used simulations to investigate how the optics of the aspheric cornea influences the accuracy of stereo eye tracking methods. We demonstrated that the shape of the cornea has a significant influence on the accuracy of stereo eye-tracking methods. Pupil size, gaze direction and head position all influence the accuracy of stereo eye-tracking methods. The gaze reconstruction that uses the center of the pupil and reflections of IR light sources (VP-CC) is more accurate than the conic algebra method, which uses the shape of the pupil-iris boundary. The gaze errors resulting from the VP-CC method remained within about 1.3 deg. The largest gaze errors were found for large translations of the head (6 cm) and large rotations (15 deg) of the eye. This effect is mainly caused by the error in estimating the center of corneal curvature. In conclusion, stereo eye-tracking methods that assume a spherical cornea with one refractive surface can be an option in situations where reliable calibration is not possible. However, more accurate measurements require the use of a more elaborate model of the eye geometry in which the optics of the cornea are better taken into account.

In Chapter 5, we successfully developed and validated a high-speed stereoscopic eye-tracker with two USB 3.0 cameras and two infrared light sources that can track both eyes at ~ 350 Hz for eccentricities of up to 20°. Previous work has shown that with two cameras one can estimate the orientation of the eyes’ optical axis directly. Consequently, only one calibration point is needed to determine the deviation between an eye’s optical and visual axes. A user interface allows for online monitoring and threshold adjustments of the pupil and corneal reflections. We validated this tracker by collecting eye movement data from nine healthy participants and compared these data to eye movement records obtained simultaneously with an established eyetracking system (EyeLink 1000 Plus). The results demonstrated that the two-dimensional accuracy of our portable system is better than 1°, allowing for at least ± 5-cm head motion. Its resolution is better than 0.2° (SD), and its sample-to-sample noise is less than 0.05° (RMS). We concluded that our stereo eyetracker is a valid instrument, especially in settings in which individual calibration is challenging. Finally, it could be beneficial to use the stereo eye tracker in settings in which the relative gaze angles provide sufficient information (e.g., to quantify the amplitude and frequency of a nystagmus).
Chapter 7

Part 3: Saccade latencies in children with and without visual impairments

In Chapter 6, we further explored how visual processing speed develops during childhood and whether visual processing is slower in children with visual impairments. To that end, we determined the latencies of orienting responses during a preferential looking task in children with normal vision and in children with visual impairments between 6 and 12 years old, and assessed the feasibility of scoring grating detection in these populations with video-based eye tracking. Children performed a computerized preferential looking test, while our remote eye tracker (Chapter 5) measured the children’s eye movements. The stimuli consisted of 2×2 grids with three uniform grey fields and one target field containing a black-and-white square wave grating. The grating was presented randomly at one of the four grid locations. The spatial frequencies (1.05, 2.11 and 7.02 cyc/deg) were randomly interleaved, with 10 trials per spatial frequency. Three different methods were used to score the accuracy of the responses: 1. primary saccade ends on target, 2. gaze 50% of the presentation time on target, and 3. a combination of method 1 and 2 (i.e., primary saccade ends on target, or gaze 50% of the presentation time on target). The combined scoring method was most reliable to determine whether children could resolve the gratings. We confirmed that eye tracking in participants with ophthalmological problems can be challenging, because their eyes can be more difficult to track. For 19/34 of children with visual impairments we could not obtain reliable eye tracking data. Children with visual impairments had significantly lower accuracies than children with normal vision with all three scoring methods. In addition, we found that saccade latencies decreased with age and were significantly longer (62±15 ms) in children with visual impairments. We therefore conclude that the use of eye tracking to assess grating detection with a preferential looking task in clinical populations provides valuable additional information, including objective detection measures and developmental delays in saccade latencies.

7.2 Development of visual processing speed

On all reaction-time tasks we found large developmental effects for children with normal vision. Manual reaction times on the detection tasks and the speed-acuity test (Chapter 2) as well as saccade latencies in the preferential looking task (Chapter 6) decrease substantially with age in children between five and twelve years. This is in line with previous findings that children become faster on visual tasks as they grow older (Chapter 1). The question is whether the improved speed of perceiving and
reacting to a simple optotype could explain, at least part of, the increase of reading speed with age\textsuperscript{66,67}, or the decrease with age in the time needed for visual search tasks\textsuperscript{145–147} or matching letters or numbers\textsuperscript{148–150}. Due to the limited available testing time, we could not include more tests to assess whether this is true. Future research is necessary to answer this question.

The developmental effects on the reaction time measures could reflect an improvement and optimization of visual processing, but another explanation could be that non-visual factors become more efficient. Although we kept the motor and cognitive demands as simple as possible, we cannot exclude that motor processes or cognitive development played a role in the decrease in reaction times with age. Motor skills\textsuperscript{251,252}, cognitive processes\textsuperscript{253,254}, and visuomotor skills\textsuperscript{255} all develop during childhood. The effect of non-visual factors would most likely be most prominent in the speed-acuity task, because it requires a more complex cognitive judgment and a more complex motor response. The development of non-visual factors might also account for, at least part of, the decrease with age of the extra time needed for discrimination compared with detection (Chapter 2). The reduction in saccadic latencies with age may reflect optimization of visual processing, or development of several visuomotor and attentional processes, such as shifts of visual attention to the new target, disengaging visual attention from the fixation dot, or translation from sensory to motor coordinates\textsuperscript{161,164}.

Alternatively, faster responses in older children could result from a general decrease in noise levels within the visual system, because in that case lower decision bounds can be adopted while maintaining a similar accuracy. This is supported by the results of the mixed-model analysis on the reaction times on the speed-acuity test, which revealed that the shape of the chronometric curves changed with age (Chapter 2). These shape-changes can only be attributed to an optimization of the visual discrimination process, because they only relate to the size of the visual stimulus. The difference between the reaction time on the largest Landolt-C compared with the reaction time on the smallest Landolt-C was larger for younger children, while their visual acuity did not significantly differ from the visual acuity of older children. This suggest that the younger children applied a higher decision bound (i.e. needed more evidence to reach a decision) to compensate for increased noise levels. Applying the entire drift diffusion model, which includes fits to both the reaction time data and accuracy data simultaneously\textsuperscript{42,62,180}, could potentially provide more insight in whether young children applied higher decision bounds. We have
considered such an approach, but we decided against it, because we do not want to get trapped in an oversimplification of the visual discrimination processes. Drift diffusion models, like the ones in Palmer et al.\textsuperscript{62}, model the trade-off between speed and accuracy by assuming a lawful coupling between reaction time and choice, but this lawful coupling does not hold in general\textsuperscript{185,256}. In clinical settings, the proposed one-to-one coupling between choice behavior and reaction time might be violated even more. Therefore, we did not perform simultaneous fits of the reaction time and accuracy data.

Furthermore, the decrease in reaction time as a function of optotype size was steeper in older children, which demonstrated that the sensitivity of extracting the relevant information from the stimulus was higher for older children. We did not find such an interaction between age and stimulus strength on the saccade latencies in Chapter 6. However, this could be due to the limited number of spatial frequencies we used in the preferential looking task. All gratings were well above visual acuity, while the shape effects on the chronometric curves are most pronounced for stimuli close to visual acuity.

