TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

What drives you - vision or cognition?

Exploration and validation of new methods to test real-world visual functioning

Iris Tigchelaar

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1907 | MEDICA – ODONTOLOGICA | TURKU 2025

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

WHAT DRIVES YOU – VISION OR COGNITION?

Exploration and validation of new methods to test real-world visual functioning

Iris Tigchelaar

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1907 | MEDICA – ODONTOLOGICA | TURKU 2025

University of Turku

Faculty of Medicine
Department of Clinical Medicine
Ophthalmology
Visual Neuroscience
Doctoral Programme in Clinical Research (DPCR)

Supervised by

Dr. Markku Leinonen, M.D., Ph.D.
Ocusweep
Turku, Finland

Professor Eija Vesti
Department of Ophthalmology
Turku University Hospital
Turku, Finland

Reviewed by

Dr. Tuomo Kujala, Ph.D.
University of Jyväskylä
Jyväskylä, Finland

Docent Ville Saarela, M.D., Ph.D.
Oulu University Hospital
University of Oulu
Oulu, Finland

Opponent

Docent Juho Wedenoja, M.D., Ph.D.
University of Helsinki
Helsinki, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Cover Image: I. Tigchelaar & Bing Image Creator (bing.com/create)

ISBN 978-952-02-0332-0 (PRINT)
ISBN 978-952-02-0333-7 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku Finland, 2025

To Hugo and Lasse

UNIVERSITY OF TURKU
Faculty of Medicine
Department of Ophthalmology
Ophthalmology
IRIS TIGCHELAAR: What drives you – vision or cognition. Exploration and validation of new methods to test real-world visual functioning.
Doctoral Dissertation, 218 pp.
Doctoral Programme in Clinical Research (DPCR)
November 2025

ABSTRACT

Current measures for assessing vision, such as charts and visual field testing, do not correlate well with how patients experience daily life tasks such as driving. However, the regulations on fitness to drive in glaucoma patients are still relying on those types of measurements.

The main objective of this work was to assess new ways of measuring vision related to real world tasks, with a focus on driving in glaucoma patients. First, a literature review showed that the relationship between driving and visual and cognitive impairment is mediated by several factors, including degree of impairment, compensation, and research design. Then, the effect of visual field defects on neuropsychological testing was studied in a glaucoma group and a healthy participant group. We found that caution is advised when making conclusions about cognition when glaucomatous visual field defects could influence performance, for example in the Trail Making Test A. This study continued with driving in a driving simulator, in which glaucoma patients did not perform worse than the healthy group. Finally, four new functional vision tests were evaluated, including a test for basic reaction time, face discrimination, visual search and a visual field test using a built-in webcam that monitors compliance. All of these four new ways of testing functional vision were found to be reliable and stable measures of functional vision.

These studies show that the relationship between visual impairment and real-world functioning is complex and mediated by many different factors. Therefore, more tests are needed to evaluate real-world functioning of glaucoma patients instead of isolated visual acuity and visual field tests. Tests outside of an ophthalmological setting on portable devices that take into account vision and cognition could be used to help explain the relationship between visual impairment and real-world performance.

KEYWORDS: cognition, driving, glaucoma, visual field, vision

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Kliininen laitos
Silmätautioppi
IRIS TIGCHELAAR: Mikä meitä ohjaa – näkö vai kognitio? Uusien menetelmien soveltuvuus toiminnallisen näkökyvyn testaamiseen
Väitöskirja, 218 s.
Doctoral Programme in Clinical Research (DPCR)
Marraskuu 2025

ABSTRAKTI

Nykyiset näkökyvyn arviointimenetelmät, kuten näöntarkkuus- ja näkökenttätestit eivät kerro miten hyvin potilaat suoriutuvat näkönsä turvin arkipäivän tilanteista, kuten autoilusta. Glaukoomapotilaiden ajokelpoisuutta koskevat säädökset perustuvat kuitenkin edelleen tämän tyyppisiin mittauksiin.

Tämän työn päätavoitteena oli arvioida uusia tapoja mitata käytännön tilanteisiin liittyvää toiminnallista näkökykyä ja erityisesti glaukoomapotilaiden ajokykyä. Ensinnäkin kirjallisuuskatsaus osoitti, että ajokyvyn sekä näkökyvyn ja kognitiivisen heikkenemisen välinen yhteys on riippuvainen useista tekijöistä, kuten heikkenemisen määrästä, vian kompensoimisesta ja testausasetelmasta. Ensin tutkittiin näkökenttäpuutosten vaikutusta neuropsykologisten testausten tuloksiin glaukoomaryhmässä ja terveiden osallistujien ryhmässä. Havaitsimme, että on syytä olla varovainen tehtäessä johtopäätöksiä kognitiosta, silloin kun glaukoomasta johtuvat näkökenttäpuutokset voivat vaikuttaa testisuoritukseen, esimerkiksi tmt-a-testissä. Tutkimusta jatkettiin ajamisella ajosimulaattorilla, jossa glaukoomapotilaat eivät suoriutuneet huonommin kuin terveiden ryhmä. Lopuksi arvioitiin neljää uutta toiminnallista näkötestiä, mukaan lukien perusreaktioajan mittaus, kasvojen tunnistus, visuaalinen haku ja näkökenttätesti, jossa käytettiin sisäänrakennettua web-kameraa valvomaan testisuoritusta. Kaikki nämä neljä uutta toiminnallisen näkökyvyn testiä osoittautuivat luotettaviksi ja vakaiksi toiminnallisen näkökyvyn mittareiksi.

Nämä tutkimukset osoittavat, että näkövian ja reaalimaailman toimintakyvyn välinen suhde on monimutkainen ja että siihen vaikuttavat monet eri tekijät. Siksi tarvitaan lisää testejä glaukoomapotilaiden todellisen toimintakyvyn arvioimiseksi pelkkien näöntarkkuus- ja näkökenttätestien sijaan. Näkövammaisuuden ja käytännön toimintakyvyn välistä suhdetta voitaisiin selvittää myös kannettavilla laitteilla tehtävillä testeillä, joissa otetaan huomioon sekä näkökyky että kognitio.

AVAINSANAT: ajaminen, glaukooma, kognitio, näkökenttä, näkökyky.

UNIVERSITEIT VAN TURKU
Medische Faculteit
Afdeling Oogheelkunde
Oogheelkunde
IRIS TIGCHELAAR: Wat beweegt ons – visie of cognitie. Verkenning en validatie van nieuwe methoden om het zien gerelateerd aan het dagelijks leven te meten.
Proefschrift, 218 p.
Doctoral Programme in Clinical Research (DPCR)
November 2025

ABSTRACT

Om het zicht van patiënten met oogheelkundige aandoeningen te testen worden onder andere letterkaarten of visuele veld tests gebruikt. Deze metingen correleren echter niet goed met hoe patiënten verschillende taken in het dagelijks leven ervaren.

Het doel van dit proefschrift was het onderzoeken van nieuwe manieren om het zien te meten op een manier die rekening houdt met het dagelijks leven. Het literatuuronderzoek liet ten eerste zien dat de relatie tussen autorijden en visuele en cognitieve beperkingen beïnvloed wordt door vele factoren, zoals de ernst van de beperking, compensatie strategieën en de onderzoeksopzet. Verder werd het effect van visuele velddefecten op neuropsychologische tests onderzocht waaruit bleek dat we voorzichtig moeten zijn met conclusies trekken over cognitie wanneer visuele velddefecten invloed kunnen hebben gehad op de scores, bijvoorbeeld in de Trail Making Test A. Vervolgens reden de deelnemers in de rijsimulator, waarin de glaucoom patiënten niet slechter reden dan de gezonde groep. Als laatste werden er vier nieuwe visuele tests onderzocht, waaronder een reactietijdtest, een gezichtsherkenning taak, een zoektaak en een visuele veld test die gebruikt maakt van een webcam om naleving van de regels van het onderzoek te meten. Deze vier nieuwe tests bleken betrouwbare en stabiele tests te zijn.

Deze onderzoeken laten zien dat de relatie tussen visuele beperkingen en functioneren in het dagelijks leven complex is en wordt beïnvloed door vele factoren. Er zijn meer van deze nieuwe soort visuele tests nodig, in plaats van geïsoleerde gezichtsscherpte tests en visuele veld tests. Onderzoeken buiten een oogheelkundige setting met draagbare apparatuur die zowel het zien als cognitie meten kunnen helpen om de relatie tussen visuele beperkingen en dagelijks functioneren te verduidelijken.

SLEUTELWOORDEN: autorijden, cognitie, glaucoom, realistische tests, visueel veld, zien.

Table of Contents

Abbreviations

ANOVA	Analysis of Variance
BVRT	Benton Visual Retention Test
CBR	Centraal Bureau voor Rijvaardigheidsbewijzen (Dutch agency for driving licenses)
CI	Confidence Interval
CRT	Choice Reaction Time
dB	Decibel
DLS	Differential Light Sensitivity
DSM	Diagnostic and Statistical Manual of Mental Disorders
DST	Digit Span Test
EU	European Union
FACS	Facial Action Coding System
FAS	Functional Ability Score
HFA	Humphrey Field Analyzer
Hz	Hertz
ICADTS III	The International Council on Alcohol, Drugs, and Traffic Safety category 3 (potentially dangerous drugs)
IPS	In-plane switching
IQR	Inter Quartile Range
MCI	Mild Cognitive Impairment
MD	Mean Deviation
MGS	Memory Guided Saccades
MMSE	Mini Mental State Examination
mNCD	Mild Neurocognitive Disorder
MoCA	Montreal Cognitive Assessment
MS	Mean Sensitivity
MVA	Motor Vehicle Accident
NHTSA	National Highway Traffic Safety Administration
NTG	Normal-Tension Glaucoma
PDF	Probability Density Function
POAG	Primary Open-Angle Glaucoma

RT	Reaction Time / Response Time
RTP	Reaction Time Perimetry
SAP	Standard Automated Perimetry
SD	Standard Deviation
SDLP	Standard Deviation of the Lateral Position
SITA	Swedish Interactive Threshold Algorithm
SRT	Simple Reaction Time
TMT	Trail Making Test
UFOV	Useful Field of View
UMCG	University Medical Center Groningen, Netherlands
VA	Visual Acuity
VF	Visual Field
VTS	Vienna Test System
WAIS-IV	Wechsler Adult Intelligence Scale IV
ZEST	Zippy Estimation by Sequential Testing

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Mäntysalo T, Leinonen MT, Suominen J, **Tigchelaar I**, Järvinen P, Määttä E, Pentillä K, Kuutti R & Lahdenperä S (2018). *Method and system for evaluating reliability of results in a visual reaction test.* WO2019/122533 A1.

II **Tigchelaar I**, de Waard D, Jansonius, NM and Leinonen, MT (2021), Exploring the effect of glaucomatous visual field defects of current drivers on a neuropsychological test battery. Acta Ophthalmol.
doi: https://doi.org/10.1111/aos.14975

III **Tigchelaar I**, de Waard D, Jansonius NM, Leinonen MT. Evaluating the relationship between vision and cognition on driving performance in glaucoma patients. Manuscript.

IV Jones PR, **Tigchelaar I**, Demaria G, Wilson I, Bi W, Taylor DJ & Crabb DP (2020): Refinement and preliminary evaluation of two tablet-based tests of real-world visual function. Ophthalmic Physiol Opt 40: 35–46.
doi: https://doi.org/10.1111/opo.12658.

V Jones PR, Demaria G, **Tigchelaar I**, Asfaw DS, Edgar DF, Campbell P, Callaghan T, Crabb DP (2020); The Human Touch: Using a Webcam to Autonomously Monitor Compliance During Visual Field Assessments. Trans. Vis. Sci. Tech. 8:31.
doi: https://doi.org/10.1167/tvst.9.8.31.

1 Introduction

Currently, in ophthalmological clinical settings, vision is measured by assessing its various components separately. For example, visual acuity and contrast sensitivity are measured with charts (Figure 1) and visual field defects are measured using a perimeter. Currently, vision tests are often done without allowing eye movements called saccades, but normal visual perception is based on accurate saccades to aim the fovea to the target. This targeting is based on information from peripheral vision and visual memory. Therefore, even though these measurements are useful for clinicians and aid in diagnosis, they are not a good reflection of real-world visual functioning. Vision in real life is often time-limited, objects are moving, and objects are crowded by other objects. Therefore, new ways to measure vision that are more related to real-life functioning are needed to evaluate how patients with various ophthalmic conditions cope in their daily life.

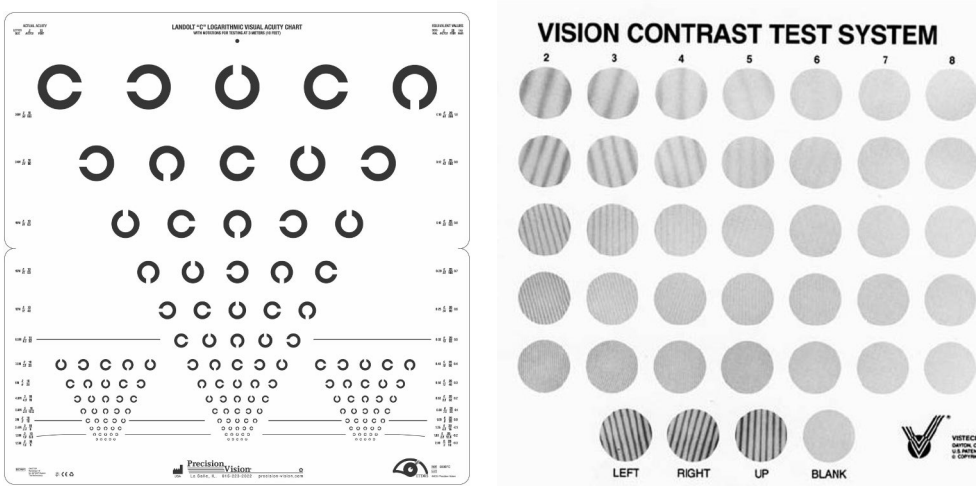


Figure 1. Landolt C Visual Acuity Chart (left) and a contrast sensitivity chart (right) used in clinical settings to evaluate visual acuity and contrast sensitivity. Source: Precision Vision & JK Kumagai, S Wiliams & DW Kline (2005)

One ophthalmic condition that causes visual field defects is glaucoma. Glaucoma is an ophthalmic neurodegenerative disease that damages the optic nerve. The global prevalence of glaucoma is around 3% to 4% (Quigley & Broman 2006), and risk factors include high intraocular pressure, high myopia, and older age (Coleman & Miglior 2008, Weinreb & Khaw 2004). In Primary Open-Angle Glaucoma (POAG), fluid builds up in the eye, and intraocular pressure increases. However, there is also a form of glaucoma, called normal-tension glaucoma (NTG), where the optic nerve is damaged, but the intraocular pressure is not elevated. Both forms of glaucoma lead to irreversible visual field loss. Since the visual field defects usually start small and appear in the midperiphery first, glaucomatous defects often go unnoticed for a long time.

The golden standard for assessing glaucomatous visual field defects is a type of perimetry called Standard Automated Perimetry (SAP). In SAP, patients have to keep fixating on a fixation object around which flashes of light will appear. The patients then respond to seeing the flash by clicking a button. This test is done in a dark room, with the patient's head on a chinrest and they are not allowed to move their eyes (Figure 2, left). While this is an effective way of monitoring glaucoma progression, this type of test is largely unrelated to real-life functioning. Glaucoma patients often experience difficulties in daily life, including face recognition (Glen et al. 2012, Roux-Sibilon et al. 2018), searching for something, for example in the supermarket (Lee, Wood & Black 2020, Wiecek et al. 2012), and driving (Haymes et al. 2008, Lee, Black & Wood 2018, Wood et al. 2016) but current methods do not relate well to these difficulties.

Figure 2. Two ways to measure the visual field. The left image shows a device used for Standard Automated Perimetry (source: https://www.zeiss.com/meditec/en/products/perimetry.html). When the test starts, the lights in the room are turned off. The right image shows the Ocusweep device.

Ocusweep is a device that can measure different aspects of vision, including visual acuity, contrast sensitivity, and the visual field. Additionally, Ocusweep can measure functional vision by using tests that require both intact vision and cognition, such as attentional capabilities and saccades (Figure 2, right). We hypothesize that combining visual and cognitive measures are more related to real-world functioning.

This thesis focuses on real-world effects of visual impairment, with emphasis on glaucomatous visual field defects, driving and new ways to test functional vision with new portable tests. Several novel ways to measure functional vision are introduced and evaluated. The first new test is a reaction time test to assess reaction times to basic stimuli and decreased attention. This test was then included in a study on driving performance of glaucoma patients and a control group in a driving simulator, in which the effect of vision and cognition was evaluated. Secondly, current methods for evaluating cognition usually involve visual stimuli, but the effect of glaucomatous visual field defects on those tests has not been extensively examined before. Therefore, glaucoma patients and a healthy group of age-similar participants were compared on performance on neuropsychological tests, aiming to identify the role of visual field defects on performance on a neuropsychological test battery. Next, new tablet-based tests for visual search and face recognition were evaluated in a young healthy control group showing that tablet-based tests could be feasible to use, and a built-in webcam for monitoring compliance during vision testing showed potential for future vision testing.

2 Review of the Literature

2.1 Regulations concerning driving with visual and cognitive impairment

Most tests for vision and cognition don't have a safe driver cut-off, which means that researchers and clinicians can make different conclusions based on the same test results. Policy makers try to define a set of criteria that allow safe drivers to stay on the road, and unfit drivers to stop driving. There are usually different standards for personal car and professional drivers and these standards can differ per country. The requirements for professional driving are usually stricter than for individuals, since for individuals the reason to drive is mobility and they deserve the benefit of the doubt, and for professional drivers the emphasis is on professionalism and protection of other road users (Bredewoud 2008). Even when criteria are available, there is no consensus as to who has to address the issue of driving throughout the European Union (EU). Primary care physicians have mixed feelings about the appropriateness of their role in fitness to drive assessments and many feel uncomfortable and unsure about advising those with cognitive impairment (Sinnott et al. 2018).

The golden standard for evaluating functional vision loss in glaucoma is the visual field test, usually a Standard Automated Perimetry (SAP) test. Both the horizontal and vertical extent of the visual field are related to driving performance (Bowers et al. 2005). The minimal horizontal extent of the visual field in the United States differs between states and ranges from 105 to 150 degrees wide (Carr et al. 2011, Steinkuller 2010). In the European Union the regulation is a visual field extent of minimal 120 degrees wide, 20 degrees vertical and no central scotomas. Despite evidence that visual acuity has a weak relationship with driving performance and fitness to drive, most regulations in the United States as well as in Europe include a minimum visual acuity for driving as the only requirement.

In the Netherlands, the Centraal Bureau voor Rijvaardigheidsbewijzen (CBR) asks drivers older than 70 years old for a health statement. This is a questionnaire that assesses whether the person is still medically capable of driving. The questionnaire includes questions about a range of domains, including but not limited to physique (use of arms and legs), vision, cardiac conditions, and neurological disorders. For glaucoma patients, when they don't fulfil the criteria for driving but are a borderline case, they

are sometimes offered an on-road driving test. Then, the patients have a chance to prove their fitness to drive. These tests however are not standardized and depend on the examiner, the time of day, the weather conditions, etc.

2.2 The gap in knowledge and contents of this thesis

Because mobility is important and the population of older drivers is increasing (Box, Gandolfi & Mitchell 2010), there is also an increasing need for driving regulations in the older population with visual and cognitive impairments. Currently, there is a gap between the knowledge gathered in scientific studies and the current regulations. Where the current regulations for drivers with glaucoma are focused on the extent of the visual field, research has shown that there are better predictors available for identifying safe drivers. To close the gap, a multidisciplinary approach is needed that takes both vision and cognition of older drivers into account. The cognitive resources might predict which of the glaucoma patients could compensate for their visual field defects. This way, those patients with compensational skills would still be able to drive safely. To assess both vision and cognition, and to fully evaluate compensational strategies, an individualized approach is recommended (Economou et al. 2020).

Another aspect could be the availability of training. One possibility is providing education about the risks that drivers with visual field defects are facing, or learning about the possibility of behavioural adaptations, such as avoidance of certain driving situations. Increasing public awareness related to driving and aging could optimize driving safety in this population (Rudman et al. 2006). Another approach would be to improve the processing speed through training, which can reduce crash risk (Wood & Owsley 2014). Since some drivers with glaucoma are able to compensate for their visual field defects, Racette & Casson (2005) argue for an individual approach when assessing patients with visual field defects. Finding solutions that fit each patient might make it possible to retain mobility, and therefore a higher quality of life, for longer. This literature review will introduce the factors that play a role in fitness to drive in older people, driving with visual and cognitive impairment and the effects of different study designs and personal factors.

2.3 Driving and older people

The older population is continuously growing (Australian Bureau of Statistics 2005, UNDESA 2019), and therefore the number of older drivers is increasing too (Box, Gandolfi & Mitchell 2010). In 2020 there were over 47 million licensed drivers over the age of 65 in the US and in 2019, approximately 8,000 persons aged over 65 were killed in traffic crashes in the US, and over 250,000 persons over age 65 were treated

for crash injuries (CDC 2020). In Finland, data from 2020 (Table 1) paints a similar picture (Tilastokeskus 2020). Next to young and unexperienced drivers, older drivers are at greater risk to die or get seriously injured in Motor Vehicle Accidents (MVAs). Higher crash risk in those groups gives the impression that age is a risk factor for MVAs (Insurance Institute for Highway Safety 2012). However, on the other hand, studies show that older drivers are less likely to be involved in motor crashes. This discrepancy can be explained when driving exposure is considered as well. Older drivers are more likely to be fatally injured during a MVA, when the numbers are corrected for the distance driven (McKnight & McKnight 1999).

Table 1. Data from Finland in 2020 showing deaths and serious injuries due to road traffic incidents. The data show that young and older drivers are at a relatively higher risk for death and serious injuries following a road traffic accident. Source: pxnet2.stat.fi/PXWeb/p

Age	Population	Death	Seriously Injured
< 14	860861	3 (0.00035%)	24 (0.0028%)
15 – 24	606642	37 (0.0061%)	108 (0.0178%)
25 – 34	708057	29 (0.0041%)	49 (0.0069%)
35 – 44	712225	32 (0.0045%)	45 (0.0063%)
45 – 54	665036	25 (0.0038%)	35 (0.0053%)
55 - 64	725034	28 (0.0039%)	54 (0.0074%)
65 - 74	708103	29 (0.0041%)	51 (0.0072%)
> 75	547835	40 (0.0073%)	42 (0.0077%)
Total	5533793	223 (0.004%)	408 (0.0074%)

Next to the growing number of older drivers, the traffic environment is getting more and more complex, with busier roads and more complex driving situations. Driving is an important aspect of retaining a high quality of life for older people. Musselwhite & Haddad (2018) proposed a (renewed) model that incorporates three levels of mobility needs of older drivers, all related to self-perceived quality of life (Figure 3). Most important is the utilitarian need to drive, which makes going to different places easy and quick. The second most important is the need for independence and social needs. Lastly, aesthetic needs to travel are important as well. This shows that driving is important for the older population on multiple levels, and driving cessation in older drivers has thus been linked to decreased out of home activities, admission to long-term care, depression and functional impairment in older people (Chihuri et al. 2016, Fonda, Wallace & Herzog 2001, Marottoli et al. 2000, Marottoli et al. 1997, Windsor et al. 2007). However, continuing to drive when a person is no longer fit to drive is hazardous for themselves as well as for others.

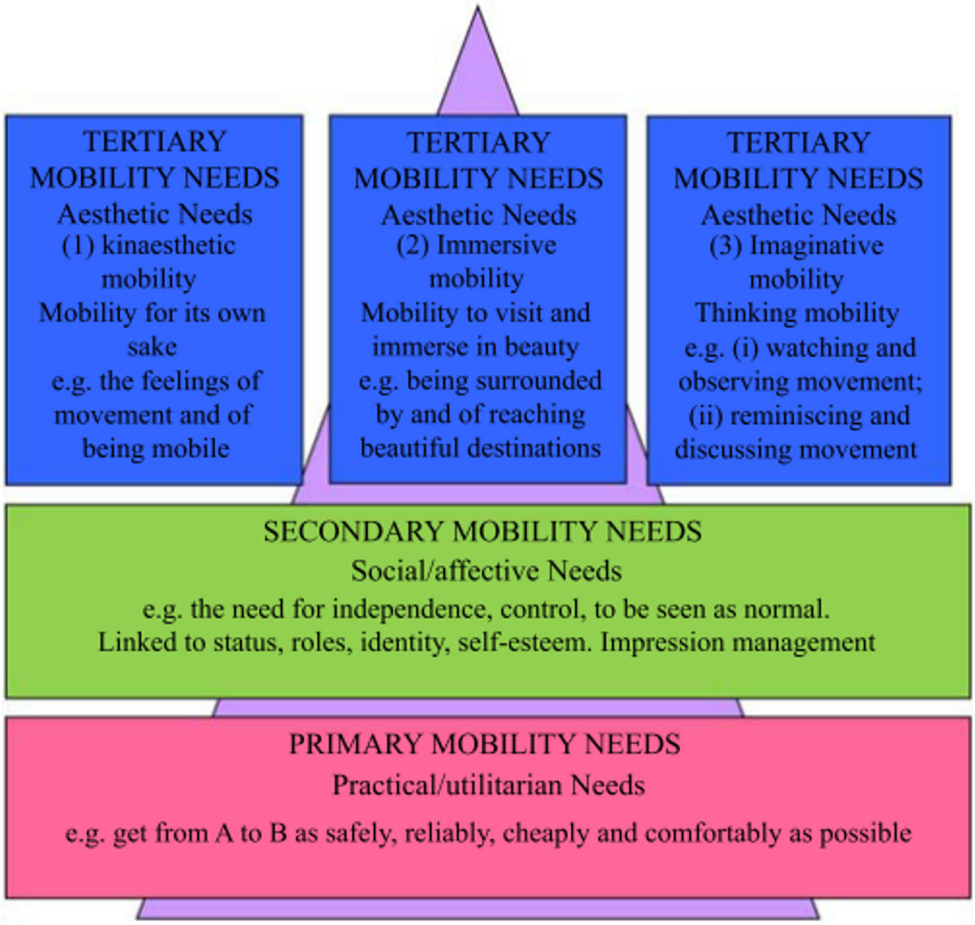


Figure 3. The hierarchical mobility model by Musselwhite & Haddad with three levels of mobility needs of older drivers, with the most important needs at the bottom. This model shows that driving is important on multiple levels, beyond the need to get from A to B.

This highlights the importance of fitness to drive assessments, especially since drivers cannot accurately estimate their driving abilities due to optimism bias and (Dalziel & Job 1997, Roy & Liersch 2013) the overconfidence effect that affects peoples judgement on their own performance (De Bondt & Thaler 1995, Pallier et al. 2002). This means that drivers overestimate their driving skills and underestimate their chances of being involved in negative events, including car crashes. This could lead to situations in which older drivers expose themselves to high risk situations even though they experience age-related cognitive and visual dysfunction (Merickel et al. 2019).

2.4 Study designs in research on driving

There are several possible outcome measures in driving research. One of them is crashes, or MVAs, either self-reported or through government or insurance data. In the case of MVAs, driving safety is reported in the number of MVAs in a certain time, or related to the amount of driving exposure in kilometres driven in a certain time frame. The most informative are at-fault MVAs, where the driver had an active role in the accident that happened. However, most studies include all MVA's since they are relatively rare events (Owsley, Wood & McGwin 2015).

The downside of using self-reported data is that it can be an underestimation of the real numbers, since participants might be reluctant to disclose their MVAs in the past because of privacy concerns (Owsley, Wood & McGwin 2015) or simply because they do not remember. Government data is both expensive to acquire and not always available. Next to this, crashes don't happen often enough to use them as an outcome in many study designs and there is only a moderate level of agreement between self-reported and state recorded MVAs (McGwin, Owsley & Ball 1998).

The golden standard in driving research is to use on-road tests as the outcome measure, since it has the best ecological validity. However, sometimes the use of a driving simulator is preferred. The driving simulator can, for example, be used to test hazardous situations that would be too unsafe to reconstruct on-road. Another reason is to get more data, or a different type of data. Variables that can be easily downloaded from a driving simulator include reaction times, lane positioning and lane position deviations, such as the Standard Deviation of the Lateral Position (SDLP), speed and steering activity. Next to that, the driving conditions, like traffic, time of day and weather conditions, can be chosen and are held constant between all participants. A driving simulator has even be used as a performance-based test for evaluation of functional impairment in glaucoma (Medeiros et al. 2012). There are however also downsides to using a driving simulator. Correlations between driving simulator performance and on-road driving performance are usually moderate to high and results are comparable, but not always similar in an absolute way (Bédard et al. 2010, Blana & Golias 2002, Eramudugolla et al. 2016, Lee, Cameron & Lee 2003, Mayhew et al. 2011, Törnros 1998, Yan et al. 2008). Another downside is the occurrence of simulator sickness, which is a type of motion sickness experienced during, for example, virtual reality or driving simulators. Particularly older individuals are susceptible to simulator sickness and can often lead to missing data (Brooks et al. 2010).

There are a lot of confounding factors in driving research, including driving experience, driving exposure and comorbidities. Driving experience is defined as how many years someone has been driving, where driving exposure is related to current driving behaviour. Driving experience is commonly thought of as a protective factor when assessing MVAs, since more experience and more exposure

leads to well-practiced drivers with automated (and therefore quick) actions. However, one can also argue that drivers with a lot of experience and exposure have more chances to be in a MVA simply because they are on the road more often (Tatham et al. 2015). Therefore, it is important that studies take driving exposure into account when assessing MVAs. Comorbidities exist and can also influence driving performance. Some of these comorbidities are physical, such as missing a limb, some are related to mental disorders like depression and anxiety and some relate to cognition such as slow reaction times or processing speeds (Huisingh, McGwin & Owsley 2016).

Driving is a complex task and requires integration of several domains, such as vision, cognition, and psychomotor skills. These functions commonly decline with older age, which contributes to an elevated crash risk of older drivers (Anstey et al. 2005, Mathias & Lucas 2009).

2.5 Common accident types in older drivers

Among the most common errors older drivers make in on-road situations are the ones related to speed (Selander et al. 2011). Other errors often emerge at intersections. Older drivers are more often involved in MVAs that emerge from the driver's side of the vehicle (Langford & Koppel 2006) and crashes happen mostly when failing to give right of way to other traffic (Clarke et al. 2010) and they also result in more severe injuries (Boufous et al. 2008). Pollatsek, Romoser and Fisher (2012) summarized the abilities that are needed for safe driving, and they state that diminished cognitive abilities, like a narrow field of view, not directing attention to what is important, poor judgment of vehicle speed, diminished physical abilities to control the vehicle and failure to turn their heads are related to these crashes. Next to this, there is also evidence that scanning behaviour of older drivers is less effective than in younger experienced drivers (Bao & Boyle 2009, Clarke et al. 2010, Romoser & Fisher 2009). A brief training program negated these effects in Romoser's and Fisher's (2009) study, which led them to believe that an active training program might work to improve driver safety.

2.6 Driving and vision

Vision is the sense that is valued as most important (Enoch et al. 2019) and it is also the most important sense for driving. Several different tests are used to assess different domains of vision, such as visual acuity, contrast sensitivity and the visual field. These tests are useful to determine what deficits might be present and to diagnose diseases, but do not show real-world problems that patients may have, for example when driving. Vision related problems can compromise driving

capabilities, as driving is a multifaceted skill and requires careful specification when doing research. The skills of driving can be subdivided into several levels, according to Michon's model (Table 2). For example, driving behaviour is related to the strategic level of driving and the tactical level. It encompasses both planning of the route and time of day as well as the chosen speed and following distance. Driving skill is related to both operational and tactical levels of driving. Compensational strategies can be incorporated both on the strategic level and the tactical level. Fitness to drive is a medical suitability to drive and is independent of skill. Fitness to drive depends partly on country specific regulations and on the judgment of the clinician. There are several eye conditions in which it is unclear whether driving is still safe under which circumstances, and glaucoma is one of them.

Table 2. Michon (1985) proposed a hierarchy of driving and proposed a model with three levels. Visual impairment can have an effect on driving at the operational level and the tactical level, and compensational strategies can be implemented at the strategic and tactical level.

Level	Description
Operation level	Controlling the vehicle through physical actions, like operating the steering wheel and the pedals. These skills are overlearned and habitual and therefore largely automatic.
Tactical level	Manoeuvre control. It includes the ability to complete goal directed actions, such as lane position, speed and gap acceptance.
Strategic level	General planning of a trip, including routes and adapting to changing situations.

2.6.1 Glaucoma

The glaucoma's are a group of progressive optic neuropathies characterized by degeneration of retinal ganglion cells and resulting changes in the optic nerve head. This damage to the optical nerve leads to visual field loss. Usually, the midperipheral visual field is lost first, and the centre is relatively spared. However, the centre of the visual field can also be affected in later stages of the disease, and sometimes even in the beginning (Aulhorn & Karmeyer 1977, Brusini & Johnson 2007). Therefore, the visual acuity is often intact for a long time after first diagnosing glaucoma. A side effect of this is that glaucoma may go unnoticed for a long time. About 3% of all people between 40 and 80 years old have glaucoma (Tham et al. 2014). Glaucomatous visual field defects can be present in one eye but not in the other, or if it is in both eyes, the visual field defects can be in different locations (Huang et al. 2014), allowing compensation by the other eye. When the defects are in the same location in both eyes, the eyes cannot compensate. However, strategies such as head

movements and additional or modified saccades might compensate for the defect (Kübler et al. 2015, Smith, Glen & Crabb 2012, Wood et al. 2011).

Research has demonstrated that some drivers with glaucoma can be safe traffic participants, but that tests are needed to differentiate between safe and unsafe drivers. Assessing fitness to drive in glaucoma is complicated, because both the degree and location of the visual field defect vary in the patient group. Glaucoma was a risk factor for MVAs in state records, but not in self-reported outcomes (McGwin, Owsley & Ball 1998). This study however used a cut-off p-value of 0.10. When comparing police record data, glaucoma patients do not have more crashes than a control group (McGwin et al. 2004). In an on-road study of driving behaviour in glaucoma patients with bilateral moderate to severe glaucoma, 52% of the patients had a marginal or fail score, compared to 21% of controls (Bhorade et al. 2016). Moderate to severe loss in the worse eye is related to state-record crashes (McGwin et al. 2005). In a study on 2000 licensed glaucoma drivers, the MVA rate was 1.65 times higher in glaucoma patients compared to controls, after adjusting for age, gender, and mental status (MMSE). The rate of at-fault MVAs was 2.5 times higher in those with severe visual field defects compared to normal visual fields to moderate visual field loss (Kwon et al. 2016). Not only severe, but also mild to moderate visual field loss can have a negative impact on driving performance (Bowers et al. 2005). Glaucoma patients were worse at making certain driving manoeuvres than controls and notice and respond later to hazards later (Lee, Black & Wood 2017). Drivers with glaucoma are rated as less safe during an on-road driving test and made more errors than their age-matched controls. The types of errors that glaucoma patients typically make included slightly worse lane positioning and planning and approaching traffic lights and give-way intersections (Wood et al. 2016). Problems with speed matching when changing lanes, lane keeping, taking turns and lane changing have been mentioned in other studies as well, next to gap judgement (Bowers et al. 2005, Kasneci et al. 2014). One study did not find differences in lane keeping between glaucoma patients and controls, and they think this is because drivers with glaucoma prioritize different parts of the driving task compared to controls. For glaucoma patients, lane keeping was most important but to controls it was steering smoothness. This could be the reason why glaucoma patients miss more peripherally presented stimuli but that lane keeping was not different from controls in this relatively small sample (Prado Vega et al. 2013). When increasing cognitive task demands, the gaze of a driver concentrates more towards the centre. This type of inattentional blindness is called tunnel vision or cognitive tunnelling (Mack & Rock 1998). Glaucoma patients often report higher cognitive workload when driving, possibly due to compensational strategies that require attention and executive functioning (Engström, Johansson & Östlund 2005). Cognitive tunnelling due to high cognitive demands of glaucoma drivers is therefore an additional

explanation of why glaucoma patients miss peripheral stimuli more often (Adrian et al. 2019).

The driving simulator has proven itself as a valuable research tool in older drivers, but also in glaucoma driving simulator performance is comparable to on-road driving (Ungewiss et al. 2018). Several studies using simulators have found that glaucoma patients have impaired driving performance. They, for example, have a higher crash incidence compared to controls (Szlyk et al. 2005) and glaucoma patients also have more difficulties with vehicle control than controls, as they had more jerky steering behaviour (Prado Vega et al. 2013). Cognitive demands disproportionally affect glaucoma drivers performance in a functional visual field test related to driving (Gangeddula et al. 2017, Prado Vega et al. 2013). This is similar to what has been found in Useful Field Of View (UFOV) research (see section 2.7.3 Attention). Differences between glaucoma patients and controls are also found in hazard perception tests, where glaucoma patients had poorer overall driving scores and hit more hazards than controls (Lee, Black & Wood 2018). The experimental designs of studies investigating driving in glaucoma patients vary in quality, sample size and glaucoma severity and location of the visual field defect which makes comparing different studies complicated.

2.6.1.1 Compensational strategies when driving with visual field defects

One explanation for finding different results in studying glaucoma and driving, next to differences in visual field defects, is compensational mechanisms. There are several ways in which glaucoma patients may compensate for their visual field defects. For example, some glaucoma patients drive slower to allow more time to scan the environment. This idea has led to several studies on the eye and head movements of glaucoma patients. Head movements did not differ between the control group and glaucoma patients in closed circuit on-road driving (Lee, Black & Wood 2018) but it is important to combine the both head and eye movements when evaluating driving, since they can successfully occur separately and also together. Another compensational mechanism could be self-regulation of driving behaviour and exposure. Glaucoma patients often self-regulate their driving behaviour and avoid demanding driving situations more often than controls (Yamasaki et al. 2020). This is an adaption of driving behaviour at the level of strategy in Michons model (Table 2). Situations that glaucoma patients often avoid are driving in bad weather, such as rain or fog, driving during rush hour, driving on the highway and driving in at night (McGwin et al. 2004). The chances that a glaucoma patient stops driving increases with increasing Visual Field (VF) loss severity (Van Landingham et al. 2013). There are, however, glaucoma patients who continue driving with severe VF loss (Ramulu et al. 2009), which can be explained by the fact that glaucoma patients

can be unaware of their visual field defects (Crabb et al. 2013). Often, drivers cannot accurately estimate their own driving skills. Most drivers think they should be able to continue driving and that they drive better than average, even when an on-road driving test deems them as unsafe drivers (Selander et al. 2011). For self-regulation to occur, awareness of the visual field defects is necessary.