It is almost impossible to differentiate between the attribution of visual, visuomotor, attentional, cognitive, and motor processes to the developmental effects on reaction time and saccade latencies in this thesis, because it is almost impossible to test one of these factors in isolation. Most neuropsychological tests may be designed to test one construct, but require multiple cognitive and perceptuomotor skills for completion\textsuperscript{257–259}. In particular, most test batteries for (fine) motor skills and neuropsychological tests highly depend on visual skills and require attention to perform the task\textsuperscript{258}. Conversely, most visual tests require verbal and/or motor responses, and cognitive aspects such as attention and memory\textsuperscript{259}. We used the detection tests to gain more insight in whether the increase in symbol discrimination speed with age can be attributed to optimization of visual processes, or to more general cognitive processes or processes related to the motor response. The results indicate that the general increase in reaction times with age cannot be uniquely attributed to developmental effects within the visual system. However, at least part of the reaction time decrease does appear to be due to an increase in the speed of the visual discrimination process.

Whether or not the developmental effects are purely related to visual processing, or whether other processes play an important role, might be irrelevant in daily life. In general, we need to respond to, and act upon, stimuli in our environment, and
not just perceive rapid changes in visual information. Many cognitive, visual, and visuomotor functions have to work properly at the same time in order to perform daily-life activities successfully and fast. It has been suggested that this may also be the reason for the UFOV test’s high ecological validity, as performance on the UFOV test depends on many cognitive and visual functions\(^4^7\). It has to be determined whether the preferential looking task correlates with activities in daily life as well.

It has been argued that white matter maturation in the brain may underlie the development of cognitive processes, including visual processing speed\(^1^5^5,2^6^0–2^6^2\). White matter is important for the efficient transmission of neuronal signals\(^2^6^3–2^6^5\) and matures during childhood\(^2^6^6–2^6^8\). White matter integrity is related to the speed of visual processing in adults\(^2^6^9,2^7^0\), and in children\(^1^5^5,2^6^0,2^6^1\). For instance, maturation of white matter connections is correlated with the decrease of simple reaction times with age in children\(^1^5^5,2^6^0\). In addition, in children, white matter maturation correlated with the increased speed of visual search with age\(^2^6^1\), with inspection time in young children\(^2^7^1\), and with the decrease of saccade latencies with age\(^2^6^8\). Furthermore, the latency of visually evoked potentials were correlated with white matter maturation of visual and motor areas of the brain\(^2^6^0\). Thus, it is likely that the increase in visual processing speed with age in Chapter 2 and 6 is in part caused by white matter maturation.

The children included in our studies were between five and twelve years old. The data suggest that at twelve years of age the children did not yet perform at maximal speed, but inclusion of older children and adults is necessary to determine at what age maturation of the reaction times and saccade latencies is finished. Previous research indicated that saccade latencies reach adult levels around 10-15 years of age\(^1^5^8,1^6^2,2^7^2\). In addition, that would allow assessing the shape of the developmental effect of age on reaction times and saccade latencies further. In the age range of our studies, the data could be well described by linear regressions. However, previous research has shown that for a number of visual tasks the developmental course of response time follows a curvilinear trajectory, i.e., rapid changes through childhood and slower rate of change later in development\(^1^5^5,1^5^8,2^7^3,2^7^4\). In addition, we used a cross-sectional design. While this design did provide important information, it would be beneficial to use longitudinal designs to further explore the age effects on the reaction times on the speed-acuity test and on the saccade latencies in the preferential looking task.
7.3 Visual processing speed in children with visual impairments

Thus far there were no studies that assessed whether children with visual impairment are slower in discerning visual details than children with normal vision. We used new tools to assess the speed and accuracy of visual processes simultaneously in these children. The results revealed that many children with visual impairment needed more time on the speed-acuity test (Chapter 3) and had longer saccade latencies on the preferential looking task (Chapter 6) compared to children with normal vision. This was in line with previous studies on visual search, orienting responses, and reading speed. Our results suggest that at least part of the extra time children with visual impairments need on visual tasks may be explained by difficulties in the ability to discriminate symbols fast. Further research is necessary to assess whether the results on the speed-acuity test correlate with the time needed for other visual tasks. In patients with macular degeneration, a significant association between increased stimulus duration threshold for letter recognition and reduced reading speed has been found, which supports the hypothesis that slower letter recognition speed could be one of the factors limiting reading speed. As described above (section 7.2), we cannot exclude that delays in motor or cognitive processes underlie part of the observed delays in symbol discrimination and saccade latencies. It has been shown that children with visual impairments need more time to perform fine motor tasks. Problems with fine motor skills may have contributed to delays on the speed-acuity test, however, reduced speed of visual processing may also account for the extra time to perform fine motor skills. Furthermore, cognitive factors could have played a role as well, especially for the children with CVI. However, the comparison of the reaction times on the speed-acuity test with the reaction times on the detection tasks (Chapter 3) demonstrated that for most children with visual impairments the delays can be largely attributed to visual processes. In addition, the delayed saccade latencies on the preferential looking task (Chapter 6) provides additional evidence that in these children the visual processing speed is indeed reduced.

A large group of children with CVI, i.e. the ones with large motor problems and/or large cognitive developmental delays, could not be included in this study because of the minimum response requirements needed to perform the tasks. Therefore, our results may not generalize to children with more severe CVI. Most likely, their speed of visual processing is more severely affected. It has been suggested that reaction times on detection tasks may serve as an indication of the level of cognitive development. Therefore, detection tasks were used as an indication of general
cognitive and sensorimotor delays. The results on the detection tasks did suggest that general sensorimotor deficits and/or developmental delays played a more prominent role in children with CVI, whereas the impairments in children with visual impairment due to ocular diseases seemed mostly due to delayed visual processing. However, on top of the general sensorimotor delays in children with CVI, visual processing speed was reduced as well.

We included children with a wide variety in ophthalmological diagnoses and underlying causes of CVI. This reflects the heterogeneity of visual impairments that is present in children in western Europe. However, it was difficult to draw conclusions on whether children with certain diagnoses or characteristics have a higher risk of being slower due to the limited number of children with a particular diagnosis. Inclusion of more children with visual impairments is necessary to provide more insight in these questions, but the low prevalence of most ocular diseases makes it complicated to find enough children who are willing to participate.

In the speed-acuity test we only assessed how fast children could discriminate a single Landolt-C of varying sizes. We have not assessed how fast children could discriminate optotypes in a crowded setting. Children with visual impairments often have more problems with crowding and crowding is a predictor of reading speed. During testing it was noticeable that children with VI needed even more time for the crowded version of the FrACT compared to the uncrowded version. Therefore, including a crowded version of the speed-acuity test could provide valuable additional information, as in daily life most visual information consist of stimuli surrounded by contours or objects. Even if there are no differences between the crowded and uncrowded VA of a child, the time needed to discriminate symbols could still differ. It is likely that the differences between children with and without visual impairments would become even more pronounced with a crowded speed-acuity test. This may have important implications for educational settings, as most textbooks and exercises consist of crowded information.