2.7 Vision of glaucoma patients and driving

There are several domains in vision, each with a different test and importance to driving. Below, research on driving and visual acuity, contrast sensitivity, the visual field and saccades will be discussed.

2.7.1 Visual acuity and central vision loss

Numerous studies have investigated the role of visual acuity and central vision loss in driving performance. Despite the intuitive link between visual acuity and driving, these studies show only a weak or non-significant relationship with driving performance regardless of the visual acuity thresholds applied (Keeffe et al. 2002, Mccloskey et al. 1994, Owsley, McGwin Jr & Ball 1998, Rubin et al. 2007).

Several factors may explain this weak relationship. First, driving is a visually complex task that relies on a wide range of visual functions beyond acuity alone, such as peripheral vision, contrast sensitivity, and motion perception. Another reason could be a selection bias: most drivers with a very low visual acuity are legally not allowed to drive and are therefore not present in the datasets of studies on visual acuity and driving although a large percentage of drivers continue to drive when they don't meet the requirements (Levecq, De Potter & Jamart 2013). Additionally, because low visual acuity is easy to notice, those with low visual acuity often self-regulate their driving behaviour and quit driving (Keeffe et al. 2002). In glaucoma, visual acuity is usually not affected in the beginning stages of the disease.

2.7.2 Contrast sensitivity

Contrast sensitivity has been shown to be a stronger predictor than visual acuity for driving safety. For example, cataract patients have cloudy lenses which reduce contrast (Ball et al. 1998, Owsley et al. 1999) and often self-regulate their driving behaviour by avoiding situations like driving at night or in bad weather (Rubin et al. 1994). But despite self-regulating their driving behaviour, drivers with reduced contrast are still more at risk for MVAs (Rubin et al. 1994, Singletary et al. 2017, Tatham et al. 2015), have poorer manoeuvre execution and scored lower on summary driving measures (Bowers et al. 2005). Chances of a history with one or more MVAs

in cataract patients is 2.5 times higher than in an age-matched control group (Owsley et al. 1999) and drivers with a MVA history are 8 times more likely to have a loss of contrast sensitivity (Owsley et al. 2001). Other studies have confirmed these findings by conducting on-road driving tests with simulated and true cataracts (Wood & Troutbeck 1994, Wood & Troutbeck 1995). A reason for the relationship between reduced contrast and driving performance is that with low contrast, people need more time to react. This was studied with artificially lowered contrast sensitivity, which increased reaction times (Wood et al. 2010). Glaucoma patients may also experience reduced contrast sensitivity (Bhorade et al. 2016, Hawkins et al. 2003, Lahav et al. 2011, Richman et al. 2010).

2.7.3 Visual field

Visual field loss is the most prominent symptom of glaucoma, affecting one or both eyes. A normal visual field spans 180 degrees horizontally and 130 degrees vertically, of which about 95% is peripheral and therefore has poor spatial resolution. Nevertheless, the peripheral visual field plays a critical role in detecting motion and guiding attention. It registers movement and directs the eyes toward changes in the environment through visually guided saccades (Martin, 2017). While monocular visual field loss is a poor predictor of crash involvement, binocular visual field defects are related to crash risk (Johnson & Keltner 1983, McGwin et al. 2015, Rubin et al. 2007). The severity of the visual field defects is related to self-reported MVAs in the past 5 years (Ono et al. 2015) but there is no consensus on the specific type or severity of visual field defects most strongly linked to crash involvement (Correa et al. 2019).

Artificial defects in the upper half of the visual field are related to lower Hazard Perception Test (HPT) scores (Glen, Smith & Crabb 2015), due to the higher probability of events outside of the car being visible in the upper half compared to the lower half as the lower visual field is mostly occupied by the features of the car, such as the dashboard and steering wheel. Some studies (Szlyk et al. 2005), report that the horizontal and peripheral extent of visual field loss correlates with the number of accidents in a driving simulator, while others don't find that association (Kasneci et al., 2014). This can again be explained by the ability to compensate for visual field defects by early scanning and more head movements (Coeckelbergh et al. 2002).

2.7.4 Saccades

Saccades are rapid eye movements towards a target and can either be voluntary and endogenously driven, or visually guided (McDowell et al. 2008). Saccades have been

associated with driver safety in research because older drivers are at increased risk of intersection crashes (US Department of Transportation, 1997) and self-report difficulties with intersections (Eck & Winn 2002). One explanation for these difficulties is that they are less likely to make a saccade outside of the intended path of the vehicle before turning at an intersection (Bao & Boyle 2008, Fisher & Romoser 2009, Pollatsek, Romoser & Fisher 2012). When combining a visual search task, which includes eye movements to scan the visual field, with steering on a highway, older drivers perform worse (Yamani et al. 2016). Glaucoma patients also have different saccades than controls, which could be the reason for their increased risk of crashes. However, findings regarding the specific nature of these saccadic differences remain inconsistent across studies.

Studies on saccades focus mainly on three components of a saccade: the length or amplitude, the location and the number of saccades. Some studies find that glaucoma patients make larger saccades, which was related to better driving scores. However, in a control group made smaller saccades and within this group large saccades were not related to better driving performance (Lee, Black & Wood 2018). The location of the fixation is also important for driving safety. Glaucoma patients that passed an on-road driving test had different exploration patterns than those that failed. Those that focused on the central area longer and made more saccades towards the area of their visual field defect passed the test more often. Saccades can be a form of compensation, in which the driver is gaining a larger field of view by moving their eyes more of a way to compensate for the missing part of the visual field. Head movements also play a role in successful driving behaviour in those with a visual field defect (Kasneci et al. 2014). Because the three aspects of saccades can differ per person, but can all help create a better overview and therefore a safer driver, research has yielded incomplete and sometimes incongruent information. For example, in one study glaucoma patients made more eye movements, with shorter fixations but equivalent amplitudes (Crabb et al. 2010). However, other studies contradict this and say that glaucoma patients make smaller saccades (Lee, Black & Wood 2017), that they have longer saccade reaction times in all eccentricities (Mazumdar et al. 2014) and on average make fewer saccades during a visual search task (Smith, Glen & Crabb 2012). Other studies found no differences in eye movement patterns between glaucoma patients and controls (Prado Vega et al. 2013). Glaucoma patients did however miss peripheral stimuli more often than controls. In this study the speeds were fixed, which removes the opportunity to compensate for visual field defects by slowing speed. Glaucoma patients were 6 times more likely to have the driving instructor intervene because they missed pedestrians in the periphery (Haymes et al. 2008). What can be concluded is that there is a lot of variability in research on saccades and that differences can be small between participant groups.

All the research on vision and driving shows that vision tests alone cannot predict driving performance accurately enough to predict safe and unsafe drivers (Owsley, McGwin & Ball, 1998, Charman 1997, Wood 1999). An intrinsic feature of glaucoma is variability in vision tests, which also complicates research. Combining the measures of vision might help to improve their predictive abilities (Decina & Staplin 1993) and multidisciplinary test batteries with both vision tests and functional tests that include reaction times or visual processing speeds have more predictive value (Wood et al. 2013).

2.8 Neuropsychological tests for driving and cognitive abilities

Other aspects than vision alone might help to explain why some glaucoma patients are still safe drivers and others are not. The ability to compensate for visual field defects by making quick adequate saccades or turning your head for example. The ability to react timely to a hazard or to learn to make saccades to the area of the visual field defect depend on the cognitive resources and self-awareness of the locations of the visual field defect the driver has.

In the elderly population it is common to a form of cognitive decline. In the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) by the American Psychiatric Association, cognitive decline beyond what can be expected based on age alone was called Mild Cognitive Impairment (MCI) and a lot of research has been done on this topic. In the new DSM-V, the term is mild neurocognitive disorder (mNCD). Both terms essentially mean the same, in that one of the cognitive domains is affected, but the complaints do not fulfil the requirements for dementia. The impact on fitness to drive of dementia is often clearly regulated, but that is not the case with mNCD. Guidelines vary in Europe, with some countries giving out lifelong driver's licenses with no medical checks, some countries rely on self-reporting of medical conditions, and some require a check after a certain age is reached. In the United States of America, the National Highway Traffic Safety Administration (NHTSA) recommends drivers with dementia to get a health check, but mNCD is not mentioned. Overall, there is no consensus on what steps should be taken and who is responsible.

Drivers with mNCD score lower on off-road and on-road assessments of driving safety (Anstey et al. 2017). When testing fitness to drive in the case of cognitive impairments, there are several neuropsychological tests that can assess all the cognitive domains relevant for driving. Some countries have already adapted their screening procedure. For example, in Denmark a short version of the cognitive screening test called the Mini-Mental State Examination (MMSE) and a clock drawing test were added to the screening procedure for older drivers. However,

adding these tests to the procedure did not lead to less accidents (Siren & Meng 2011). One test like this cannot identify safe or unsafe drivers, but a test battery is needed that covers all important cognitive domains (Lincoln et al. 2006). The most tested domains for driving are attention, executive functioning, slowed psychomotor speeds and visual and spatial deficits. All of these will be discussed in the following section.

2.9 Visual and cognitive impairments

Both glaucoma and cognitive impairment are common in the older population (Fukuoka, Nagaya & Toba 2015). The association between glaucoma and cognitive impairment has been studied, but without a definitive answer. There are biological similarities in glaucoma and Alzheimer's Disease, but a common cause is not yet identified. There are several explanations for the overlap in pathology. The first explanation could be that glaucoma and Alzheimer's Disease have common pathogenic mechanisms, for example altered cerebrospinal fluid circulation, intracranial pressure of genetic differences in gene coding for certain proteins, such as optineurin (Mancino et al. 2017). Another explanation is that glaucoma is not a neurodegenerative disease like Alzheimer's Disease, but does have secondary effect on the central nervous system (Danesh-Meyer & Levin 2015). This relates to the phrase "use it or lose it", which means that cells that are not activated for an extended period die off. Some studies find no difference in cognitive functioning between glaucoma and controls (Jonas et al. 2018, McCoskey et al. 2018). Although McCoskey found no difference between glaucoma patients and controls, there was an additional effect of visual field loss and cognitive impairment (McCoskey et al. 2018). Others do find that glaucoma patients often have mild cognitive impairment (Harrabi et al. 2015, Yochim et al. 2012).

Even without a defined common pathology, cognitive impairment and glaucoma can co-exist, and this is something that should be considered when testing fitness to drive in glaucoma patients. For example, Alzheimer patients often have visual complaints (Ho et al. 2012) which can slow down perception, such as increased latencies of all saccades compared to controls, impaired smooth pursuit of targets and more anti-saccade errors (Garbutt et al. 2008, Rashbass 1961, Rottach et al. 1996). Researchers need to develop a reliable and valid composite battery that can determine driver safety in individuals with forms of mild cognitive impairment (Bennett, Chekaluk & Batchelor 2016). One study looked at fitness to drive in both glaucoma patients with and without cognitive impairment (Rosen et al. 2018). They investigated whether there is an additive effect of glaucoma on cognitive impairment. Results showed that glaucoma patients are generally not more cognitively impaired than a control group. A combination of tests for vision,

cognition and motor domains such as reaction times are able to classify fit and unfit drivers when using on-road driving as an outcome measure (Wood et al. 2008). Careful interpretation of cognitive scores is however needed because most tests have visual components that could be a disadvantage for glaucoma patients.

2.9.1 Cognitive screening tests

Cognitive screening tests have been used a lot in determining fitness to drive in research, as well as in the clinic. The Mini Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA) are the most used screening tests.

The MoCA assesses several different domains, including visuospatial skills, attention, memory, language, and abstract reasoning. The maximum score on the MoCA is 30 points. Scores of 26 or higher are considered in the normal range, scores between 22 and 25 indicate the possibility of cognitive impairment, and scores of 21 or below indicate the possibility of a more significant impairment (Nasreddine et al. 2005).The sensitivity of the MoCA is high, around 90% for mNCD patients and 100% for patients with Alzheimer's disease (Nasreddine et al. 2005, Wittich et al. 2010). The MoCA is one of the few tests that has a version specifically for patients with lower vision. This MoCA-Blind test is the same as the MoCA, but the items with visual stimuli are removed from the analysis. In this version, the components visuospatial skills, executive skills and naming are deleted. Together, these tasks are worth 8 out of 30 points. The MoCA-Blind therefore has a maximum score of 22. The suggested cut-off point for cognitive impairment in the MoCA-Blind is any score equal to or below 18 (Wittich et al. 2010). This reduces the sensitivity of the test to about 44% for mNCD and 87% for Alzheimer's disease but yields a better specificity of 98% instead of 87% than the original MoCA for detecting mNCD. Scores on both versions of the test are corrected for education by adding one point for all participants with less than 12 years of education.

As the MoCA does not use the peripheral visual field and there is no time limit, it is not expected that glaucoma patients will experience difficulties on this test due to visual field defects, and they can take longer to answer the visual items without lowering their score. In general, with glaucoma patients, the normal MoCA is used (McCoskey et al. 2018).

Research has shown variable results when using the outcomes of these short screening tests in predicting driving safety (Dawson et al. 2009, Whelihan, DiCarlo & Paul 2005). A reason for this could be that the tests are too short to be related to real world situations, they are insensitive to subtle and early cognitive decline and that the full version with good sensitivity often uses visual stimuli, potentially interfering with the reliability of the assessment for those with visual impairment

(Kempen, Mark & Feldman 1994). Therefore, next to a screening tool, specialized neuropsychological tests are needed to further assess cognitive impairment.

2.9.2 Memory

Working memory can be assessed with the Digit Span Test, which is part of the WAIS-IV neuropsychological test battery. In this test, the participant repeats a series of numbers after the test leader has read them out loud. First, the sequence is three numbers long and the longest sequence is ten numbers. There is a forward condition, with normal repetition, and a backward condition in which the numbers have to be repeated in reverse order. The Digit Span test therefore measures memory without a visual component and can be used as an indicator of cognitive impairment. A lot of neuropsychological tests have a visual component but when assessing fitness to drive with visual impairments, this might be overestimating their cognitive impairment. In glaucoma the centre of the visual field is usually spared but adding non-visual tests to the test battery allows for a comparison with other measures of cognition. Poor Digit Span performance is related to driving cessation (Pyun et al. 2018) and crashes (Park et al. 2011). In a study with individuals with mild cognitive impairment, the neuropsychological tests including the Digit Span scores were indeed lower for the cognitive impaired group but did not have an effect on driving performance (Kawano et al., 2011).

2.9.3 Attention

The term divided attention refers to the ability to effectively pay attention to more than one thing at a time. This ability is a significant predictor when assessing fitness to drive. Attention skills are related to curve coherence, which is the ability to maintain your position on a road with many curves (Gracitelli et al. 2015). The Useful Field of View test (UFOV) (Ball & Owsley 1993) is one of the most extensively studied tests in research on fitness to drive (Wood & Owsley 2014). The focus of the test is to find the minimum duration of the presented stimuli, with increasing task complexity. UFOV has high variability, especially in glaucoma patients (Bentley et al. 2012). However, compared to other tests, the UFOV has yielded relatively consistent results. The UFOV is a test that measures different aspects of attention. The three subtests measure processing speed, divided attention, and selective visual attention. In the first subtest, the subject has to recognize the central target (either car or truck) using a 17-inch touchscreen. The shortest presentation time that gets a correct answer for 75% of the time or more is the score. In the second subtest, which is the divided attention test, is similar but with an additional peripheral localization task. Attention resources are limited, so

performance on this task is usually slower and with more mistakes. In the third subtest, which assesses selective attention capabilities, distractors are added to the second task.

The UFOV is associated with increased crash risk in older drivers (Ball et al. 1993, Clay et al. 2005, Cross et al. 2009, Owsley, McGwin Jr & Ball 1998), delayed hazard response times (Lee, Black & Wood 2017) and overall driving performance (Bowers et al. 2005). The divided attention subtest is the best predictor of the three subtests for crash involvement and the reaction times in the test for selective attention were significantly slower in a group of older drivers who passed an on-road driving test (Selander et al. 2011). The UFOV has better predictive abilities for driving safety than static automated perimetry (Gracitelli et al. 2015, Haymes et al. 2008). The UFOV Divided Attention subtest and the risk category can predict pass/fail on an on-road driving test (Classen et al. 2013). Research has shown that glaucoma patients perform worse on the UFOV than controls (Bhorade et al. 2016). A population based study controlled for confounding variables found no effect of UFOV and rate of MVAs (Friedman et al. 2013). Correcting for confounding variables like age and contrast sensitivity is very important since this might influence the results.

Misdirecting attention might be related to behaviour and could therefore be improved by training. Traning these attentional skills could help improve driver safety (Pollatsek, Romoser & Fisher 2012). An important and relevant example of divided attention related to driving is telephone use while driving. Glaucoma patients have a greater decline in their ability to detect peripheral events, like hazards, when using a phone compared to controls. Glaucoma patients might lose the ability to compensate for their visual field defects when cognitive load is increased. But research shows that glaucoma patients use a phone as much as controls (Ogata et al. 2019) which might put them at increased risk for MVAs.

2.9.4 Visuospatial skills

Mazes have been used in research on driving because they represent the skill of planning a route, requiring memory, executive functioning visuospatial skills. Visuospatial skills are a group of skills that help to process the space and manipulate it. The time to complete a maze and the number of errors have been linked to driving ability (Carr et al. 2011, Ott et al. 2003, Silva, Laks & Engelhardt 2009) and fitness to drive (Staplin et al. 2013). However, in an Alzheimer group there was no relationship between mazes and driving performance (Grace et al. 2005). Research has shown that glaucoma patients performed worse on a maze task than controls (Bhorade et al. 2016). One example of a maze test is the Snellgrove Maze Task (SMT) (Snellgrove 2005), which is a paper-based test in which the participant completes a maze without touching the walls or going into dead ends. Advanced

glaucoma patients are significantly slower to complete the MST (Bhorade et al. 2016).

The Benton Visual Retention Test (BVRT) (Sivan 1992) is a test for visual perception and memory. In this test, the participant is allowed to see the figure for 10 seconds before it is removed and has to be reproduced from memory. There are 10 different images in this test, and the score is the number of correct drawings (max. 10) and the number of mistakes (max. 40). The BVRT has demonstrated an association with driving in healthy older adults (Hunt et al. 1993), driving cessation (Emerson et al. 2012) and driving with both glaucoma and cognitive impairment (Rosen et al. 2018). The study by Rosen et al. (2018) included glaucoma patients with and without cognitive impairment and a control group with and without cognitive impairment found that the BVRT shows an additional cognitive impairment in glaucoma patients, compared with those with cognitive impairment without glaucoma.

2.9.5 Executive functioning

Executive functioning is related to the control of behaviour, such as selecting, monitoring and inhibiting actions. For a long time, executive functioning was not part of fitness to drive research, despite it's almost obvious relation to driving. The Trail Making Test (TMT) (Reitan 1958) is well-known and used in clinics, in research to assess executive functioning, processing speed and (visual) attention and in research on fitness to drive. The TMT is a paper- and pen-based test on an A4 sized paper and is done at reading distance. At that distance, this corresponds to a visual angle diameter of at least 29° by 41° degrees. In the TMT-A the participant has to connect all the numbers with a pen in ascending order, as fast as possible. The total time to complete the test and the number of errors are the outcome measures.

The time to complete the TMT-A was significantly related to passing or failing an on-road driving test (Bhorade et al. 2016, Carr et al. 2011, Hunt et al. 1993). In the second part of the test, the TMT-B, both numbers and letters have to be connected in ascending order, where the numbers and letters have to be alternated (A-1-B-2-C-3). The main difference between TMT-A and TMT-B, is that TMT-A is a measure of processing speed, and the TMT-B requires more cognitive processing related to executive functioning. Research on the effects of simulated low visual acuity and cataract on TMT performance have demonstrated a negative impact on performance (de Haan, Tucha & Heutink 2019, Wood et al. 2010). Glaucoma patients also performed worse on the TMT-A than controls (Bhorade et al. 2016) and are significantly slower on the TMT-B compared with controls (Gangeddula et al. 2017). However, these low scores can probably be explained by the significant role of the

peripheral visual field in this test. It can be argued that the role of vision is larger in TMT-A compared to TMT-B but also in the TMT-B glaucoma patients are slower than controls (Lee, Wood & Black 2020). Altogether, the TMT alone does not have enough predictive power to discriminate those with unsafe driving behaviour (Dobbs & Shergill 2013, Fox et al. 1997).

2.9.6 Reaction time and processing speed

Reaction time tests are used to measure processing speed. Reaction times are important for several different aspects of daily life. Relatively short delays in reaction times can lead to hazardous driving situations (Horswill et al. 2008). For example, when driving 80 km/h a 500-millisecond delay in braking reaction time equals to stopping the car 11 meters later. Quick reaction times are also relevant in sports and physical fitness (Atan & Akyol 2014, Reigal et al. 2019, Sant'Ana et al. 2017). Slow reaction times have been linked to a number of conditions, such as mild cognitive impairment (Andriuta et al. 2019), dementia (Bailon et al. 2010) and are also related to mortality (Hagger-Johnson et al. 2014). Next to traditional simple and choice reaction time measures, there are several ways to measure visual processing speed that are more related to real world situations. However, results can then be confounded by either other cognitive domains that are needed to complete the task, a general cognitive impairment, or visual impairment.

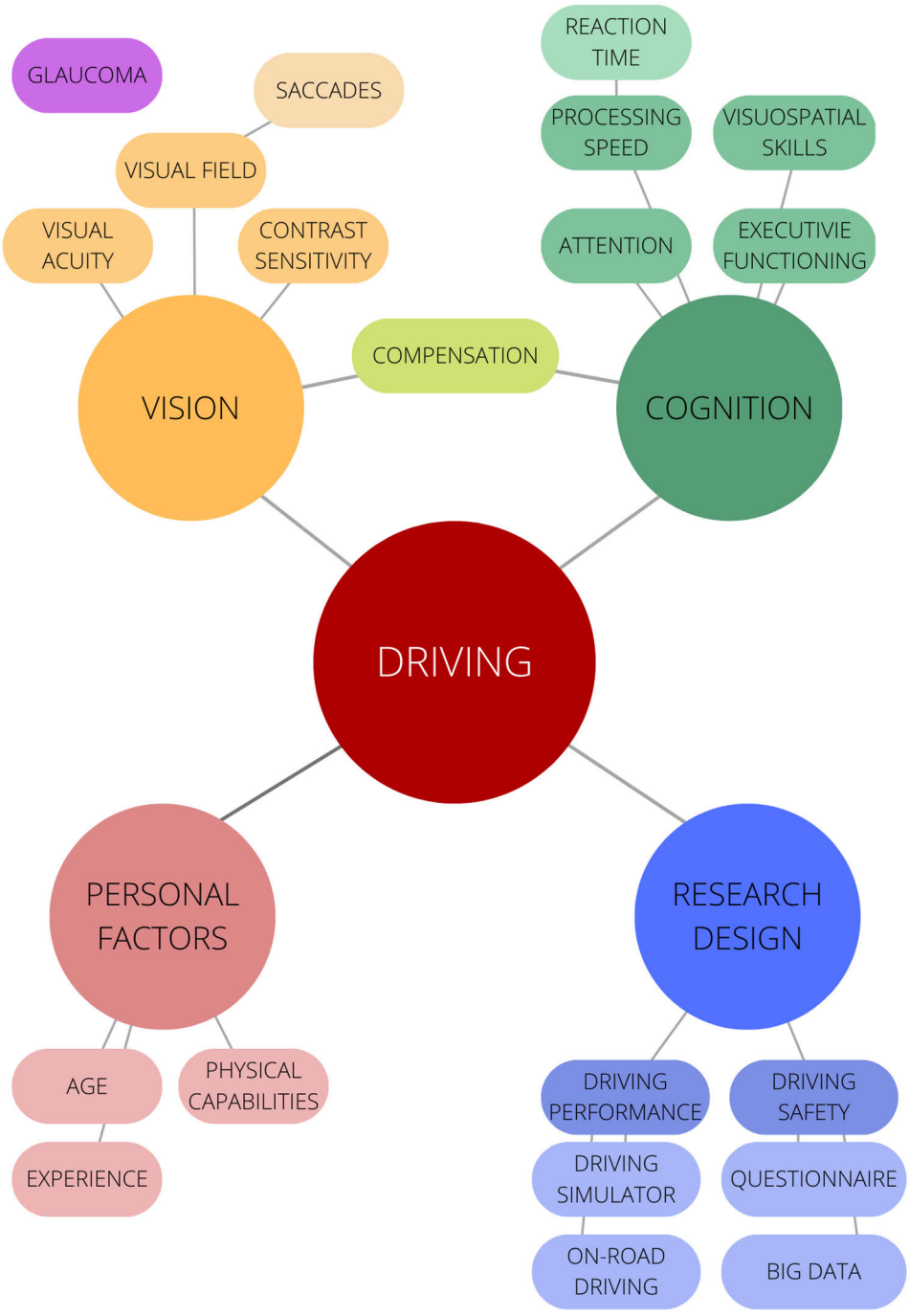
Slower reaction times in mNCD are related to accident probability (Economou et al. 2020). Reaction times to low contrast stimuli while driving a driving simulator are related to a history of (self-reported) MVAs and even outdid conventional perimetry and UFOV (Tatham et al. 2015). Slower reaction times on a reaction time test indicate slower reaction times in traffic, which has an influence on the pass/fail rate of an on-road driving test in older patients with dementia (Piersma et al. 2016).

Despite the importance of quick visual reaction times, speed of vision generally has a small role in clinical vision testing. To evaluate functional vision, both traditional measures of vision should be combined with a timing component that is not heavily confounded by cognitive impairment and is fast and easy to complete. To measure visual processing speed of a subject in real-world situations we must first be aware of the reaction time to a very basic stimulus, far above the visual threshold of seeing. Reaction time tests generally have small learning effects and high test-retest reliability (Lemay et al. 2004, Sakong et al. 2007, Vora et al. 2016).

2.9.7 The effect of glaucomatous visual field defects on a neuropsychological test battery

Neuropsychological test batteries that encompass several cognitive domains are needed to assess cognitive decline. However, one problem in assessing cognition in older people is comorbidity, including visual impairment. Over 2.2 billion people worldwide have some form of visual impairment, caused by, for example, unaddressed refractive error, cataract or glaucoma (World Health Organization 2019). Visual impairment can influence performance on neuropsychological testing as both lowered contrast sensitivity and low visual acuity can have negative effects on neuropsychological test scores, and therefore overestimate the cognitive impairment (Cronin-Golomb, Corkin & Growdon 1995, Toner et al. 2012). Glaucomatous visual field defects might influence performance on neuropsychological tests that utilize a larger part of the visual field. These defects might lead to a slower response because compensation by making additional head or eye movements is needed to find stimuli in the area of the visual field defect. As mentioned before, cognitive impairment is common in the older population (Fukuoka et al. 2015) and can therefore also occur in the glaucoma population. Because neuropsychological testing is often used in the assessments of fitness to drive when cognitive impairment can be expected (Bennett, Chekaluk & Batchelor 2016, Piersma et al. 2016) or for older drivers (Vaucher et al. 2014), the influence of visual field defects needs to be taken into account.

2.9.8 Visual summary



2.9.9 The relationship between vision and cognition with driving performance in glaucoma patients

In studies using state records and police records, glaucoma is sometimes considered a risk factor for crashes (Haymes et al. 2007, Kwon et al. 2016, McGwin, Owsley & Ball 1998, McGwin et al. 2005, Vingrys et al. 2016) but not in all studies (McGwin et al. 2004) and differences between glaucoma patients and controls may be small. Additionally, both the location and the size and depth of these visual field defects vary, and glaucoma may be present monocularly as well as binocularly (Huang et al. 2014). As a result, cut-off values for driving allowance for both visual acuity and visual field tests remain a topic of debate.

In studies that compare driving performance in controls and patients with severe glaucoma (Bhorade et al. 2016, Kwon et al. 2016, McGwin et al. 2005), as well as with moderate glaucoma (Bowers et al. 2005) showed higher crash rates and fail scores. There are specific driving errors that glaucoma patients make more often than controls, such as delayed hazard response times (Lee, Black & Wood 2017), lane positioning (Bowers et al. 2005), the way traffic lights and give-way intersections are approached (Wood et al. 2016), changing lanes (Bowers et al. 2005), and gap judgment (Kasneci et al. 2014).

Fitness to drive in glaucoma patients depends on more than vision alone. Glaucoma patients might be able to compensate for their visual field defects. For example, in case of peripheral visual field defects, glaucoma patients can make extra saccades to be able to see peripheral cues or hazards. While this allows them to see what happens in their blind visual field areas, it might make them slower to react in hazardous situations. Glaucoma patients report higher cognitive workload while driving than controls (Engström, Johansson & Östlund 2005) and they are more affected by additional cognitive tasks (Engström, Johansson & Östlund 2005, Gangeddula et al. 2017, Prado Vega et al. 2013) which might hinder their compensational abilities. Glaucoma patients also self-regulate their driving behaviour. Self-regulation can occur only for specific circumstances (McGwin et al. 2004, Yamasaki et al. 2020), such as avoiding driving during rush hours, in fog, at night or in rain, or it could be complete driving cessation (Blane 2016, Takahashi et al. 2018, Van Landingham et al. 2013).

2.9.10 Real-world visual function testing

Methods for directly assessing performance on everyday visually guided tasks may be more appropriate than traditional clinical measures of basic visual function. There is substantial clinical interest in developing more ‘patient friendly’ tests that can quickly and easily assess real-world visual function (Bastawrous et al. 2015, Dabasia et al. 2014, Glen, Baker & Crabb 2014, Nesaratnam et al. 2017, Strutton et al. 2016).

Recently, tablet tests measuring more basic aspects of visual function were introduced, such as visual acuity (Aslam et al. 2016, Bastawrous et al. 2015, Black et al. 2013, Zhang et al. 2013), contrast sensitivity (Bodduluri et al. 2018, Dorr et al. 2017, Dorr et al. 2013, Rodríguez-Vallejo et al. 2015, Wu et al. 2015), visual fields (Anderson et al. 2017, Johnson et al. 2017, Jones et al. 2019, Kong et al. 2016, Nesaratnam et al. 2017, Prea et al. 2018, Schulz et al. 2018, Vingrys et al. 2016), stereopsis (Rodríguez-Vallejo et al. 2017), and colour vision (Bosten et al. 2019, de Fez et al. 2018).

Face discrimination and visual search are two tasks that are particularly important to patients and are often impaired in people with glaucoma (Glen et al. 2012, Glen, Smith & Crabb 2013, Mazzoli, Urata & Kasahara 2019, Roux-Sibilon et al. 2018, Smith, Crabb & Garway-Heath 2011, Smith, Glen & Crabb 2012). Currently, however, the equipment required for these tests is bulky, and the tests themselves can be time-consuming and are often only used in research and not in a clinical setting. Rapid, tablet-based versions of these real-world measures would be more clinically applicable. Tablet tests are particularly attractive as they are inexpensive, and because patients in the waiting room can potentially complete them while they wait to be seen – thereby minimising any burden to patients and staff. Indeed, the idea of using tablet-based activities to utilise the time patients spend more productively in waiting areas is a growing area in other medical disciplines. For example, tablets are being increasingly used in healthcare to collect questionnaire data, as an educational tool (Brinker et al. 2018, Hassan et al. 2016, Patel et al. 2015, Stribling & Richardson 2016), or for assessing hearing loss (Kelly et al. 2018). However, this concept remains relatively unexplored in ophthalmology.

2.9.11 Compliance and reliability during visual field assessments

Visual field assessments are central to the diagnosis and management of glaucoma and in deciding the fitness to drive of glaucoma patients. However, visual field assessments are often demanding for patients (Glen, Baker & Crabb 2014), require sustained concentration, and patients can become bored, confused, or fatigued, sometimes leading to unreliable data (Lee, Zulauf & Caprioli 1994, Montolio et al. 2012, Rao et al. 2015). If perimeters could recognize when a patient's attention towards the test is decreasing, more reliable data would be collected. The machine could then automatically take pre-emptive steps to minimize the acquisition of bad data, such as by repeating trials, discounting suspect responses, pausing the test, or offering encouragement (McKendrick et al. 2019, Wong, Dodge & Remington 1995), just as a human clinician would do.

There are several types of reliability checking, for example checking if the blind spot is detectable, positive and negative catch trials, errors in the recognition of the stimuli, time limits in the duration of the tests (indicating difficulties for the algorithm in finding the threshold value with minimal number of crossings of the threshold because the person is giving fluctuating answers) and the use of video camera for checking the fixation. Reliability can be checked post-hoc with, for example, the Heijl–Krakau method which detects fixation loss by measuring (false-positive) responses to stimuli presented to the blind spot, measuring false positives as responses made after the end of the current response window and measuring false negatives as failures to respond to stimuli more intense than those responded to previously. These measures are far from complete in capturing non-compliance and often "depend more on visual field status than on the patient attentiveness (Heijl et al. 2019)."

Research in this field could also focus on improving reliability of perimetric tests by creating fully autonomous (non-technician led) assessments necessary for home monitoring (Anderson et al. 2017), mass screening or assessment of fitness to drive. For example, Henson and Emuh (2010) used near-infrared eye tracking and found that certain eye-tracking parameters were related to vigilance in glaucoma patients. Similarly, Wang et al. (Wang et al. 2011) examined blinks and found that the probability of seeing was reduced when blinks overlapped with a stimulus presentation, although there was no association between overall blink rates and threshold variability.

3 Aims

The main objective of this work was to assess new ways of measuring vision related to real world tasks, with a focus on driving. The specific aims of this thesis were:

1. To establish whether the current regulations for fitness to drive in older glaucoma patients agree with what has been found in research **(Literature review)**

2. To evaluate the influence of visual impairment on cognitive testing, specifically the effect of visual field defects on neuropsychological tests that are commonly used in assessments of fitness to drive of those with cognitive impairment **(Study II)**

3. To evaluate visual functioning, cognitive functioning and driving performance in a driving simulator of glaucoma patients and healthy participants **(Study III)**

4. To evaluate novel vision tests related to real-world functional performance in a healthy population. This includes assessments of reaction time, visual search, face recognition, and compliance monitoring during tablet-based visual field testing (**Studies I, IV, and V)**

4 Materials and Methods

4.1 Ocusweep

Ocusweep was designed because ophthalmological diagnosis and disease progression can be more accurate by measuring functional visual compared to physical changes alone. Another goal was to develop one device which could measure all the important domains of the vision. Since Ocusweep is a new device, the tests used in this thesis will first be introduced. Ocusweep has been tested in internal clinical validation studies, but not all these results have been published, and not all tests have been officially validated. In this chapter, the background, technical details, and results from clinical validation studies will be presented.

4.1.1 Visual Acuity

Visual acuity can indicate refractive error and opacities in ocular media. Ocusweep measures visual acuity using the tablet, mounted on the perimeter. The background luminance is set to 200 cd/m^2. The stimuli are increasingly smaller Landolt C stimuli (Figure 4). Visual performance is measured by using the Ocusweep algorithm with limited viewing time, similar to real life situations. In this thesis, the test was performed either for near vision at 40cm distance or at 3m distance for far visual acuity. Clinical validation studies showed that visual acuity measures with the automatic Ocusweep custom algorithm yielded highly standardized, repeatable measurements with a repeated measurement correlation of 0.92 when tested with 23 cataract patients (Ocusweep Oy. 2015) and 0.88 to 0.90 in another study with 63 healthy participants (Ocusweep Oy. 2015).

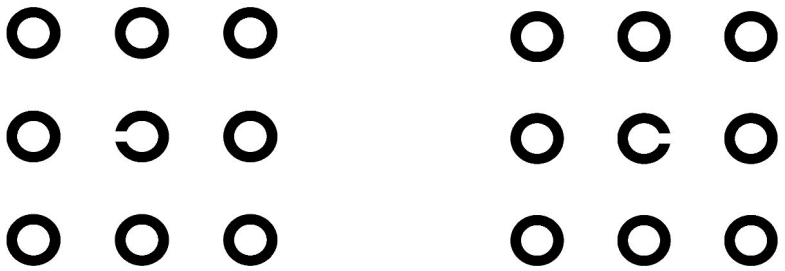


Figure 4. Visual acuity stimuli used in Ocusweep visual acuity testing. The left image indicates a Landolt C with the gap on the left side, the right image shows a Landolt C with the gap on the right side. The stimulus is always surrounded by flankers. The participant reacts to the stimulus by indicating whether the gap is on the left or right side.

4.1.2 Ocusweep time limited algorithm for central vision testing

Visual acuity and contrast sensitivity thresholds (central vision testing) were measured using a time limited algorithm developed by the Ocusweep company (Mäntysalo T., Leinonen MT. et al. Patent US 10,835,117 B2, 2020). The Ocusweep begins central vision testing from the visually most clear image and reduces the visibility according to the 4-2-1 staircase method following a correct response (Figure 5, 208). The response waiting timeout is 2.5 seconds. If the test subject does not respond within this timeout, the algorithm interprets this as if the test subject did not see the image (Figure 5, 203, 204). If the test subject responded with the wrong direction, a wrong response indicator image will be shown in place of the test optotype for approximately 2.5 seconds, and no interpretation on whether the test subject saw or did not see the image is made (Figure 5, 206). Multiple wrong responses are, however, an indication of possibly unreliable test result.

The Ocusweep central vision tests always require three test rounds. This means that at minimum seven threshold-crossings will happen during a measurement. If the results of the three test rounds yield a too high mean deviation, more test rounds will be done until the mean deviation of the three last test rounds tested is small enough (Figure 5, 207). The Ocusweep central vision tests always have at least two false negative catch trials and at least two false positive catch trials.