For some of the children with visual impairments, the delays in visual processing could be caused by delayed white matter maturation, or damage to white matter. Previous studies revealed that most children with CVI have damage to both gray and white matter, which could have contributed to longer reaction times on the speed-acuity test and the saccade latencies on the preferential looking task. For children with visual impairment due to ocular diseases, abnormalities of white matter in the visual pathways could also have contributed. Several studies reported
substantial changes in white matter in patients with ophthalmological diseases, including patients with Leber hereditary optic neuropathy, glaucoma, and retinal dystrophies\textsuperscript{277,278}. However, not all ophthalmological diagnoses appear to be related to white matter changes. For instance, in children with albinism no reduction in white matter volume was found\textsuperscript{279}. In addition, other factors could underlie the delays in visual processing in children with visual impairments, such as reduced contrast sensitivity. Previous studies demonstrated that reduced contrast sensitivity impaired performance on vision-based cognitive tests\textsuperscript{59,280}. Furthermore, it has been suggested that fixation instability could be an underlying cause for reduced reading speeds in people with low vision\textsuperscript{281}, and this could have resulted in longer reaction times in the children with visual impairments as well.

Criteria purely based on visual acuity and visual field often fail to capture vision-related problems in daily life\textsuperscript{1,80–82}. The ability to discern foveal details fast may be crucial for a child’s normal participation in school and society, but we did not assess whether slow visual discrimination speed, or longer saccade latencies, are more related to vision-related problems in daily life. Questionnaires of functional vision\textsuperscript{82,282}, or test batteries to assess functional vision exist, to gain insight in the functional visual capacities of people with visual impairments\textsuperscript{81,82,283,284}. These tools could be used to determine the ecological validity of the speed-acuity test in combination with the detection tasks, and preferential looking tasks.

### 7.4 Stereo eye-tracking methods

Several researchers have developed video-based stereo eye trackers\textsuperscript{117–122}. However, the sampling rates of the developed prototypes was only 20-30 Hz\textsuperscript{120–124}, which is insufficient for most eye tracking applications. We successfully developed and validated a high-speed (>350 Hz) stereoscopic eye-tracker (Chapter 5). We used a spherical model for the gaze reconstruction, which resulted in promising results in terms of accuracy and precision compared to an established eye tracker. However, in Chapter 4 we demonstrated that the aspherical properties of the cornea impact the accuracy of stereo eye tracking methods (Chapter 4). Therefore, using an aspherical model for the gaze reconstruction might improve the accuracy and precision even more. At present, there are no aspherical gaze reconstruction algorithms available for stereo eye trackers, but for single-camera eye-trackers an aspherical model significantly improved the gaze reconstruction\textsuperscript{285}. The stereo gaze reconstruction could result in more accurate gaze estimations, reduced noise levels, and higher robustness against head movements with an aspherical eye model. Preliminary
results of gaze reconstructions with an aspherical eye model that we developed indicate that this is most likely the case. This would especially benefit testing clinical populations or (young) children, or in settings with high demands in terms of accuracy and precision.

We used the eye tracker to quantify the saccade latencies in children with and without visual impairments in Chapter 6. The reduced calibration procedure proved to be sufficient and highly facilitated testing children. In general, we could start our experiment within 30 seconds after starting the eye tracking software, instead of starting with an extensive calibration procedure, which would have taken at least a couple of minutes. It was even possible to use the stereo tracker in children with nystagmus. However, detection of the eyes was not always optimal with the stereo eye tracker. For 45% of children with visual impairments, we could not obtain reliable eye movement recordings (Chapter 6). Especially in case of glasses, detection proved to be challenging. Different methods to detect the eyes might result in more robust tracking, but in case of ocular malformations such as aniridia, coloboma of the iris (in which the pupil is not round), or iris transillumination, other methods would most likely fail as well. Alternatively, a chin rest and manual selection of the eyes, or using a sticker to facilitate the eye detection could be helpful additions to the stereo eye tracker.

7.5 Clinical application

We demonstrated that it is feasible to use the speed-acuity test in children with visual impairments and that the results provide valuable insight in the effect of a child’s visual impairment. In addition, we showed that the use of eye tracking to assess grating detection with a preferential looking task in clinical populations does provide valuable additional information, such as developmental delays in saccade latencies and objective detection measures. The capacity to discern visual details fast and accurately is indispensable for a child’s ability to perform visually guided tasks, and thereby essential for participation in school and society. Therefore, including measures of visual processing speed in clinical practice could add valuable criteria to the existing diagnostic tools, by not only assessing whether children can perceive details, but also on how long it takes to perceive and respond to symbols. Especially for children with CVI the diagnostic criteria are limited and there is no consensus on when to give the diagnosis of CVI. The broad spectrum of visual deficits makes it difficult to diagnose CVI on the basis of ophthalmological examination and to characterize and quantify visual problems. Furthermore, the
impact of CVI in education and rehabilitation is less understood than in the case of ocular impairment. The speed-acuity test may be more sensitive in characterizing the visual deficits in children with CVI compared to traditional methods. In addition, the speed-acuity test could provide clear recommendations for rehabilitation and educational settings for children with visual impairments, by providing insight in whether they need more time for visual tasks and whether magnification could compensate for the delays in visual processing.

It has been argued that more automated methods of acuity testing are needed and that technological advances have increased the feasibility of such methods. In this thesis we provide support for this statement and we developed these methods. Computerized tests have several advantages compared to traditional charts: (i) scoring does not rely on an observer, because it is fully automated, (ii) parameters such as luminance or contrast can be adapted, (iii) different stimuli can be used and optotypes or gratings can be presented in random order or at random positions, (iii) results are not limited to a single outcome measure (e.g. visual acuity or contrast sensitivity), reaction times can be included as well. In addition, the children included in our studies often preferred the speed-acuity test compared to the standard visual acuity test, because they could press the buttons themselves. For future research, it would be an option to include a crowded version of the speed-acuity test, or use different optotypes. However, in that case normative values have to be acquired for the adapted test. In addition, we chose to use Landolt-C’s with their opening to the left/right as optotypes, because of the intuitive link with left and right mouse buttons. The same could apply to tumbling E’s. In clinical settings, other optotypes are often preferred, such as letters (EDTRS, or HOTV), or Lea symbols in young children. However, to associate letters or Lea symbols with left/right mouse buttons might be less intuitive because of the more complex, rule-based stimulus-response mapping, and could thus be more mentally demanding.

Similarly, the preferential looking task we used can be adapted as well. For example, by measuring saccade latencies to other stimuli, such as cartoons, motion patterns, or stimuli to assess color perception or contrast sensitivity. Similar stimuli have already been used in combination with eye-tracking in children with and without visual problems, but previous studies used traditional eye-trackers which rely on an elaborate calibration procedure. In addition, combining the stereo eye-tracker with existing versions of digital displays for preferential looking tasks with high spatial resolution, would allow to determine saccade latencies to gratings.
closer to visual acuity.

Although it proved to be difficult to track the eyes of children with visual impairments, eye-tracking could provide valuable additional information in clinical settings. As described in the introduction, assessment of eye movements in clinical practice is in general based on visual inspection\(^{91,99}\). One of the underlying reasons is that the available methods are not suitable for use in patients and/or children. The stereo eyetracker could be the solution for objective eye movement recordings in the clinic, as it proved to be highly suitable for children, even with visual impairments. However, due to ocular abnormalities and/or the presence of glasses, camera-based eye-tracking remains challenging in clinical populations. But if eye tracking is possible, the stereo eyetracker could be applied to quantify the amplitude and frequency of nystagmus, assess fixation stability, and determine properties of saccades, such as saccade latencies or peak velocity, even without the one-point calibration procedure.

### 7.6 Conclusion

It is only fitting that the research reported in this thesis was performed at the Donders Institute and was finished 200 years after the birth of F.C. Donders, because we combined three of his main research interests: eye movements\(^{290,291}\), reaction times\(^{39}\), and visual acuity\(^{8}\).

We revealed important developmental effects on visual processing speed in children between five and twelve years old with new methods to assess the speed and accuracy of visual processes simultaneously. In addition, we showed that a large number of children with visual impairments need more time to discern visual details, even if their reduced visual acuity is taken into account. The ability to discern visual details fast may be essential for a child’s normal participation in school and society, but standard visual acuity tests do not adequately capture this aspect of visual processing. We therefore argue that methods which assess the speed and accuracy of visual processes simultaneously offer valuable insight in the effect of a child’s visual impairment.
References
7. Snellen H. Test-types for the determination of the acuteness of vision. Printed by P.W. van de Weijer; 1862.


40. Luce RD. Response times: Their role in inferring elementary mental organization. Oxford University Press on Demand; 1986.

References


110. Dodge R, Cline TS. The angle velocity of eye movements. Psychological Review. 1901;8(2).


283. Taba IB. Improving eye-gaze tracking accuracy through personalized calibration of a user’s aspherical corneal model. University of British Columbia; 2012.


9

Nederlandse Samenvatting
Dankwoord
Publications and Presentations
Biography
Donders Series
Samenvatting

Een visuele beperking wordt vastgesteld op basis van metingen van de gezichtsscherpte (een maat voor de kleinste details die iemand nog net kan onderscheiden) en het gezichtsveld (het gebied dat het oog zonder te bewegen in één keer kan overzien). Bij deze metingen wordt echter geen rekening gehouden met de tijd die het kost om visuele informatie te kunnen onderscheiden. In het dagelijks leven ontstaan echter wel problemen wanneer iemand te laat reageert op visuele informatie, bijvoorbeeld bij deelname in het verkeer. Kinderen met een normale gezichtsscherpte die op school toch veel tijd nodig hebben om te zien, bijvoorbeeld bij lezen en op het bord kijken, worden verwezen naar instellingen voor slechtzienden, maar volgens de huidige definitie zijn deze kinderen niet slechtziend. Het doel van het onderzoek in dit proefschrift was: i) het in kaart brengen van de ontwikkeling van visuele verwerkingsnelheid bij goedziende kinderen tussen de 5 en 12 jaar, en ii) het vaststellen of kinderen met visuele beperkingen trager zijn in het onderscheiden van visuele details dan goedziende kinderen. Om dat te meten, werden nieuwe methodes ontwikkeld die de snelheid en de nauwkeurigheid van visuele processen gelijkvrij gemeten. Daarnaast zijn er nieuwe methodes ontwikkeld om de oogbewegingen van deze kinderen in kaart te kunnen brengen.

Deel 1: De tijd die het kost om visuele symbolen te onderscheiden bij goedziende kinderen en bij kinderen met visuele beperkingen

Voor kinderen is het van belang om visuele informatie snel genoeg te kunnen onderscheiden, bijvoorbeeld op school en bij andere activiteiten in het dagelijks leven. Maar hoe snel kinderen visuele details kunnen onderscheiden en hoe deze vaardigheid zich ontwikkelt gedurende de kindertijd is nog onbekend. In Hoofdstuk 2 hebben we de ontwikkeling van de visuele verwerkingsnelheid van goedziende kinderen tussen de vijf en twaalf jaar in kaart gebracht. We gebruikten daarvoor een visuele discriminatie-reactietijd test: de speed-acuity test. Dit is een test waarmee zowel de gezichtsscherpte als de tijd die de kinderen nodig hadden voor het onderscheiden van de aangeboden symbolen gelijkvrij bepaald kon worden. Op een computerscherm werden Landolt-C symbolen gepresenteerd van verschillende groottes en de kinderen moesten op 5 meter afstand zo goed en zo snel mogelijk met muisknoppen aangeven of de opening van de C aan de rechterkant, of aan de linkerkant zat. Op die manier konden we de reactietijden bepalen voor symbolen van
verschillende groottes en tevens bepalen wat de kleinste C was die de kinderen nog konden zien. Daarnaast bepaalden we de reactietijden van de kinderen in een visuele detectietaak (VDT) en een auditieve detectietaak (ADT). In de ADT moesten de kinderen zo snel mogelijk op de muisknop drukken bij het horen van een hard geluid en in de VDT moesten ze zo snel mogelijk op de muisknop drukken bij het verschijnen van een grote O op het scherm. Bij oudere kinderen werd dezelfde gezichtsscherpte gemeten als bij de jongere kinderen, maar oudere kinderen hadden veel minder tijd nodig dan jongere kinderen voor het onderscheiden van symbolen van gelijke grootte. Oudere kinderen hadden weliswaar ook minder tijd nodig voor het detecteren van de stimuli in de VDT en de ADT dan jonge kinderen, maar deze reactietijden waren niet alleen korter, ook de leeftijdseffecten waren kleiner dan die in de Landolt-C discriminatie-reactietijd test. Dit geeft aan dat het onderscheidingsvermogen van de oudere kinderen sneller is dan dat van jongere kinderen. Daarnaast was de reactietijd voor alle leeftijden voor grote C’s korter dan voor kleine C’s. Dus naarmate details kleiner worden, duurt het langer voordat kinderen deze details kunnen onderscheiden. Dit grootte-effect was echter het sterkst in de jongste kinderen. Concluderend: de resultaten laten zien dat de visuele verwerkingssnelheid sterk verbeterd naarmate kinderen ouder worden. Daarom zou de test een waardevolle toevoeging kunnen zijn voor klinische toepassingen.

In Hoofdstuk 3 hebben we onderzocht of kinderen met visuele beperkingen tussen de vijf en twaalf jaar trager zijn in het onderscheiden van visuele details dan goedziende kinderen. Verder hebben we bepaald of eventuele vertragingen verklaard kunnen worden door hun verminderde gezichtsscherpte. De groep kinderen met visuele beperkingen bestond uit slechtziende kinderen, bij wie de visuele problemen zijn ontstaan door aandoeningen aan de ogen, en slechtziende kinderen met een cerebrale visusstoornis (CVI), bij wie de visuele problemen een gevolg zijn van een hersenaandoening of hersenletsel. We gebruikten dezelfde speed-acuity test als in Hoofdstuk 2 om gelijktijdig de gezichtsscherpte en de reactietijden te kunnen meten. In totaal reageerden 88% van de kinderen met visuele beperkingen later dan de goedziende kinderen op de speed-acuity test. Deze vertraging werd gedeeltelijk verklaard door hun verminderde gezichtsscherpte, maar 40% van de kinderen met
visuele problemen bleek nog steeds trager dan je op grond van hun verminderde gezichtsscherpte en de normale afname van reactietijden met stimulusgrootte mag verwachten. Dit geeft aan dat voor deze kinderen het vergroten van tekst niet voldoende is om voor de vertraging te compenseren. Daarnaast reageerden kinderen met visuele beperkingen ook later op stimuli in de VDT, maar deze toename in reactietijd kon niet geheel verklaren waarom zij meer tijd nodig hadden voor het onderscheiden van de C-symbolen. Kinderen met CVI waren gemiddeld ook trager op de ADT, wat suggereert dat er bij deze kinderen een meer algemene vertraging aanwezig is. De resultaten laten zien dat de gecombineerde test waardevolle informatie geeft over de gevolgen van visuele aandoeningen.