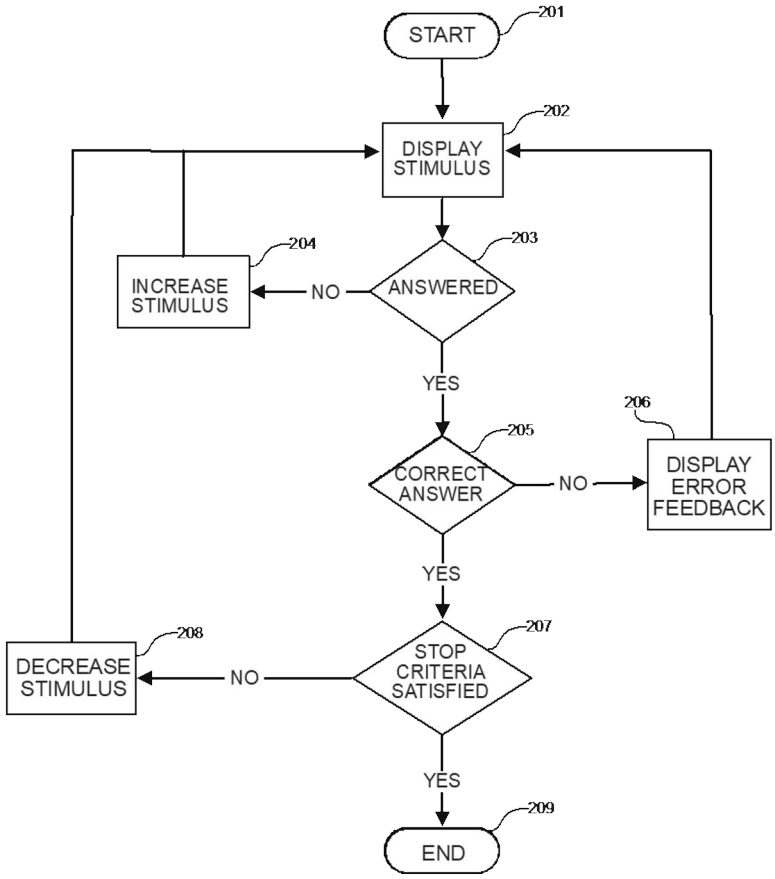


Figure 5. The time limited algorithm used in Ocusweep visual acuity and contrast sensitivity measurements. The stimulus is displayed, then the subject answers by pressing one of the two buttons. If no answer is received in 2.5 s, the stimulus visibility is increased (visual acuity – size; contrast sensitivity – contrast).

4.1.3 Contrast Sensitivity

Similar to visual acuity, Ocusweep measures contrast sensitivity using the tablet, mounted on the perimeter. The background luminance is 150 cd/m^2. The stimuli are gratings (spatial frequency 1 cycle per degree) with increasingly lower contrast levels (Leinonen & Mäntysalo 2018) (Figure 6). The thresholds are measured with the Ocusweep time limited algorithm. When comparing the golden standard contrast test (Pelli-Robson) to Ocusweep's contrast test, agreement is good except for three outliers with good contrast vision. Correlation between the results of the two tests is moderate (0.40) and the internal consistency measured by Cronbach alpha is poor (0.52) (Ocusweep Oy. 2015). In a study with 19 cataract patients scheduled for operation, repeatability of the measurements was good (correlation 0.86).

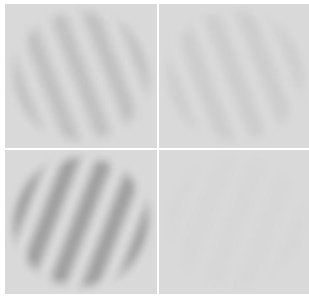


Figure 6. Contrast sensitivity is measured on the Ocusweep by using gratings. This figure shows four different stimuli, at four different contrast levels. The participant responds to this type of stimulus by indicating which way the gratings lean – left for the top two gratings, right for the bottom two gratings.

4.1.4 Standard Automated Perimetry on Ocusweep

Ocusweep also has the capability to measure the visual field, similar to Standard Automated Perimetry (SAP). In various locations in the visual field a light will be shown to which the participant responds by pushing a button. When there is no button push, it means that the light was not perceived. The locations of the Ocumap the visual field that Ocusweep uses are displayed in Figure 7.

The clinical validation study compared Octopus measurements with Ocusweep's SAP using the Ocumap grid in a heterogenous group of 30 participants. The difference between the threshold values of Ocusweep and Octopus was calculated, and the blind spot was omitted from the results. The mean of the differences of thresholds (Ocusweep – Octopus) was -0.54 dB (std 0.78 dB) and the mean of standard deviations was 3.47 dB (std 1.15 dB). Therefore, in this thesis, Ocusweep's SAP was used as a baseline measurement for visual field integrity.

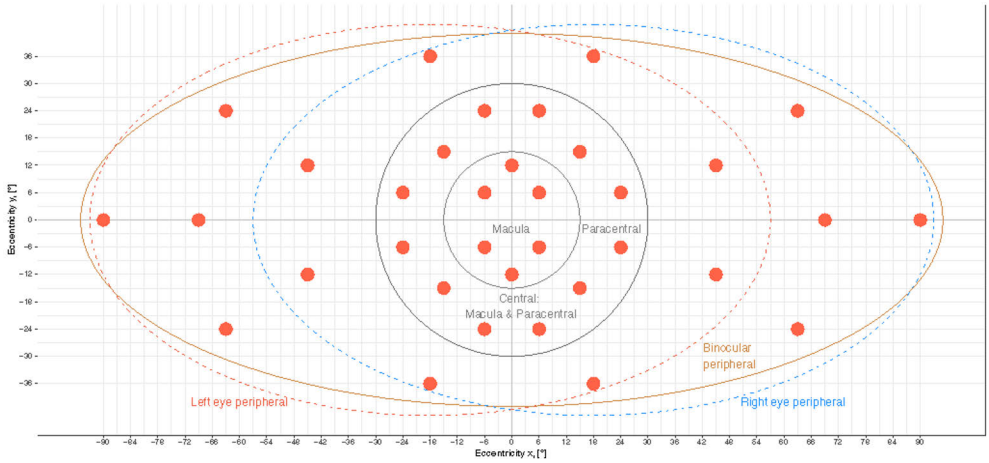


Figure 7. An overview of the grid locations used for standard automated perimetry testing on the Ocusweep device.

4.1.5 Reaction Time Perimetry

Reaction Time Perimetry (RTP) is similar to SAP but additionally takes reaction time into account by using a multifixation perimetry method and foveal recognition task with 36 peripheral suprathreshold saccade triggering stimuli across visual field spanning horizontally ± 78° and vertically ± 36° (Figure 8). By including reaction times to responses to stimuli across the visual field, this measure can be seen as a combination of vision and cognition. Therefore, this test may be more related to real-world functioning than a common visual field assessment (Leinonen M., Mäntysalo T. 2018). This test does not need to be done in a dark room using a chin rest, as the luminance of the stimulus is adjusted based on the environment. The test is explained in Figure 9. By combining reaction times, weighted reaction times for different regions plus carefulness in responses, the Functional Ability Score (FAS) of the RTP test can be calculated. For good performance on RTP the visual field has to be normal, with normal eye movements and quick visual processing speed.

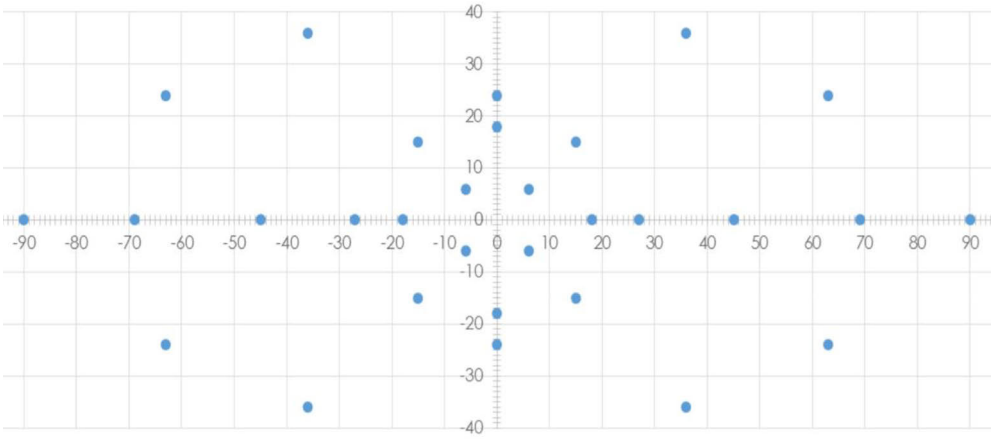


Figure 8. An overview of the grid locations used in Reaction Time Perimetry testing on the Ocusweep device, from -90 to +90 degrees eccentricity horizontally, and -40 to +40 degrees eccentricity vertically.

A study with 34 healthy participants has shown that the RTP method with threshold bracketing can be used for assessment of visual field sensitivity. Sensitivity of the visual field corresponds to SAP, as RTP and SAP visual field index values in the central (< 30°) visual field are very close to each other with a mean difference of 1.4 dB (SD = 1.0) in Mean Sensitivity (MS). The difference and variability in single location sensitivities is larger. RTP visual field measurements are repeatable, especially in global visual field indices, with a MS difference of 0.8 dB (SD = 2.9).

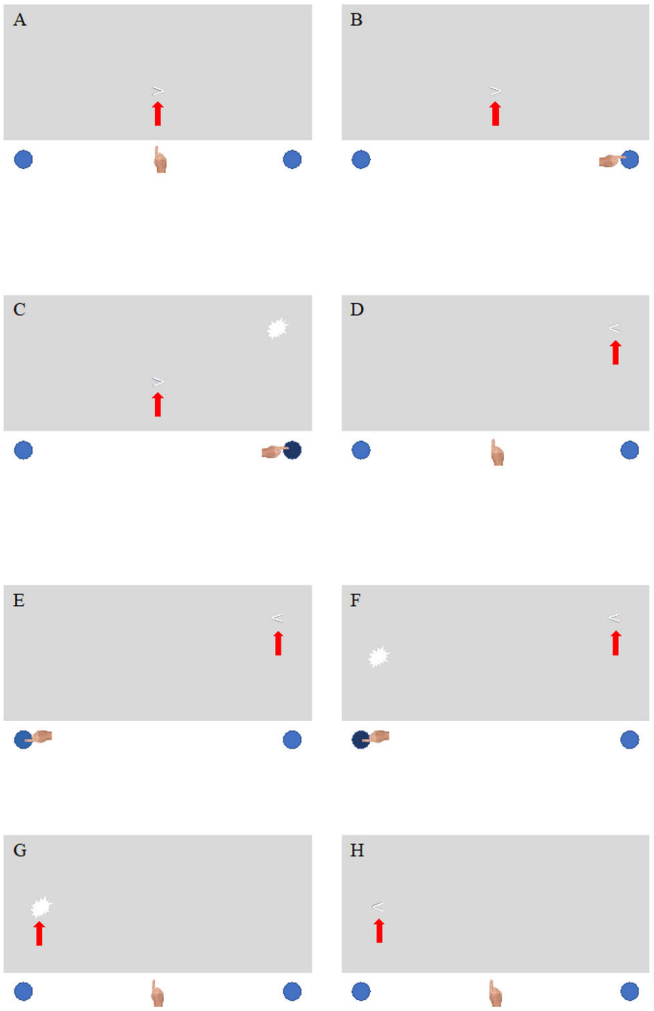


Figure 9. A visualization of different steps in the Reaction Time Perimetry Test. First, a white arrow (>) is displayed in the centre of the perimeter surface as a foveal recognition task (A). The red arrow indicates the location of the fixation, after the arrow has been found. When the white arrow is recognized, the participant responds by pushing the button on the side that the white arrow points to, in this case the right side (B). Immediately when the button is clicked, there is a flash (saccade triggering stimulus) on the surface of the perimeter (C) on a location on the RTP grid presented in Figure 8. With the help of saccades, the eyes then move to the flash location, where a new white arrow will be displayed (D) as a new foveal recognition task. This is the start of the next trial, after which the participant again presses the button on the side that the arrow points to, now the left side (E). Again, a flash (saccade triggering stimulus) will be presented somewhere in the visual field (F), the eyes move to the new location (G) where the new arrow for foveal recognition is presented (H).

4.1.6 OcuRT

To get a better evaluation of processing speed, a reaction time test should also have a way to distinguish between valid and invalid measurements such as slow reaction times due to decreased attention. OcuRT is a new test (Mäntysalo et al., 2019) and the patent is part of this thesis (see publication I). The stimulus in this test is a simple and easy to see very large Landolt-C with soft cosine filtered edges and a stimulus-background contrast of 43% to minimize spatial requirements of the visual system and maximize temporal performance (Figure 10). The radius of the C is constant (39.4 mm), resulting in a stimulus size of 11 degrees of visual angle at 40 cm distance. This way, also persons with low vision can see the stimulus, and an invalid result may indicate the inability to direct attention to the test. To take the test, the participant is seated in front of an Ocusweep device standing on a table with a tablet at approximately 40 cm distance at eye-level. The test is done binocularly. Participants are instructed to push the left button when the opening in the stimulus was on the left side (mirrored Landolt C) and the right button when the opening was on the right (Landolt C). They were also instructed not to guess. After a correct answer, the next stimulus was shown after a short delay of variable length (Figure 10). Following a wrong answer, a red circle appeared which was feedback to the participant that they made a mistake.

The time delay of the radio link between the remote and the Ocusweep is constantly monitored and kept between ±10 ms (SD = ± 0.2). Every next item is presented after a variable delay time. This delay time adapts based on the speed of answers from each participant, so that those with fast reaction times have shorter delays than those with slower reaction times. For all participants the shortest possible delay is 100 ms and the longest lies between 400-560 ms. After three measurement rounds the test was terminated and the shortest median reaction time of six consecutive correct answers was calculated.

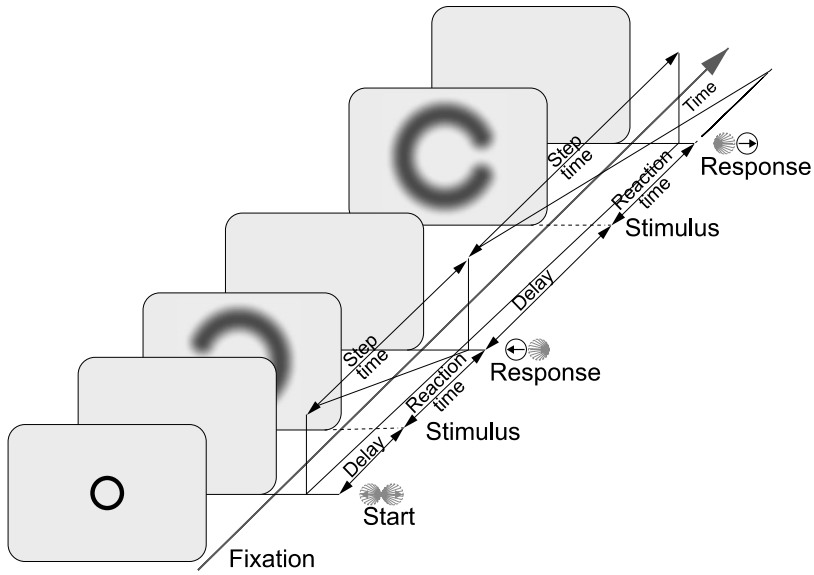


Figure 10. Design of the OcuRT test. The participants focuses on the circle, there is a delay of random duration, and then the stimulus is shown. The step time consists of the delay time plus the reaction time. After the response to the previous stimulus, the next stimulus presents with a random delay duration between 100 and 560 ms.

4.2 Evaluation of a new reaction time test (Ocusweep® Reaction Time Test) for assessing visual performance and attention (Study I)

4.2.1 Participants

Study I, the evaluation and validation of OcuRT, included 59 participants with a median age of 43 years (IQR = 31.5, 51.0). Of the included participants, 46% were male and had a median of 15 years of education (IQR = 12.3, 16.0). The Montreal Cognitive Assessment (MoCA) (Nasreddine 2004) was used to ensure cognitively healthy participants. No subjects were excluded based on cognitive impairment. Further exclusion criteria were history of neurological diseases or a (corrected) visual acuity lower than 0.5 decimal as measured with both Ocusweep visual acuity tests and a Landolt-C wall chart.

4.2.2 Methods

The OcuRT test was done four times in this study. The first three times were repeated measures of OcuRT (OcuRT1, OcuRT2 and OcuRT3) to analyse the repeatability

and learning effect of the test. The fourth time OcuRT was done with an induced cognitive load. Cognitive load refers to the strain a task can have on the working memory. According to the cognitive load theory, tasks take up a certain space in the working memory, which leaves less resources for other tasks (Sweller 1988, Miller 1956). In this study, cognitive load was induced by asking the participant to count back from 200 in steps of seven, out loud, while completing OcuRT.

The variable delay time was used to detect answers slower than the subjects' true response time and to ensure the subject was reacting to the appearing visual stimulus. The total time of stimulus presentation and the following response, including the induced delay before the stimulus presentation, was named step time (Figure 10). An indication of validity was calculated from the correlation between the total step time and delay. It was expected that, when the subject correctly directed attention to the stimulus, the length of the delay between two stimuli had no effect on the reaction time. Then, the variance in total step time was only due to the variation in the delay time because the reaction time would be similar for each trial. Consequently, the correlation between the delay and step time was high. When a person did not have full attention for the stimuli, there would have been a low correlation between step time and delay length, as the step time is now influenced by more factors than delay time only. To assess this, this measure of validity was calculated for each participant using this correlation.

Correlations between OcuRT and other tests that measure cognitive domains relevant to reaction times and attention were calculated using Kendall's Rank Correlation. This study used the Trail Making Test (TMT) (Reitan 1958) and the Vienna Test System (VTS) (Gierczuk & Ljach 2012, Prieler 2008, Schuhfried 1992) for Simple Reaction Time (SRT) and Choice Reaction Time (CRT). All VTS test scores were converted to percentile scores that are corrected for age, gender, and education.

The order of testing was the same for each participant. OcuRT1 was the first, followed by VTS SRT and CRT. After that, OcuRT2 was completed, followed by the MoCA test. Lastly, OcuRT3, the TMT and OcuRT with cognitive load were completed. This order was chosen to diminish the learning effect of OcuRT repetitions by doing something else between measurements, as well as reduce boredom that can be induced by doing repeated tasks.

4.2.3 Analysis

All relevant variables were tested for normality using Shapiro-Wilk normality tests and appropriate tests were used for the following analyses. The reaction times followed a normal distribution, so mean reaction times, duration of the test, errors, and the number of tests below the validity cut-off value were established. The cut-

off for valid measurements was set after exploring the data visually. To evaluate cognitive load, the mean reaction times were compared to the mean reaction time on the first valid measurement of OcuRT. The number of errors in the cognitive load were be compared to the errors in OcuRT1, OcuRT2 and OcuRT3. Cloud data was compared to mean reaction times in the current study. Results from OcuRT's first valid reaction times were compared using Kendall's tau rank order correlations to the Trail Making Test and Vienna's reaction time tests.

4.3 OcuDrive (Study II & III)

The first aim of this study was (1) to evaluate whether it is warranted to use vision-fair testing in glaucoma patients. Vision-fair testing refers to tests that aim to measure cognition where the requirements on the visual system are minimal. In glaucoma this means no items or stimuli in the peripheral visual field. We hypothesize that visual field defects interfere with scores on neuropsychological tests that use visual items that are spread across the visual field. A second hypothesis is that on tests with peripheral stimuli, severe visual field defects are related to lower scores than mild visual field defects. The second aim (2) is to evaluate driving performance in a group of older glaucoma patients compared to an age-similar healthy group and the third aim (3) is to identify if there are visual and cognitive measures that can distinguish between safe and unsafe drivers. In this study, we focused on older current drivers with a valid driving license and recent driving experience.

4.3.1 Participants

For Study II, 19 glaucoma patients and 19 healthy participants were included. The glaucoma patients were recruited through the University Medical Center Groningen. Participants in the healthy group were recruited in Groningen and Leeuwarden using flyers. Exclusion criteria were motor disorders and medication use that prohibits driving following the ICADTS (The International Council on Alcohol, Drugs and Traffic Safety) III criteria. This study was conducted in accordance with the Medical Ethical Committee of the University Medical Center Groningen (UMCG). Participants signed informed consent before starting the study. This study is in accordance with the Declaration of Helsinki. For all participants the inclusion criteria were age over 65 years old, current drivers, a binocular visual acuity of 0.5 decimal or better (logMAR 0.3 or lower) while wearing their habitual correction and speaking Dutch. For the glaucoma group, all disease stages and visual field defect locations were included. The sample of 19 glaucoma patients included both patients with monocular visual field defects and patients with binocular visual field defects, as well as different stages of the disease (Table 3).

Table 3. Demographic and visual function characteristics between glaucoma patients and healthy controls. Age, gender, education, and general visual acuity did not significantly differ between groups. Binocular defects were defined as a deviated in any location, in either eye. There were no requirements for minimal defects. Glaucoma patients were significantly more likely to avoid certain situations.

Participant characteristics	Glaucoma (n=19) Median (IQR)	Healthy (n=19) Median (IQR)	p-value*
Age (years)	74 (72.0, 77.5)	72 (70.5, 75.0)	.19
Gender (% male)	63%	74%	.49
Education (years)	15 (15.0, 15.0)	15 (14.0, 15.5)	.66
Near Visual Acuity (logMAR)	0.16 (0.11, 0.22)	0.19 (0.12, 0.22)	.66
Far Visual Acuity (logMAR)	-0.05 (-0.07, 0.05)	-0.08 (-0.18, -0.00)	.50
Contrast Sensitivity (log(1/threshold contrast))	2.08 (1.93, 2.22)	2.15 (2.08, 2.18)	.42
Better Eye MD (dB)	-4.51 (-6.67, -1.05)	NA	NA
Worse Eye MD (dB)	-11.96 (-16.22, -9.22)	NA	NA
Binocular defects	79%	NA	NA
Recent experience (min. number days per week)	3 (1.5, 4.0)	3 (2.0, 5.0)	.55
Total experience (years)	52 (50.0, 57.5)	53 (50.0, 55.5)	.73
Self-reported fines (%)	16%	0%	.07
Self-reported accidents (%)	11%	11%	1.00
Avoiding situations (%)	47%	11%	.01

* Wilcoxon Signed Rank Test or Chi-squared Test

4.3.2 Methods

All participants filled in a questionnaire on driving experience, history of accidents and if they had any comorbidities. Participants were compared based on recent driving experience, defined as the number of days they usually drive per week in the last 6 months, total driving experience, defined as the number of years they hold a driver's license, the number of accidents in the past five years and fines in the past year (Table 3).

4.3.2.1 Neuropsychological tests

In this study, tests that are commonly used in research and clinical practice around fitness to drive were included (for rationale and explanation of each test, see Literature Review). These tests include the Montreal Cognitive Assessment (MoCA) (Nasreddine 2004), which is a cognitive screening test for assessing mild cognitive

impairment, the Trailmaking Test (TMT) (Reitan 1958) as a measure of cognitive flexibility and processing speed, the Benton Visual Retention Test (BVRT) (Benton 1945) as a measure of visual perception and visual memory, the Snellgrove Maze Test (SMT) (Snellgrove 2005) as a measure of attention, visuoconstructive abilities and planning a route and foresight and lastly the Digit Span Test (DST) which measures working memory without any visual items, to allow comparison of cognition between glaucoma patients and a healthy group without a visual component.

For the second part of the study in which the participants drive in the driving simulator, analysis also includes the Vienna Test System (VTS) (Schuhfried 1992) Reaction Time tests and the Useful Field of View Test (UFOV). VTS can measure simple reaction times to both auditory (SRT-A) and visual stimuli (SRT-V). Next to that, it also measures choice reaction time (CRT) that combines visual and auditory measures. All scores are converted to percentile scores that are corrected for age, gender, and education. Scores are separated for reaction time and motor time. Slower reaction times on this test might be related to slower reaction times in traffic (Piersma et al. 2016). The Useful Field Of View Test (UFOV) (Wood & Owsley 2014) is a measure of the useful visual field of the participant and depends on attentional capabilities, processing speed, executive functioning (ignoring distractors) and the visual field.

4.3.2.2 Visual function tests

For the glaucoma patients, the Humphrey Field Analyzer (HFA) 24-2 or 30-2 SITA (Swedish Interactive Threshold Algorithm) visual field measurement closest to the appointment date was downloaded from the hospital server. Visual fields measurements of both eyes were merged, taking the highest value for sensitivity in each location, and plotted onto a visual field map (Figure 11). The severity of glaucoma is usually quantified with mean deviation (MD), which is the average difference in decibel (dB) from the expected value in all measured locations of the visual field (as a rule of thumb, 0 dB means intact, more than 6 dB is moderate/severe glaucoma, and 30 dB means fully blind).

All participants completed vision tests on the Ocusweep for near and far VA. While far VA is the most common acuity measure, near VA is the most relevant when reading or doing neuropsychological tests (Dupuis et al. 2015)). Next to VA, Contrast Sensitivity, OcuRT and Reaction Time Perimetry (RTP) were completed on the Ocusweep.

4.3.2.3 Driving simulator

Fitness to drive was assessed with driving simulator performance. Two fixed-base Jentig50 driving simulators of ST Software were used in Leeuwarden and Groningen, in the Netherlands (Figure 12). The simulator consisted of an open mock-up with a car seat, a steering wheel, and pedals. Three screens were used to display the interior of the car, including the dashboard and mirrors, and the road (200° degrees in total). Simulated traffic adapted to the driving behaviour of the participants. All rides were driven with automatic transmission. Four different rides, including a practice ride until they felt comfortable to continue the study, were used to assess different aspects of driving performance. In the experiment similar scenarios were used, similar to Piersma et al. (2016).

The first ride after the practice ride is called ‘Lane Tracking’, in which the participant is driving on a winding rural road with no intersections. There was no speed limit, but participants were instructed to drive an appropriate and comfortable speed. Important measurements from this ride include speed, lane positioning, and collisions.

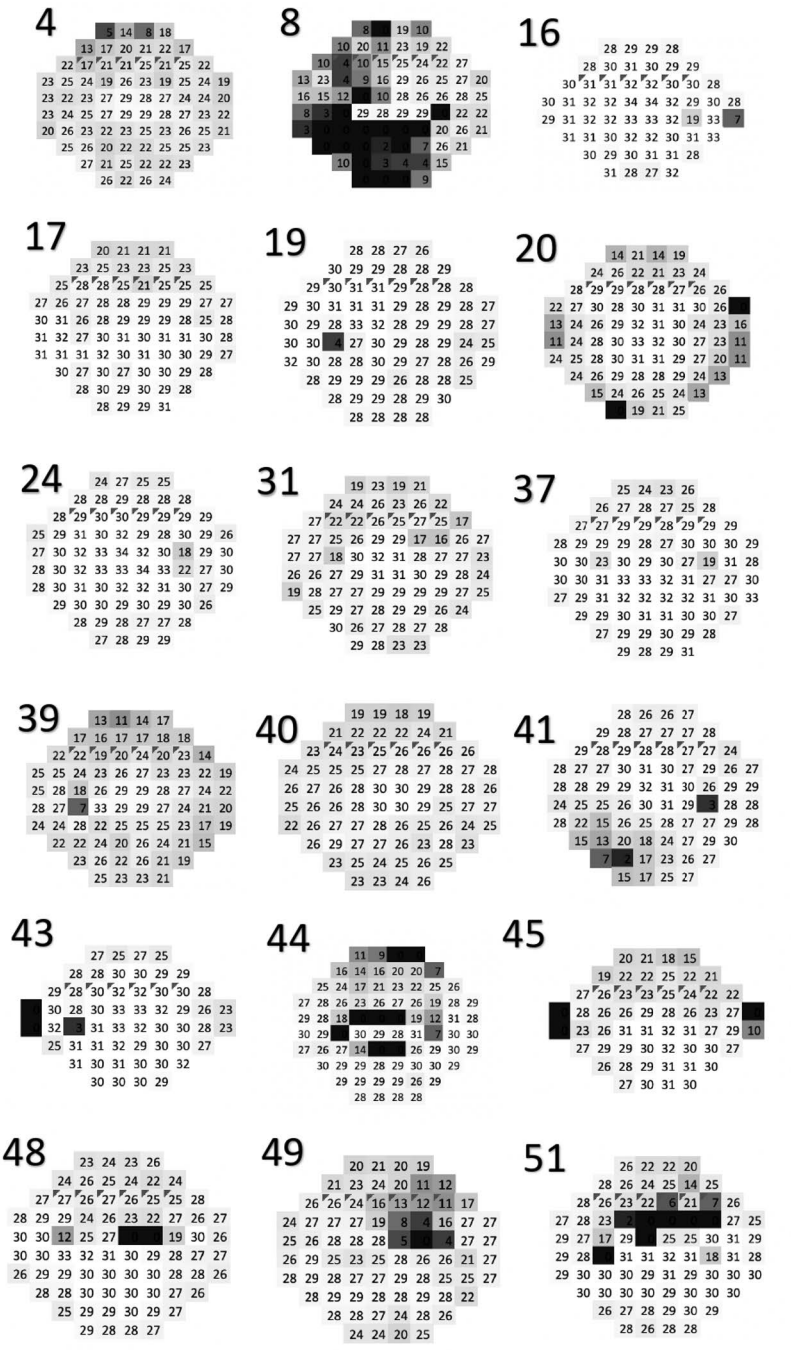


Figure 11. Merged binocular fields of the glaucoma patients in the OcuDrive study. The most recent available binocular field test from hospital servers was used. These images were made by taking the highest sensitivity for each location from both eyes and plotting the result on a grid covering the visual field. Therefore, each grid is an approximate representation of the binocular visual field of the participant.

The second ride is called ‘Intersections’. This ride was also in a rural environment, but now they encountered six intersections with different regulations. At all intersections the assignment was to drive straight ahead. The speed limits were indicated by signs, two sections had a speed limit of 60 km/h (37 mph), and two sections had a speed limit of 80 km/h (50 mph). Important measures from this ride were speed, brake reaction time to a traffic light that turns amber, braking for a car that merges onto the road from a parking spot, the number of times the participants turn their head, and collisions.

The third ride was ‘Merging’, in which the participants were required to gain speed on the ramp and then merge onto a highway (speed 120 km/h [75 mph]). After a while, the computer told the participants to go to the left lane, overtake one car, and then return to the right lane. Then, the computer told the participant to take the next exit. Measurements from this ride focused on how much the participant hindered other road users, by calculating how close the participant was to other traffic or how much other traffic had to decelerate because of the participant. Next to that, the speed, crashes, and head turns were noted. After all rides were completed, overall performance was scored by the test operator on a five-point Likert scale.

Figure 12. Driving simulator set-up. The left image shows the driving simulator used, with three of the five screens visible. On the screens, a scene was shown of which the image on the right is a screenshot during the Intersections ride of the primary screen. The primary screen is the middle screen right in front of the participant and shows what you would see through the windscreen of a normal car.

4.3.3 Analysis

First, the glaucoma group and the healthy group were compared on demographics, the questionnaire, near and far VA and CS using Wilcoxon signed-rank tests and Chi-Squared Tests. Then, the neuropsychological test scores were compared between the glaucoma group and the healthy group using Wilcoxon Signed-Rank Tests and in boxplots. The percentage of agreement between the MoCA (cut-off 26)

and MoCA-Blind (cut-off 19) was calculated to evaluate the effect of using the MoCA-Blind version in glaucoma patients. Spearman correlations between age and visual and neuropsychological measures were calculated. The effect of glaucoma severity was tested using non-parametric tests that compare the patients with none to mild visual field defects in the better eye to the patients with moderate to severe visual field defects in the better eye. Lastly, those who crashed in the driving simulator and those who did not were compared on the visual measures and neuropsychological test results. The results in this study were not corrected for multiple comparisons, since this is an exploratory study with a relatively low number of participants with a power of 0.83 using an alpha of 0.05 (McDonald 2014). To limit Type I errors, comparisons were determined a priori, and all results are communicated in this thesis. Analysis were performed using R (version 4.0.2, R Core Team 2020) and RStudio (version 1.3.1093, R Studio Team 2020).

4.4 Introducing and evaluating two new tablet-based tests of real-world visual function for visual search and face recognition (Study IV)

4.4.1 Participants

In this study, participants completed two tablet-based tests for real-world visual functioning. The goal was to describe, refine, evaluate, and provide normative control data for these two freely available tablet-based tests of real-world visual function, using a cohort of young, normally-sighted adults. To evaluate these new tests participants first completed tests for VA, CS, colour vision and working memory. After screening, 50 normally sighted adults aged with a mean age of 25 years (SD = 5) performed the two tablet tests.

4.4.2 Methods

Both tests were performed twice in one session (ABAB or BABA randomly), in order to assess test-retest repeatability. Both tests were set to run for a fixed large number of trials, allowing us to determine, post hoc, how many trials were actually required in order to obtain a statistically stable measure of performance. After testing, participants were also asked to rate the ease and clarity of each test. Both tests were run on a touchscreen tablet computer with an IPS (in-plane switching) screen measuring 26 cm by 17.3 cm at 50 cm from the screen (Figure 12). However, after starting the test viewing distance was not strictly constrained. On both tests, the screen was viewed binocularly, and participants were allowed to move their eyes freely.

4.4.2.1 Test 1 – Faces

For this test, data from a subset of 30 individuals were reported. The Faces test measured participants' ability to discriminate between four human faces. On each trial, the participant was asked to identify and then touch the odd one out. In all trials, three of the faces were identical but the face itself varied randomly between trials. The fourth face varied from the other three by parametrically manipulating photographs of real faces (Tiddeman, Stirrat & Perrett 2005). Each face had 127 feature points which could be altered by a percentage to create progressively more dissimilar versions of a particular reference image. The ears were removed on all faces. The location of the target face during testing varied randomly between each of the four positions in each trial. Participants were encouraged to respond as quickly and accurately as they can. Depending on whether the participant answered correctly or incorrectly on the previous trial, the degree of similarity between the faces was decreased or increased after every trial with a QUEST adaptive algorithm (Watson & Pelli 1983).

The outcome measure was the threshold of the smallest difference between the faces that the participant could detect reliably, measured in units of dissimilarity (*d*). This was defined mathematically as the Maximum Likelihood Estimate (MLE) of the mean of the posterior density function, which was computed after every trial. In Figure 13a, trials showing typical threshold values between 3–4 are shown, with bigger numbers indicating easier trials with larger dissimilarity.

4.4.2.2 Test 2 – Visual Search

The second test, Search, measured participants' ability to locate a particular object in a crowded scene ('find the matching object'). On each trial, a random image (the Reference) was presented in the centre of the screen and remained visible throughout the trial. After 1 second, 62 additional images then appeared also on the screen, spaced uniformly on a 7 by 9 grid (Figure 13b). One of these 62 images (the Target) was identical to the central Reference image. Participants were asked to locate the matching Target image as quickly as possible, and to touch it. All 63 images varied on every trial, drawn randomly from a previously described database of 2400 real world objects (Brady et al. 20008). The Target was presented once at each possible grid location over the course of 62 trials so that each possible location was tested once. The order of target locations was randomised each for each test. The outcome measure was Response Time (RT) in each location, and the overall median RT was computed as a summary measure of performance.

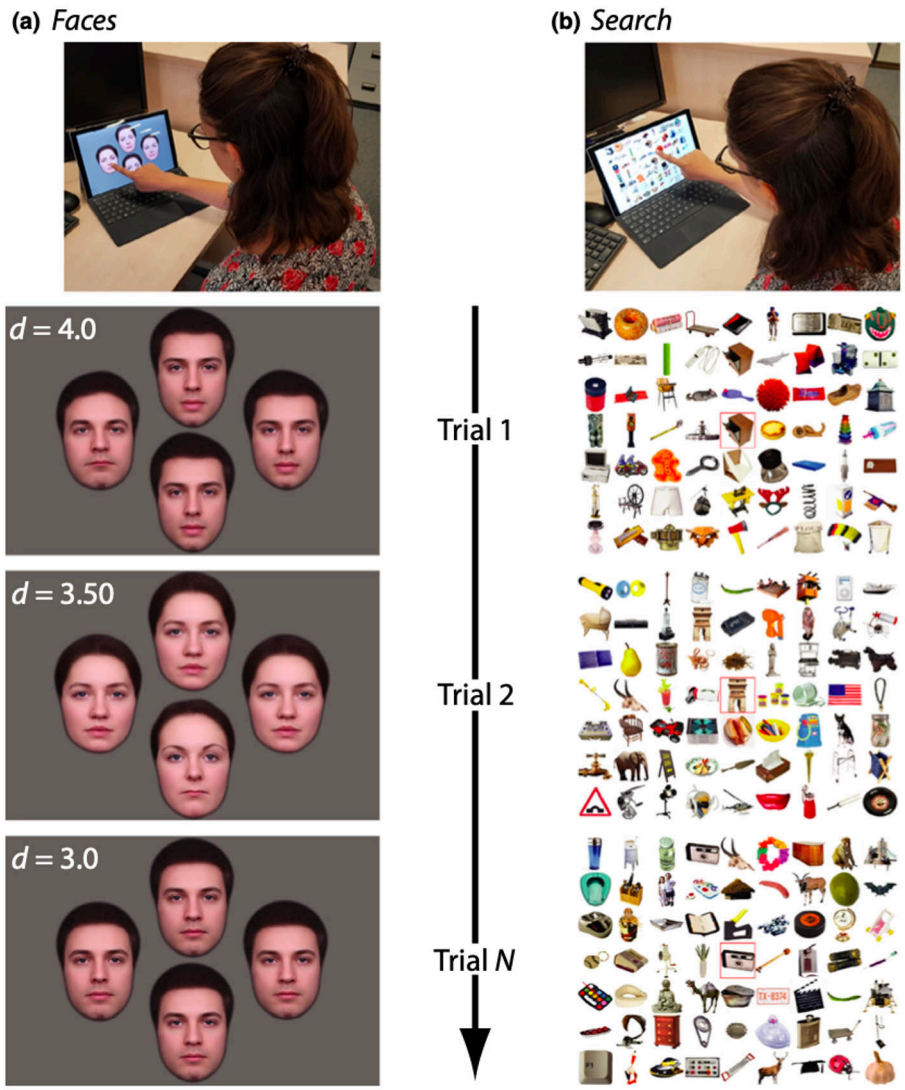


Figure 13. Setup and example trials for the (**a**) Faces test, and (**b**) Search test. In the faces test, the *d*-value indicated the similarity between the faces, in which a higher *d* resulted in more differences. For the visual search test, the reference image was always shown in the middle with a red box around it, and the target could appear in any of the other locations and was systematically varied throughout the test.

4.4.3 Analysis

Analyses were performed using MathWorks MATLAB R2016b. The Face test was programmed in Python (www.python. org) by author IW, using the OpenSesame

toolbox (Mathôt, Schreij & Theeuwes 2012). The Search test was programmed in C# by author WB.