Deel 2: Ontwikkeling van een high-speed stereoscopische eye tracker

Oogbewegingen en het verwerken van visuele informatie zijn sterk verbonden. Oogbewegingen kunnen dan ook een rol hebben gespeeld in de reactietijden in de gecombineerde test in Hoofdstuk 2 en in de tragere verwerkingssnelheid van de kinderen met visuele beperkingen in Hoofdstuk 3. Om die reden werden de oogbewegingen van de kinderen in kaart gebracht. Beschikbare commerciële eye-trackers waren helaas ongeschikt, aangezien die gebruik maken van een uitgebreide individuele calibratie. Bij zo’n calibratieprocedure wordt de proefpersoon gevraagd om naar bepaalde punten op het scherm te kijken (vaak kleine stippen). Het punt waarnaar de proefpersoon kijkt, wordt dan gekoppeld aan de stand van het oog (bijvoorbeeld door middel van de vorm en positie van de pupil), om tijdens het experiment te kunnen bepalen waar de proefpersoon op het scherm heeft gekeken. Een dergelijke calibratieprocedure is voor goedziende volwassenen geen enkel probleem, maar voor slechtziende kinderen is dit vaak erg lastig. Vooral omdat deze kinderen soms een nystagmus hebben, waarbij de ogen continue onwillekeurig heen en weer bewegen. Daarom werd een high-speed stereoscopische (d.w.z. met 2 camera’s) eye-tracker ontwikkeld, waarbij de positie en de oriëntatie van de ogen direct uit de camerabeelden bepaald kan worden. Daardoor is een uitgebreide individuele calibratieprocedure onnodig.

De huidige methodes om met een stereoscopische eye-tracker de kijkrichting te kunnen bepalen, gaan ervan uit dat het hoornvlies perfect bolvormig is. In werkelijkheid is dat niet helemaal juist, het hoornvlies is een licht afgevlakt bol. In Hoofdstuk 4 werd met behulp van computersimulaties bepaald in hoeverre de vorm van het hoornvlies de betrouwbaarheid van de huidige stereo reconstructiemethodes beïnvloedt. We hebben laten zien dat de vorm van het hoornvlies een belangrijk effect
heeft op de nauwkeurigheid van deze methodes. Verder hangt de nauwkeurigheid van deze methodes ook af van de grootte van de pupil en de kijkrichting en de positie van het oog ten opzichte van de camera’s. De grootste onnauwkeurigheden werden gevonden voor grote kijkhoeken en de grootste verplaatsing van het oog ten opzichte van de camera’s. De reconstructiemethode die het centrum van de pupil en de reflecties van meerdere lichtbronnen op het netvlies gebruikt, bleek nauwkeuriger dan de reconstructiemethode waarbij alleen de vorm van de pupil gebruikt werd. Deze methodes kunnen een optie zijn in situaties waarin het niet mogelijk is om een betrouwbare calibratieprocedure uit te voeren, maar het versimpelde model van het hoornvlies zorgt wel voor een minder nauwkeurige reconstructie van de kijkrichting. Voor nauwkeurigere resultaten is het nodig om een uitgebreider oogmodel te gebruiken, wat beter overeenkomt met de werkelijkheid.

In Hoofdstuk 5 beschrijven we de ontwikkeling van onze high-speed stereoscopische eye-tracker. De eye-tracker bestaat uit twee USB 3.0 camera’s en twee infrarood (IR) lampen. Met de camera’s worden beide ogen gefilmd met een frequentie van ongeveer 350 beelden per seconde. Voorgaand onderzoek heeft aangetoond dat het met twee camera’s en twee IR lampen mogelijk is om de positie en oriëntatie van het oog nauwkeurig te bepalen (maar zie ook Hoofdstuk 4). Met behulp van de eye-tracking software worden de ogen herkend en wordt de positie van de ogen in de videobeelden gevolgd. Vervolgens bepaalt de software de locatie van het centrum van de pupil en de locatie van de reflecties van de IR lampen op het hoornvlies uit de videobeelden. Deze gegevens worden gebruikt om de kijkrichting te bepalen. Verder geeft de software online informatie over de oogbewegingen, wat monitoring tijdens experimenten toestaat. We hebben de prestaties van onze eye-tracker vergeleken met een van de meest nauwkeurige traditionele eye-trackers (EyeLink 1000 Plus) door tegelijkertijd met beide systemen te meten. De resultaten laten zien dat de nauwkeurigheid van de ontwikkelde stereo eye-tracker vergelijkbaar is met de Eyelink (beter dan 1 graad) en daarbij ook hoofdbewegingen van 5 cm toelaat. De precisie is beter dan 0.2 graad en het ruisniveau door opeenvolgende videobeelden is kleiner dan 0.05 graad. We concluderen dat onze stereoscopische high-speed eye-tracker betrouwbaar en precis is en vooral geschikt is in situaties waarin individuele calibratie lastig is. Verder zou de eye-tracker ook een goede optie kunnen zijn in situaties waarin alleen de veranderingen in kijkrichting belangrijk zijn, bijvoorbeeld bij het bepalen van de frequentie en amplitude van een nystagmus.
Deel 3: Saccade latenties van goedziende kinderen en kinderen met visuele beperkingen