4.5 Using a webcam to autonomously monitor compliance during visual field assessments (Study V)

4.5.1 Participants

In this study, 42 eyes from 42 adults with corrected-to-normal vision with a median age of 26 (IQR = 22, 29) years were included. The healthy participants wore their habitual correction and had normal vision, as confirmed with screening VA, CS, and colour vision. Additionally, 14 eyes from seven adults with a median age of 69 (IQR = 64, 74) years with a bilateral open-angle glaucoma (n = 6) or unilateral secondary glaucoma were examined. All patients exhibited best-corrected logMAR acuity < 0.5 in the worse eye, and none had undergone ocular surgery or laser treatment within 6 months prior to participation. Severity of visual field loss (Mills et al. 2006) varied from –2 dB (early) to –18 dB (advanced), although the majority of eyes exhibited moderate loss (median = –8 dB). An example patient's visual field is shown in Figure 14B.

4.5.2 Methods

All eyes were tested twice with monocular visual field assessments. The healthy participants performed the visual field test twice, once for each healthy eye, using a custom screen perimeter (Figure 14A), implemented on an HP Pavilion x360 15-inch laptop (HP Inc., Palo Alto, CA). The test was a modified version of the Eyecatcher visual field test (Jones et al. 2019), modified in the present work to more closely mimic conventional static threshold perimetry, most notably by employing a Zippy Estimation by Sequential Testing (ZEST) thresholding algorithm (Turpin et al. 2003), a central fixation cross, and a button press response. The software was implemented using Psychtoolbox 3 (Kleiner et al. 2007) and we used bit stealing to ensure >10-bit luminance precision (Tyler et al. 1992), with extensive photometric calibration to ensure stimulus uniformity across the display.

Unlike conventional perimetry, participants received visual feedback regarding the true stimulus location after each button press, intended to keep participants motivated and alert during testing. The output of each Eyecatcher assessment was a 4 × 6 grid of differential light sensitivity (DLS) estimates, corresponding to the central 24 locations of a standard 24-2 perimetric grid (Figure 14B; ±15° horizontal and ±9° vertical). For analysis and reporting purposes, these values were transformed

to be on the same decibel scale as the HFA: dB = 10log10 (3183.1/DLScd/m2). A summary measure of visual field sensitivity was computed by mean averaging these 24 DLSdB values, resulting in two mean sensitivity (MS) values per participant (one per test; same eye). Another common summary metric of visual field sensitivity is mean deviation. None of the present findings differed if this metric was used instead of MS. Test–retest measurement error was quantified as the absolute difference in MS between each test.

The 14 patient eyes were assessed once by Eyecatcher. The healthy participants were assessed twice; one SAP test with HFA (SITA Fast 24-2) to quantify measurement error and once by Eyecatcher. The results of the two tests were highly correlated (Pearson correlation, $r12 = 0.86$; $p < 0.001$) with no significant difference in mean score (repeated measures t-test, $t13 = 1.38$; $p = .190$). For equivalence with Eyecatcher, MS from the HFA assessment was computed by averaging across the central 24 test locations only. Test–retest measurement error was computed as the absolute difference in MS between the two tests (MSEyecatcher – MSHFA).

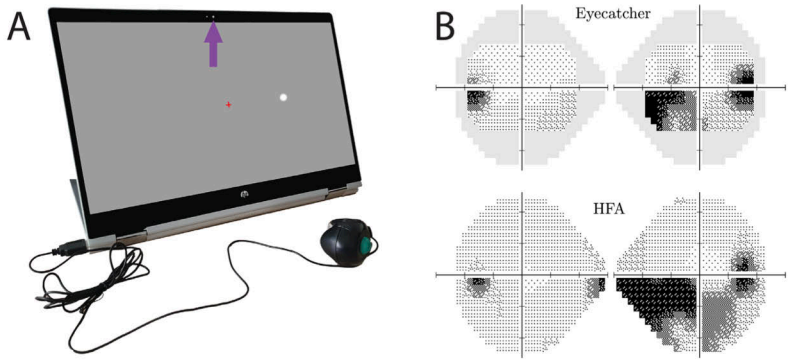


Figure 14. (**A**) A modified version of the Eyecatcher visual field test (Jones et al. 2019) was taken using the inexpensive screen perimeter. During the Eyecatcher assessment, live recordings of the participant were made via the screen's front facing camera (purple arrow). (**B**) Example measures Eyecatcher and HFA from a single participant. In all cases, only the central 24 points of the 24-2 grid were analysed when computing MS.

4.5.2.1 Biomarkers of Task Compliance

Seven potential biomarkers of task compliance were selected based on informal piloting and pragmatism and followed an initial assumption that less compliant individuals would be more likely to fidget or exhibit displeasure and could be measured without additional hardware. In the current thesis, only five out of seven biomarkers are discussed, keeping only those related to vision, where the participant was looking and response latency. The two biomarkers not discussed here are related to emotions, including *sadness* and *surprise.*

The variables were primarily derived from the video footage taken with the standard integrated camera of the HP Pavilion laptop (5 Hz with 640 × 480-pixel resolution) using OpenFace 2.0 (Baltrusaitis et al. 2018). OpenFace 2.0 has been applied previously to assess dementia (Sonntag 2018), depression (Song, Shen & Valstar 2018), and suicidal ideation (Laksana et al. 2017). It has also been used to improve automatic speech recognition (Sterpu, Saam & Harte 2018), perform video classification (Wu et al. 2017), monitor engagement with e-learning materials (Kaur et al. 2019), and inform trauma-recovery regimens (Dhamija & Boult 2017).

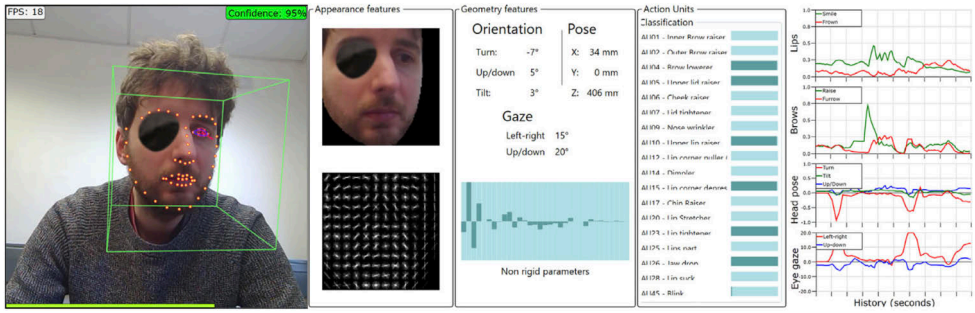


Figure 15. Biomarkers of task compliance. Various biomarkers were computed from raw video footage of the Eyecatcher assessment. The data shown here are from author P.R.J. and are for illustration purposes only.

4.5.2.1.1 Gaze Variability

Eye gaze was estimated yielding one vector of gaze coordinates (in degrees visual angle) per video frame. Gaze variability was quantified as the median (absolute) distance of every gaze point from every other gaze point.

4.5.2.1.2 Head Location Variability

The location of the head was estimated using a speed-optimized convolutional experts constrained local model (Baltrusaitis et al. 2018). This yielded one vector of location coordinates (in millimetres) per video frame. Variability in head location was computed in the same manner as gaze variability.

4.5.2.1.3 Head Rotation Variability

The rotation of the head was estimated as part of the same head pose pipeline as head location (expected mean absolute error, ~3 degrees) (Baltrusaitis et al. 2018). This yielded one vector of roll values (in degrees) per video frame. Variability in head

rotation was computed in the same manner as gaze variability and head location variability.

4.5.2.1.4 Blink Rate

OpenFace 2.0 recognizes facial expressions by detecting the intensity and presence of 18 discrete facial action units (AUs), each corresponding to entries in the classic Facial Action Coding System (FACS) taxonomy of human facial movements (Ekman & Rosenberg 2012). The primary output was 18 positive scalar intensity values, where 0 indicates the complete absence of a given AU. This yielded one value per video frame. The presence of a blink is encoded by AU45 of the FACS taxonomy.

4.5.2.1.5 Mean Response Latency

Unlike the other six biomarkers, response latency was not derived from the video footage. It was instead computed simply as the difference (in seconds) between the onset of a given stimulus presentation and the participant's button press response (τ). This was recorded only for trials where the participant responded to the stimulus.

4.5.2.1.6 Composite of All Biomarkers

A composite of all seven individual biomarkers (including surprise and sadness) was computed by standardizing each measure as a z-score and then taking their weighted linear sum where the weights were proportional to the correlation coefficient (ρ) between each measure and observed performance (Figure 15), normalized so that they summed to one.

4.5.3 Analysis

Pearson’s correlations between each biomarker and overall (test–retest) measurement error were calculated. Measurement error was quantified as the (absolute) difference in MS between two assessments, in dB. The biomarkers were computed, each yielding one value per assessment. For healthy eyes the mean of the value of both eyes was taken for each biomarker to produce a single value for each participant. For patient eyes, only one of the biomarker estimates was available (as the other measurement was done using traditional HFA).

To examine whether each biomarker could also be used in real time to identify decreased attention for the test, the final set of 24 pointwise DLS estimates from Eyecatcher were taken as the best guess of each participant's true threshold at each

location. All trials where the stimulus intensity was more than 3 dB brighter than the threshold at a given location were selected, giving 4598 trials (of 13,867 total). When the participant did not respond to these very bright trials, it was classified as a miss (false negative) and a successful response was termed a hit (true positive). Ideally, the hit rate, P(hit), for such suprathreshold stimuli should equal 1. The missed stimuli (4.6%) are possible indicators of a lapse in attention for the test.

Finally, we examined how well each biomarker predicted hits and misses on each trial. To do this, we recomputed each biomarker for each of the 3dB above threshold stimulus presentations, using only video data from that trial, and from the 20 frames (4 seconds) directly preceding it.

4.6 Availability of tests

Faces Tablet Test

https://bitbucket.org/iainrwilson/facediscrimination

Visual Search Tablet Test

https://github.com/CrabbLab/CrazySearch

Eye-catcher

https://github.com/petejonze/Eyecatcher

Results

4.7 Evaluation of a new reaction time test (Ocusweep® Reaction Time Test) for assessing visual performance and attention (Study I)

4.7.1 Characteristics of OcuRT

Distributions from OcuRT1, OcuRT2 and OcuRT3 were examined visually with a histogram and a normal quantile plot, and normality was confirmed with Shapiro-Wilk tests (OcuRT1; p=.113, OcuRT2; p=.697 and OcuRT3; p=.681). Mean reaction times and standard deviations were calculated for all three repetitions of OcuRT (Table 4) and were plotted in a scatterplot with integrated boxplots (Figure 16).

Table 4. Characteristics of the three repetitions of OcuRT. The reaction times show a small learning effect, where in the cognitive load condition the reaction times were the longest, followed by OcuRT3, OcuRT2 and OcuRT1. De number of errors followed a similar pattern of improvement, with the most errors in the cognitive load condition, followed by OcuRT1, OcuRT2 and OcuRT3.

	OcuRT1	**OcuRT2**	**OcuRT3**	**Cognitive Load**
Reaction Time (ms) (mean, SD)	464 (49)	450 (41)	441 (34)	636 (533, 843)
Duration (s) (median, IQR)	50 (48, 54)	48 (46, 51)	48 (46, 50)	87 (71, 109)
Errors (N total, mean %)	65 (2.44%)	45 (1.69%)	58 (2.18%)	231 (7.68%)
Below validity cut-off (*N*)	9	4	2	59

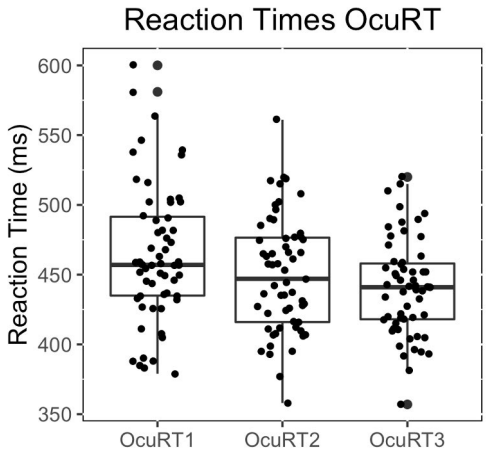


Figure 16. Boxplots showing the reaction times of all participants in OcuRT1, OcRT2 and OcuRT3. The means in the boxplots show a small learning effect, which means that each repetition of OcuRT was slightly quicker than the previous trial.

A repeated measures ANOVA (Analysis of Variance) determined that the mean reaction times were significantly different for the repetitions of OcuRT ($F(1, 58) = 12.66$, $p < .0001$), such that OcuRT1 had the longest reaction times and OcuRT3 the shortest. Pairwise comparisons revealed that OcuRT1 was significantly faster than OcuRT3, $t(113) = 3.0$, $p = .003$). The duration of the test (Table 4) was significantly shorter in OcuRT3 ($p = .003$) and OcuRT2 ($p = .002$) than in OcuRT1 as assessed with the Wilcoxon Signed Rank Test. OcuRT2 and OcuRT3 did not differ significantly ($p = .925$). A Wilcoxon Signed-Rank Test showed that participants made significantly more errors in OcuRT1 than in OcuRT2 ($V = 342.5$, $p = .020$). There were no significant differences in the total number of errors between OcuRT1 and OcuRT3 ($V = 609$, $p = .286$) and OcuRT2 and OcuRT3 ($V = 230$, $p = .144$).

Consistency through repetitions was assessed by calculating correlations between the three repetitions. The correlation between OcuRT1 and OcuRT2 was r (57) = 0.62 ($p < .001$), between OcuRT2 and OcuRT3 the correlation was r (57) = 0.79 ($p < .0001$) and between OcuRT1 and OcuRT3 it was r (57) = 0.58 ($p < .001$). The ICC (Intraclass Correlation Coefficient) was calculated to assess the agreement between OcuRT1, OcuRT2 and OcuRT3. There was moderate absolute agreement between all three trials, using a two-way mixed effects model, absolute agreement and single rater unit (kappa = 0.57, $p < .001$). There was good absolute agreement between OcuRT2 and OcuRT3, using two-way mixed effects model, absolute-agreement and single rater unit (kappa = 0.76, $p < .001$). Scatterplots and Bland-Altman plots for each repetition pair (Figure 17) showed moderate to good agreement between all repetitions of OcuRT, and the smallest difference was visible

between OcuRT2 and OcuRT3. Example data (Figure 18) showed the reaction time to each stimulus in OcuRT1, OcuRT2 and OcuRT3 as a function of the length of the delay. Each participant illustrates different effects seen in the dataset. Some participants experienced a learning effect, where OcuRT3 was faster than OcuRT1 (subject 2). Then there was an effect of delay; some participants had quicker reaction times when the delay was long compared to when the delay was short (subject 1). There were also participants who experienced neither (subject 25) or both (subject 11) effects.

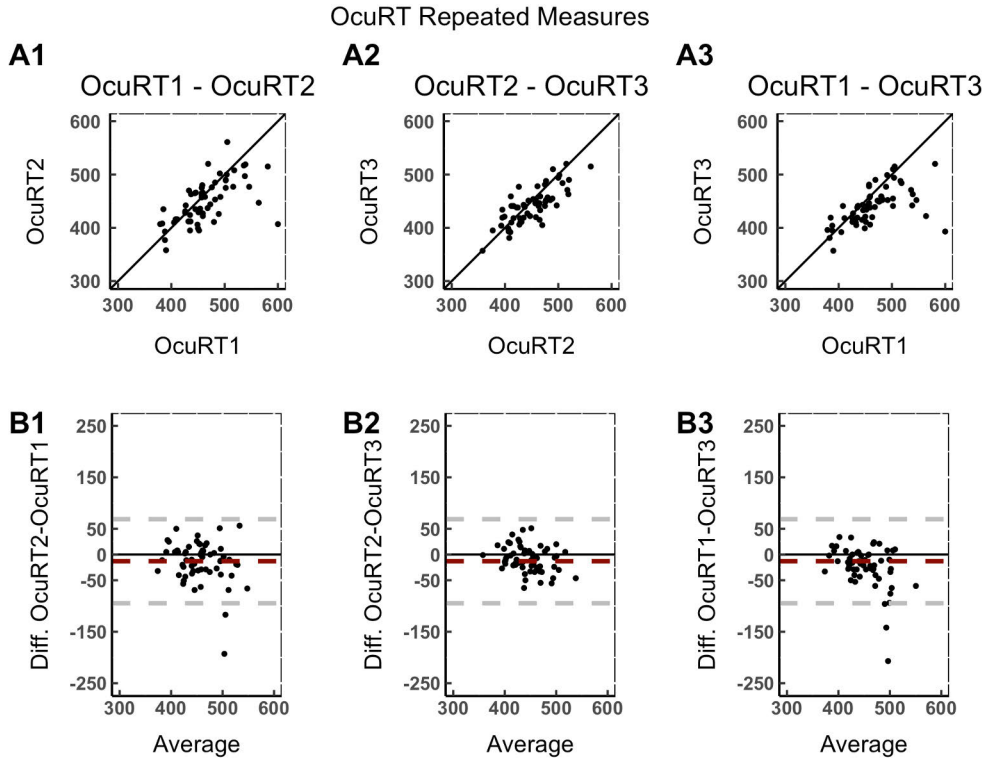


Figure 17. Scatterplots (**A**) and Bland-Altman graphs (**B**) for OcuRT1-2, OcuRT2-3 and OcuRT1-3 show that the differences in reaction time between test trials are small. The red line indicates the mean difference between two trials, the grey lines show the 95% confidence interval of the mean of the difference.