In Hoofdstuk 6 hebben we verder onderzocht hoe de visuele verwerkingsnelheid zich ontwikkelt bij kinderen en daarnaast of kinderen met visuele beperkingen meer tijd nodig hebben voor het onderscheiden van visuele informatie. Daartoe hebben we gebruik gemaakt van een “preferential looking” test. “Preferential looking” tests zijn gebaseerd op het principe dat jonge kinderen (en zelfs baby’s) liever kijken naar een streeppatroon dan naar een egaal grijs vlak. Door streeppatronen met verschillende streepdiktes te gebruiken, kan op basis van de oog- en hoofdbewegingen van het kind de gezichtsscherpte bepaald worden. Als een kind de streepjes niet meer kan zien dan zal het kind niet meer direct naar het streeppatroon kijken, aangezien dat er voor hem/haar dan ook uitziet als een grijs vlak. De kinderen moesten aan het begin van elke trial kijken naar een fixatiestip midden op het scherm. Daarna werden er 4 vlakken (2 x 2 grid) gepresenteerd op het scherm, 3 grijze vlakken en 1 streepjespatroon waarvan de streepdikte kon variëren. Met behulp van de eye-tracker uit Hoofdstuk 5 werden de oogbewegingen van de kinderen gemeten, waarbij we ons met name hebben gericht op de saccades. Een saccade, of oogspont, is een snelle verandering van de kijkrichting. De saccadelatentie is de tijd tussen het presenteren van het streeppatroon en het begin van de saccade naar dat patroon toe, dus hoe lang het duurt voordat kinderen hun blik verplaatsten van het midden van het scherm naar het streepjespatroon. Voor streeppatronen met verschillende streepdiktes werden de saccadelatenties bepaald. De resultaten lieten zien dat oudere, goedziende kinderen kortere saccadelatenties hadden dan de jongere goedziende kinderen. Daarnaast waren de saccadelatenties van de kinderen met visuele beperkingen significant langer dan die van de goedziende kinderen. Daarnaast hebben we bekeken wat de meest betrouwbare manier is om op basis van de oogbewegingen te bepalen of de kinderen de streepjespatronen nog konden zien. De meest betrouwbare methode bestond uit de combinatie van twee criteria: 1. de eerste saccade eindigt op het streepjespatroon, of 2. de blik was gedurende 50% van de presentatietijd gericht op het streepjespatroon. Verder bleek dat het gebruik van een eye-tracker bij kinderen met visuele beperkingen lastig is, met name omdat het bij deze kinderen moeilijk was om de ogen automatisch te herkennen, bijvoorbeeld door brillen met erg dikke glazen of afwijkingen aan de pupil of iris. Desondanks concluderen we dat, als het detecteren van de ogen lukt, eye tracking een waardevolle toevoeging is in “preferential looking” tests. Niet alleen biedt
eye tracking de mogelijkheid om objectiever te bepalen of kinderen de streepjes nog kunnen zien, maar het geeft ook informatie over eventuele vertragingen van saccades.

Conclusie

We hebben duidelijke leeftijdseffecten laten zien op de visuele verwerkingsnelheid van kinderen tussen de vijf en twaalf jaar, met nieuwe methodes die gelijktijdig zowel de snelheid als de nauwkeurigheid van visuele processen kunnen bepalen. Daarnaast hebben we aangetoond dat een groot deel van de kinderen met visuele beperkingen meer tijd nodig heeft om visuele details te onderscheiden, zelfs als er gecorrigeerd wordt voor hun verminderde gezichtsscherpte.
Dankwoord

Eindelijk is het zover, mijn proefschrift is af! Promoveren kan een aardige achtbaan zijn en de afgelopen jaren zaten dan ook vol met hoogte- en dieptepunten. Het was me nooit gelukt om tot dit resultaat te komen zonder de steun van veel mensen. Daarnaast zijn er veel mensen die ervoor gezorgd hebben dat het een fantastische tijd was. Ik ga mijn best doen om iedereen hierbij te bedanken en hoop dat ik niemand vergeet.

Allereerst wil ik graag alle kinderen bedanken die mee hebben gedaan aan de studies in dit boekje. Nadat studenten bij de eerste pilots klaagden dat de testen te saai waren, wisten we niet zeker of het zou lukken om voldoende data te verzamelen. Gelukkig bleek dat we daar niet bang voor hoefden te zijn, want alle kinderen waren ontzettend gemotiveerd en vonden het geen probleem om in totaal meer dan 250 keer aan te geven aan welke kant de opening van de “C” stond. Ik bedank uiteraard ook alle ouders dat hun kinderen mee mochten doen en de ouders van de slechtziende kinderen dat ze naar Zeist of Nijmegen wilden komen voor de metingen. De metingen bij de goedziende kinderen waren op hun eigen basisschool. Dank aan de drie basisscholen voor de samenwerking en de hartelijke ontvangst. Het was een leuke afwisseling om weer rond te lopen op een basisschool en het was heel fijn dat de docenten het prima vonden dat ik de kinderen uit de les haalde. Verder ben ik ook dank verschuldigd aan de deelnemers van de validatiestudie van de eye tracker. Niet alleen was dat saai, maar ik heb ze ook in een biteboard gezet en dat blijft oncomfortabel (al denk ik dat sommigen al lang blij waren dat ze geen coils in hoefden).

Jeroen, bedankt voor de fijne samenwerking. Ik had me geen betere copromotor kunnen wensen. Je gaf me veel vrijheid en had zelfs vertrouwen in me toen ik bedacht had dat ik wel zelf een eye tracker kon gaan bouwen. Als het nodig was dan kon ik altijd bij je terecht en dan nam je ook alle tijd. De lange discussies over analyses, de eye tracker en papers waren niet alleen productief, maar ook vaak inspirerend. We waren het niet altijd eens, maar we kwamen er altijd uit. Ik waardeer je kritische blik en je betrokkenheid en ik heb veel van je geleerd. Ik hoop dat we de samenwerking kunnen blijven voortzetten.
Chapter 9

Nienke, dankjewel voor je betrokkenheid en je klinische blik. Je vertaalt problemen die je in de oogheelkundige praktijk ziet naar onderzoek en probeert de resultaten ook weer terug te vertalen naar de praktijk. Daar heb ik veel bewondering voor. Ik heb veel geleerd van de woensdagen in Zeist. Verder waren de gesprekken tijdens de fietstochtjes naar het station altijd een gezellige afsluiting van de dag.

John, ik ben blij dat je mijn promotor bent. Ook jij gaf me alle ruimte, maar je gaf ook vanaf het begin aan dat ik altijd bij je terecht kon als ik hulp nodig had. En bedankt voor de mooie kans om als Postdoc aan de slag te kunnen op het Orient project. Ik hoop dat we de komende tijd mooie resultaten kunnen laten zien met behulp van de vestibulaire stoel.

Anna, Annemiek en Renée, ik ben heel blij met jullie als paranimfen. Anna, je was een geweldige collega en bent zelfs een nog betere vriendin. Je was een enorme steun tijdens mijn promotie. Ik kon altijd bij je terecht voor een goed gesprek of gewoon om gezellig bij te kletsen en daarnaast hebben we ook zeker de successen gevieren. Ik denk met heel veel plezier terug aan de vele koppen thee, de gezellige lunches, de borrels, etentjes, feestjes, en zeker ook de lol op de congressen in Barcelona en Florida. Annemiek, we kennen elkaar inmiddels al meer dan 10 jaar. Het begon als studiegenootjes bij BW in Groningen en ik vind het erg leuk dat je nu ook in Nijmegen werkt. We hebben daardoor nu de traditie van de vrijdagmiddagborrels in de Toeter voortgezet met biertjes bij Frowijn of in de Aesculaaf. Ik kan altijd bij je terecht voor een goed gesprek of gewoon om gezellig bij te kletsen en ik ben heel erg blij met onze vriendschap. En het blijft grappig dat sommige mensen het verwarrend vinden dat er twee blonde BW’ers uit Groningen zijn die Annemiek heten 😂. Renée, je hebt niet alleen erg veel voor het project betekend, maar zeker ook voor mij persoonlijk. Of het nou praktische hulp was als er weer wat mis was gegaan met een bestelling, een hart onder de riem als ik er doorheen zat, of even bijkletsen op kantoor, je stond altijd voor me klaar. Daarnaast was het altijd fijn om je aanstekelijke lach over de hele afdeling te kunnen horen. Ik heb je met name beter leren kennen door het samen organiseren (met Piray) van de CNS-borrels en de CNS-bbq en ik vind het erg fijn dat er een mooie vriendschap uit ontstaan is.

CNS collega’s; Anna, Karlijn, Joris, Leslie, Bert, Jeroen, Bianca, Renée, Erna, Inge, Ellen, Sophie, David, Douwe, Dirk, Arno, Piray, Shaha, Giuseppe, Elisavet, Piet, Yvet, Peter en Deborah, het was heel fijn om het werk even te onderbreken voor de lunch, een uitgebreide theepauze, een praatje op de gang of bij de koffieautomaat, of
een biertje tijdens de CNS-borrels of CNS-bbq. Al waren de onderwerpen van de
gesprekken met name tijdens de lunch soms erg vreemd, het was altijd erg gezellig.
En ik heb er genoeg nieuwe nutteloze feitjes over de meest uiteenlopende topics
opgedaan.

Onderzoek doen is niet mogelijk zonder de ondersteuning van veel mensen. Allereerst
dank aan Renée, Erna, Inge en Ellen. Ik kon altijd bij jullie terecht met vragen en jullie
hebben veel werk uit handen genomen. David, bedankt voor alles wat je voor mij
en het project heb betekend. Günter, Stijn en Ruurd, bedankt voor de technische
ondersteuning. Ook veel dank aan alle orthoptistes van Bartiméus, vooral aan Florine
en Marjoke. Ik heb erg veel van jullie geleerd op het gebied van visueel functie
onderzoek en het was heel fijn dat ik altijd mee mocht kijken en vragen kon stellen.
Marjoke, daarnaast ook bedankt voor de hulp met de inclusie. Ook bedank ik graag
iedereen van Vereniging Bartiméus Sonneheerdt, voor het ter beschikking stellen
van jullie vergaderzaal als meetlocatie en vooral ook voor de hartelijke ontvangst en
belangstelling in het onderzoek. De woensdagen in het koetshuis waren dankzij jullie
heel plezierig.

Tijdens mijn onderzoek hebben verschillende studenten me geholpen en die wil ik
hierbij dan ook allemaal bedanken. Twee van hen wil ik hier met name noemen.
Stephanie, je legde de basis voor het valideren van de stereo tracker en daarnaast
heb je geholpen met de metingen op scholen en bij de slechtziende kinderen. Het
was heel fijn om op je te kunnen rekenen en daarnaast was het ook altijd gezellig.
Suzanne, je deed een vrijwillige stage om meer ervaring op te doen met onderzoek.
Ook al waren de meeste van je werkzaamheden vrij saai (zeker het calibreren van de
camera’s door eindeloos in camera images de hoekjes van het schaakbordpatroon te
selecteren), je was altijd enthousiast en werkte heel hard.

Frank, Gert-Jan en Harjo, jullie hebben me begeleid tijdens mijn eerste
onderzoeksstages in Groningen. Door jullie werd ik enthousiast over het doen van
onderzoek en ik denk dat ik niet zover was gekomen zonder jullie. Bedankt daarvoor
en ik vind het erg leuk dat we contact zijn blijven houden. Hetzelfde geldt ook voor
Floor, al was ik in Groningen geen student van je maar leerden we elkaar kennen
tijdens de borrels in de Toeter. Bedankt, ook aan Willem uiteraard, voor de gezellige
bezoekjes aan Essex, waaronder tijdens jullie verhuizing, en de etentjes en borrels
als jullie in Nederland zijn.
Peter, Deborah, Maud, Martijn, Jeroen, Luc en Lisanne, bedankt voor de fantastische wintersporten, gezellige borrels in de Aesculaaf, uitgebreide barbecues en etentjes, anti-D&D avondjes, geniale 30-seconds herinneringen, en veel lol tijdens de oudejaarsavondens.

Inge en Anne-Wil, onze vriendschap begon op het introkamp van BW. We zien elkaar door de afstand nu wat minder vaak, maar ik ben erg blij dat we elkaar nog regelmatig zien. Of we nu gaan suppen, trampolinespringen, bootje varen in Gent, genieten van een high-tea in Utrecht, een slechte film kijken tijdens een bezoekje aan Groningen, of gewoon ergens gaan eten, het is altijd weer als vanouds gezellig.

Annemarie, we kennen elkaar inmiddels al 18 jaar en ik vind het heel fijn dat ik ook mijn promotietijd met je heb kunnen delen. Het bijpraten tijdens de etentjes in Utrecht, Arnhem, Veenendaal, of welke plaats op dat moment ergens in het midden ligt, is altijd gezellig. Ik bewonder je doorzettingsvermogen en ik weet zeker dat je jouw plek gaat vinden.

Piray, what a great time we had. We connected instantly and only a couple of minutes after we first met each other at the department we already came up with the idea to organize the social CNS-drinks. Luckily David supported us and with help of Renee we could actually make it happen. It was the start of lots of fun, laughter, and also serious conversations about the future. Thank so much for your support, the CNS-drinks and barbecues (with the mojito-hammer), all the dinners and parties, and also the time we spent in NYC with you. I will especially remember the “1 drankje doen”-night with Tessel, which obviously did not end after 1 drink 😂.

Shaha, we had so much fun together. Sometimes just over lunch in the canteen, other times when we ended up in town after the Friday-afternoon drinks, or during a lovely dinner with your parents at your place. It is always nice to catch up and spend time together. Thank you for all the nice memories. And I am really glad that you now have your nice house and lovely little family.

Elisavet en Dick, dank voor alle gezelligheid tijdens borrels in de Laaf, feestjes in Den Bosch en Nijmegen, de dansjes tijdens de vierdaagse en het leuke weekend in Tübingen. En natuurlijk ook voor het organiseren van je afscheidsetentje, anders had mijn leven er waarschijnlijk nu anders uitgezien.
Giuseppe and Giorgia, every time we spot a nice Russian Imperial Stout on the menu we think about you guys. It is nice to share our interest in good beers, but even better to share time together. Thank you for the good conversations (not all about beer) and the fun we have together.

Het is heel fijn om thuis te komen en te weten dat je terecht kan bij lieve mensen in de buurt. Tessel, bedankt voor de fantastische tijd aan de Javastraat, ik had me geen betere huisgenoot kunnen wensen. En Juul, onze fijne buuf, dankjewel voor de gezelligheid tijdens de avondjes Catan en de middagjes F1.

I would also like to thank all the “Social Donderians”; Ruud, Isabella, Linda, Erika, Charlotte, Simon-Jan, Laura and Andrea, Natalia, Remco and Ricarda, Katharina, Andrea, Tim, Charl, Natasha, Izabela, Sybrine, Fenny, Thomas, Sandra and Giacomo, and also the “Rehab” people; Milou, Renee, Bas, Frank, Roland, Lotte, Laura, Mariska, Teo, Vera, Jolanda, Marian, Mike, and Judith, for the Friday afternoon drinks at the Aesculaaf or Cultuurcafé, the fun during the Donders Discussions (and the after parties) and the Batavierenrace, the good conversations, and the great time during conferences. In addition, thanks to everyone who was involved with the organization of the Donders Discussions in 2014, it was great fun. Furthermore, it was really nice to be part of the DCN PhD council and to work together with the other PhD councils within the Donders. It did give some insight into the structure and organization of the Radboudumc and the Donders Institute and I enjoyed it a lot. Thanks to all the council members and to Nathalie, Nina, Femke and Tanja for the support.

My current colleagues at biophysics; Hendrik, Loes, Arjen, Elisabeth, Yagmur, Ying, Hans, Martijn, Lei, Katharina, Jesse, Guus, Seba, Snandan, Jasper, Onno, Marc, Stijn, Ruurd, Günter, Judith, Marie-Louise, Richard, and John, thank you all for the warm welcome and the nice atmosphere at the department. It was nice to share the end of my PhD time with you.

Christophe, Angélique, Clélia et Madalen, ça fait 10 ans que je suis venue travailler au camping pour la première fois. Dans les années suivantes vous êtes vraiment devenu ma « deuxième famille ». Je chéris les souvenirs de toutes les bonnes moments passés avec vous, d’abord aux Rosiers et maintenant dans la région de Vannes. Un grand merci pour tous les apéros, les diners, les fous rires, les balades en bateau, des conversations profonds, et les batailles d’eau à table. Je vous embrasse bien fort.
Chapter 9

Mijn schoonfamilie, Evert, Marlies, Sietse, Elizah, Patriek, Dennis, Yvonne, Barbara, Piet en Meike, bedankt voor alle gezelligheid, ik voelde me gelijk welkom bij jullie.

En dan natuurlijk mijn eigen familie. Lieve papa en mama, bedankt voor de onvoorwaardelijke steun. Jullie hebben me altijd aangemoedigd om vooral iets te gaan studeren wat ik leuk vond en werk te vinden waar ik echt enthousiast over werd. Dat is heel belangrijk voor me geweest, anders was ik waarschijnlijk niet op deze plek terecht gekomen. Verder vind ik het heel fijn dat jullie altijd belangstelling tonen voor alles waar ik mee bezig ben, op onderzoeksgebied (als is dat niet altijd makkelijk), maar ook op persoonlijk vlak. Ik hou van jullie en hoop dat ik alles nog heel lang met jullie mag delen. Mijn zusjes, Marlies en Lisanne, bedankt voor alle leuke herinneringen aan o.a. concertbezoekjes, vakanties in Frankrijk (met pontjes, Hans de Gans, lekker eten en kasteelbezoekjes), verkleedpartijen in Zuid-Scharwoude en tijdens het namaken van de oude foto’s, het dagje Apenheul, en de logeerpartijtjes in Groningen en Nijmegen. En Marlies, heel erg bedankt voor het maken van de mooie voorkant van dit boekje! Mike en Rinse, ik vind het erg leuk jullie nu ook deel uitmaken van ons gezin. Bedankt voor de gezelligheid tijdens de weekendjes op Texel, in Hoenderloo, tijdens verjaardagen of tijdens het bezoekje aan de dierentuin.

Lieve Geert, ik ben zo blij dat we elkaar tegen zijn gekomen. Je staat altijd voor me klaar en we hebben al zoveel mooie momenten beleefd, bijvoorbeeld tijdens de geweldige reis naar de VS, de vakanties naar Engeland en Frankrijk, de weekendjes weg en de succesvolle zoektocht naar een huis voor ons samen. Al geniet ik net zoveel van de kleine dingen, zoals samen een biertje drinken in de avondzon op het balkon of samen koken met goede muziek op de achtergrond. Daarnaast heb ik door jou zoveel lieve mensen leren kennen. Het is een feest om met jou het leven te vieren en ik kijk uit naar een mooie toekomst samen. Dankjewel voor alles. Ik hou heel veel van jou!
Publications and Presentations

**Full papers:**


**Conference presentations:**


• Barsingerhorn AD, Boonstra FN & Goossens J. *Normative values for a speed-acuity test to determine delays in visual processing speed* (Oral presentation), Perception Day, Radboud Nijmegen, December 2, 2016.


**Award:**

• *Keith Rayner Memorial Award* for best presentation at the European Conference on Eye Movements 2017
Biography

Annemiek Dorien Barsingerhorn was born on June 9, 1988 in Delfzijl, the Netherlands. She received her secondary education at the Ommelander College in Appingedam. In September 2006, she started with the Bachelor’s program in Human Movement Sciences at the University of Groningen, which was followed by the Master’s program in Human Movement Sciences at the same university. During her Bachelor and Master’s degree she specialized in Perception and Action research, with internships on decision making in volleyball under the supervision of dr. Gert-Jan Pepping, dr. Harjo de Poel and dr. Frank Zaal. In May 2013, she started working as a PhD student at the department of Cognitive Neuroscience of the Radboud University Medical Centre in Nijmegen under the supervision of dr. Jeroen Goossens, dr. Nienke Boonstra and prof. dr. John van Opstal on the topic of visual processing speed in children with normal vision and children with visual impairments. In addition, she developed an interest in eye tracking methodology, which resulted in the development of a stereo eye tracker. The results of her research are described in this thesis and were presented at different national and international conferences. Furthermore, she was part of the organizing committee of the Donders Discussions in 2014 and she was chair of the PhD Council of the Donders Center for Neuroscience. Currently, Annemiek continues her research as a Postdoc at the department of Biophysics of the Radboud University, on multisensory integration and eye-head coordination.
Donders Graduate School for Cognitive Neuroscience

For a successful research Institute, it is vital to train the next generation of young scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School for Cognitive Neuroscience (DGCN), which was officially recognised as a national graduate school in 2009. The Graduate School covers training at both Master’s and PhD level and provides an excellent educational context fully aligned with the research programme of the Donders Institute.

The school successfully attracts highly talented national and international students in biology, physics, psycholinguistics, psychology, behavioral science, medicine and related disciplines. Selective admission and assessment centers guarantee the enrolment of the best and most motivated students.

The DGCN tracks the career of PhD graduates carefully. More than 50% of PhD alumni show a continuation in academia with postdoc positions at top institutes worldwide, e.g. Stanford University, University of Oxford, University of Cambridge, UCL London, MPI Leipzig, Hanyang University in South Korea, NTNU Norway, University of Illinois, North Western University, Northeastern University in Boston, ETH Zürich, University of Vienna etc.. Positions outside academia spread among the following sectors: specialists in a medical environment, mainly in genetics, geriatrics, psychiatry and neurology. Specialists in a psychological environment, e.g. as specialist in neuropsychology, psychological diagnostics or therapy. Positions in higher education as coordinators or lecturers. A smaller percentage enters business as research consultants, analysts or head of research and development. Fewer graduates stay in a research environment as lab coordinators, technical support or policy advisors.

Upcoming possibilities are positions in the IT sector and management position in pharmaceutical industry. In general, the PhDs graduates almost invariably continue with high-quality positions that play an important role in our knowledge economy.

For more information on the DGCN as well as past and upcoming defenses please visit:

http://www.ru.nl/donders/graduate-school/phd/