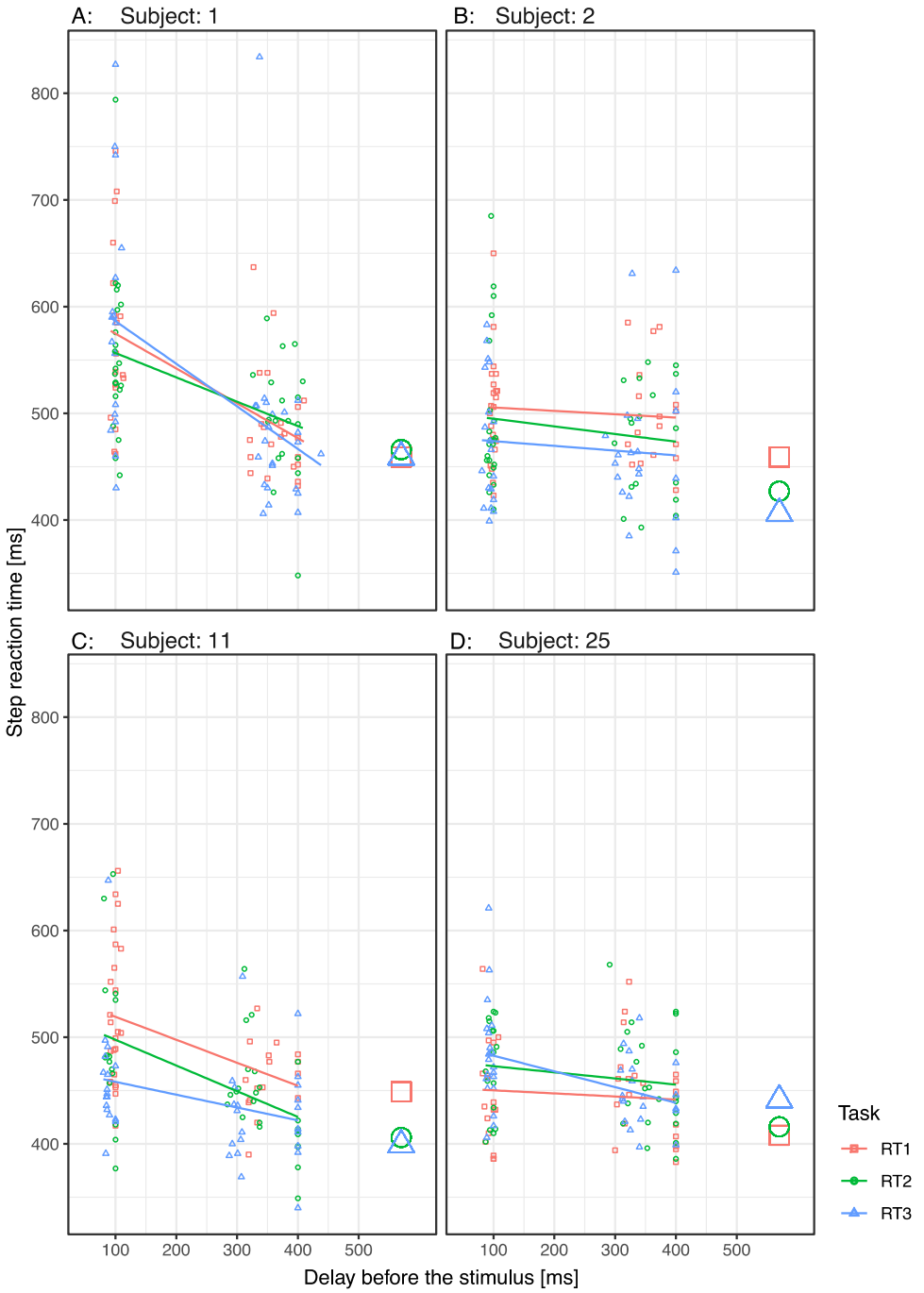


Figure 18. Example data from four different participants on all three repetitions of OcuRT. The first repetition is shown by dots in red, the second in green and the third in blue. The effect of delay is indicated by the lines. A line that is highest on the left side indicates slow responses after a short delay. A line that is horizontal indicates that the delay of the stimulus did not influence reaction time.

A validity index was calculated as the correlation between step time (sum of the delay and reaction time) and delay. A cut-off of 0.70 correlation has a sensitivity of 100% and a specificity of 95% when OcuRT3 is evaluated with the validity calculations compared to the cognitive load condition (Figure 19). This has led to the conclusion that a correlation of 0.70 or more indicates a valid result. For further calculations, the first valid measurement of each subject was used.

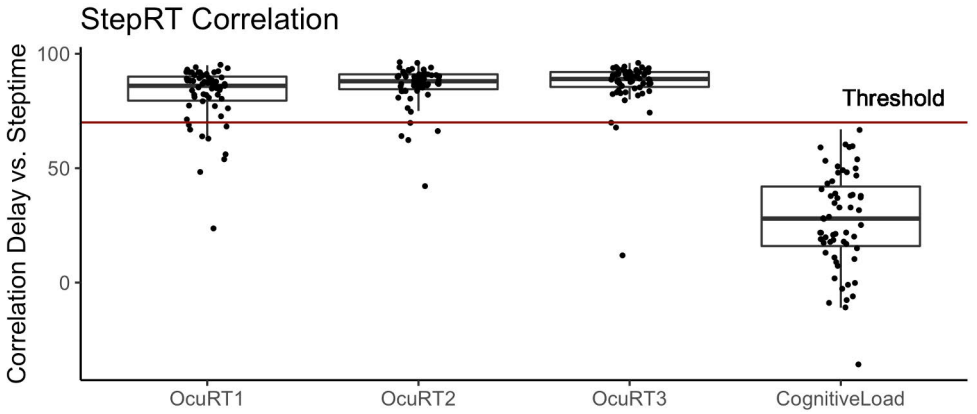


Figure 19. All four repetitions of OcuRT are plotted with the correlation between the delay and the total step time on the y-axis and the test repetitions on the x-axis. This plot shows that most participants in OcuRT1, OcuRT2 and OcuRT3 have a reliability index higher than the cut-off of 0.70, indicated with the red line, and lower than the cut-off value in the cognitive load condition.

In the cognitive load condition, participants were counting backwards during the test. Therefore, this trial can be seen as a simulation of attention problems. A Wilcoxon Signed-Rank Test showed that the reaction times in the first valid, age-normalized OcuRT result differ significantly from the cognitive load condition (V= 0, $p < .0001$), indicating a significant delay in the cognitive load condition. Figure 20 shows the reaction times in each condition for all participants. In the cognitive load condition, most subjects were about 500 to 1000 ms slower than in normal OcuRT trials, but there are approximately seven subjects who were much slower and were delayed up to 2500 ms due to counting.

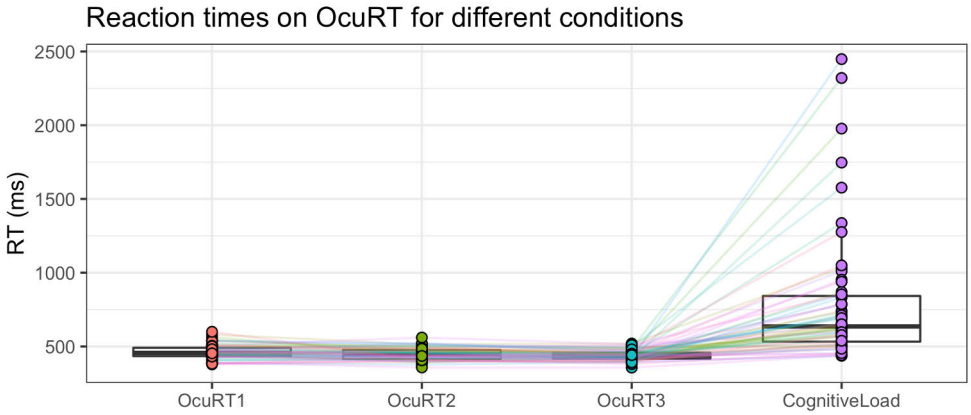


Figure 20. Reaction times on OcuRT1, OcuRT2, OcuRT3 and the cognitive load condition for all participants. The reaction times on OcuRT1, OcuRT2 and OcuRT3 are relatively similar to each other and have a smaller range compared to the cognitive load condition.

The cognitive load condition (V = 20, p <.0001) takes significantly longer to complete and participants made significantly more errors (V = 201.5, p < .0001) compared to OcuRT first valid result.

4.7.2 Comparison to cloud data

OcuRT results from the Ocusweep cloud database were downloaded and compared to the current sample of OcuRT tests. The cloud data contains 15642 anonymous measurements of OcuRT, taken from the vision assessments of the customers of Finnish optician shops (Figure 21). The cloud data was split into bins based on age (18-30, 30-40, 40-50 and 50-60 years old) before the mean reaction times were calculated for each bin. In the current study, the mean age of participants was 43 years with an IQR of 31.5 years to 51.0 years. For the bins containing ages within the IQR, the mean reaction times were similar in the cloud data compared to the current study. The cloud data contains mostly measurements similar to OcuRT1, without practice. It can also be seen that age influenced reaction times as the mean reaction time rises per bin, indicating that those that are older have longer reaction times.

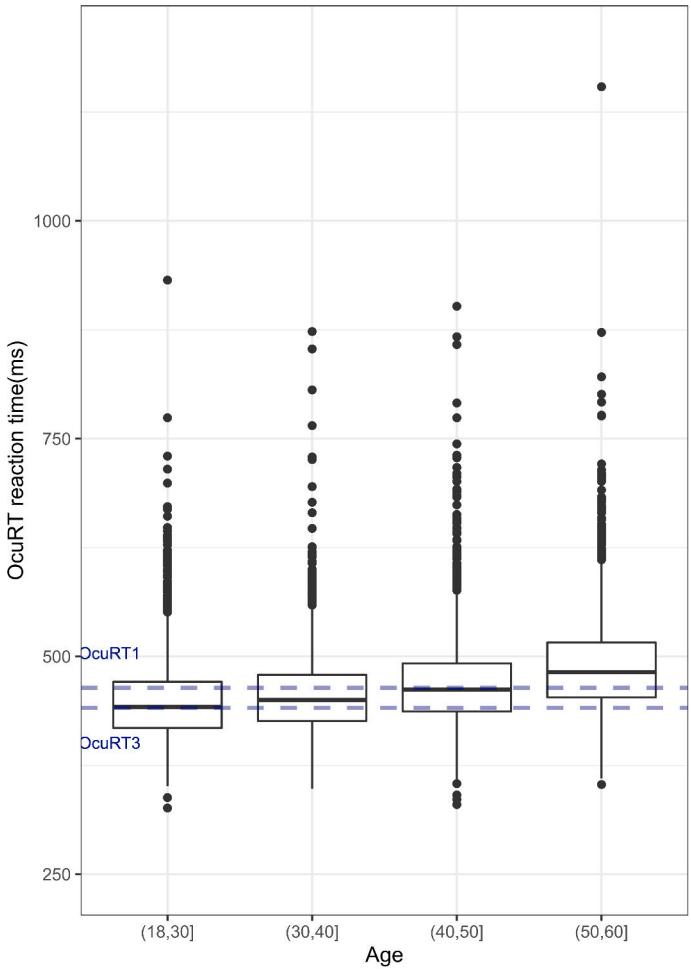


Figure 21. The data from the Ocusweep OcuRT was split into bins based on age before the mean reaction times were calculated for each bin. The cloud data is shown in the boxplots and the mean reaction times on OcuRT1 and OcuRT3 in the current study are plotted as lines.

4.7.3 Comparison to other tests

To compare OcuRT to established methods of reaction time and processing speed, the first valid OcuRT result of each participant was compared to their scores on the Trail Making Test (TMT) and two VTS reaction time tests (Table 5). The time to complete the TMT-A (τ = .27, p = .003) and TMT-B (τ = .29, p = .002) in seconds were significantly correlated to the first valid OcuRT result. Percentile scores corrected for age using the VTS norm data based on reaction times on the SRT-V (τ = .19, p = .030) and CRT (τ = .23, p = .009) were significantly correlated to OcuRT's first valid results.

A Bland-Altman plot shows the difference between CRT and OcuRT3 (Figure 22). The mean reaction time on OcuRT3 was 70 ms longer than the mean reaction

time on VTS choice reaction time test, but most reaction times were within two SD of the mean difference.

A post-hoc power analysis was done since this is an evaluation of a new test. This analysis shows it is possible to find differences of 1 standard deviation using an alpha of 0.05 with the sample size of 59 participants with a power of 0.75 for the Wilcoxon Signed-Rank Test calculating the difference between OcuRT2 and OcuRT3 and a power of 1.00 for the repeated measures ANOVA. This is a pilot study, introducing this new reaction time measure.

Table 5. Characteristics of the Trail Making Test and the Vienna Test System Reaction Time Tests.

Test	Score
TMT-A (s) (mean, SD)	24.80 (8.49)
TMT-B (s) (mean, SD)	66.88 (23.45)
Simple Reaction Time (median, IQR)	80 (52.5, 92.0)
Choice Reaction Time (median, IQR)	67 (46.5, 91.5)

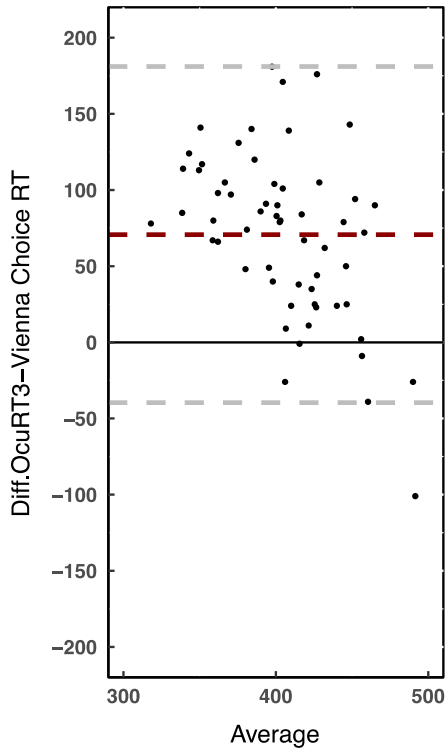


Figure 22. Bland-Altman plot showing the mean difference between Vienna Test System choice reaction time and OcuRT3. The red line indicates the mean difference between two trials and shows that the mean reaction time on OcuRT3 is longer than the VTS choice reaction time test. The grey lines show the confidence interval of the mean of the difference.

4.8 OcuDrive (Study II & III)

4.8.1 Participants

In both the glaucoma group and the healthy group, the majority of the participants was male and had received higher education (Table 3). There was no significant difference between both groups in terms of near visual acuity, measured at 40 cm distance (Wilcoxon rank-sum test, $W = 196$, $p = .66$) and far visual acuity, measured at 3m distance (Wilcoxon rank-sum test, W=157, p=.50). The cut-off point for normal CS (according to the manufacturer of the test) is logCS of 1.5. None of the participants scored below this value, and there was no significant difference between the groups ($W = 208.5$, $p = .42$). Of the participants with glaucoma, 79% had binocular visual field defects. To ensure that the healthy group did not have visual field defects, all participants performed Ocusweep's SAP. No local defects were detected, the median MD was 0.6 dB (IQR = -0.4, 1.2), and none of the healthy participants had a MD deviating from 0 by more than 2.4 dB.

4.8.2 Questionnaire

The glaucoma group and the healthy group had similar total driving experience, of 52 years for the glaucoma patients and 53 years for the healthy group. They both drove approximately 3 days per week. In the glaucoma group, 16% reported to have had a fine compared to 0% of the healthy group, but they were parking tickets and small speeding tickets (Table 3). In both groups, 11% of participants reported to have experienced an accident in the past 5 years. Of the glaucoma patients, 47% reported avoiding certain driving situations, while only 11% of the control group did (χ^2(1, N=38) =6.3, p=.01). Situations that were reported to be avoided were driving in the dark, during rush hour, and in bad weather.

4.8.3 Neuropsychological tests

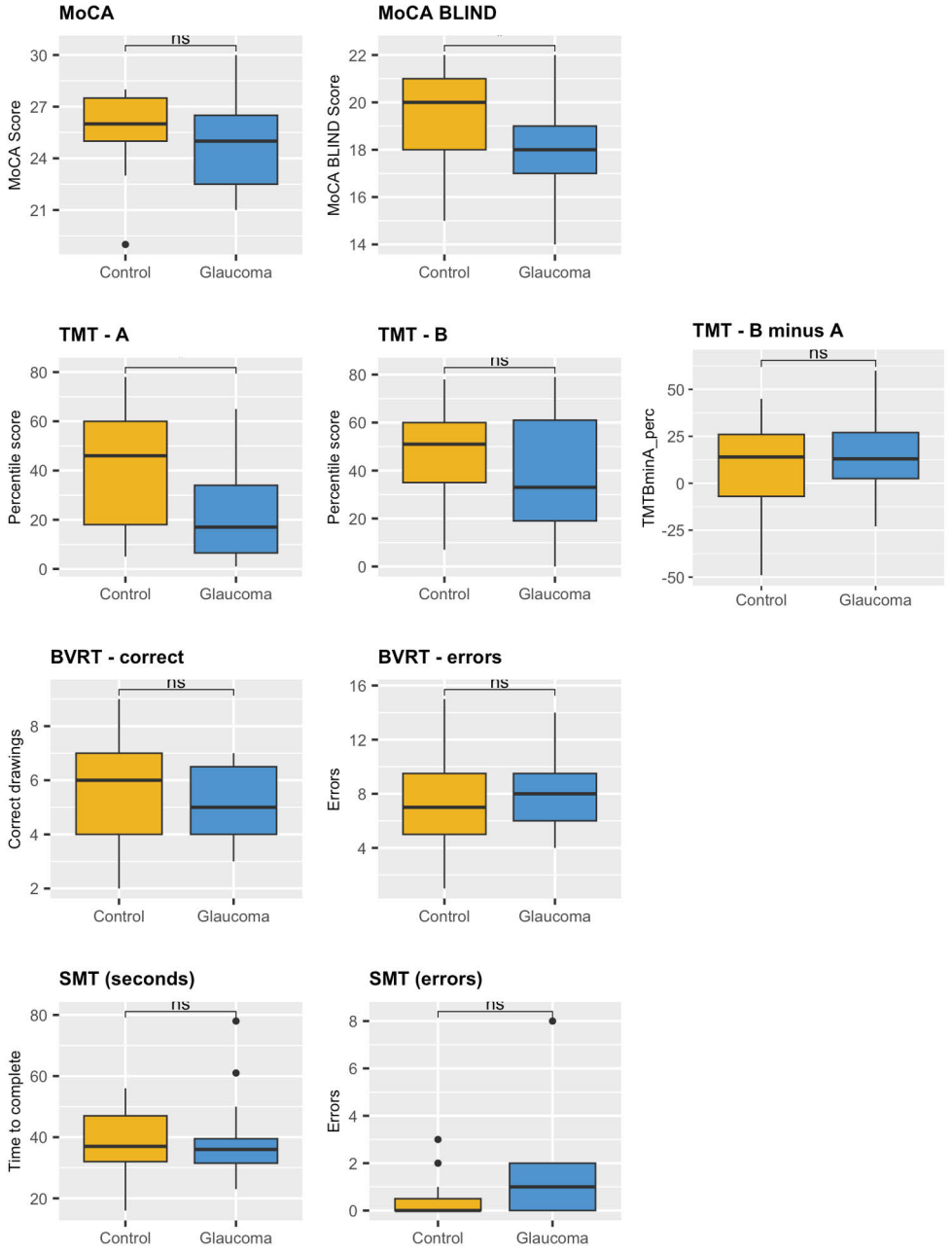
Figure 23 shows the test scores for the neuropsychological tests, for both the glaucoma and the healthy group. The boxplots showed that glaucoma patients had lower median scores on both the MoCA and MoCA-Blind, the TMT-A and TMT-B test, and they had less items correct and more errors on the BVRT and more errors on the SMT. In both conditions of the Digit Span test, they had higher median percentile scores than the control group. Table 6 presents the corresponding medians, IQRs and univariable comparisons. In the MoCA test, the median score of the glaucoma patients was below the cut-off point of 26. Removal of the visual items did not improve the performance of the glaucoma group compared to the healthy

group, but instead, they now performed significantly lower than the healthy group. The median score of the glaucoma patients on the MoCA-Blind was 18, which is the cut-off value. The agreement between the MoCA and MoCA-Blind is displayed in Table 7. A chi-square test of independence showed no significant difference between the number of participants classified as having cognitive impairment between the glaucoma group and the healthy group in the MoCA ($\chi 2(1, N=59) = 1.69$, $p = .19$). Using a cut-off of 18/22, 21% of all participants together scored below the cut-off in the MoCA-Blind. A chi-square test of independence showed a significant difference between the glaucoma group (37% of participants below cut-off) and the healthy group (5% of participants below cut-off) in the number of participants qualified as cognitively impaired by the MoCA-Blind ($\chi 2$ (1, N = 59) = 5.7, $p = .017$). There were no participants identified by the MoCA-Blind as cognitively impaired that were not classified as cognitively impaired by the MoCA. The glaucoma patients had lower median percentile score compared with the healthy participants on both the TMT-A and TMT-B (Table 6 and Figure 23), but only for the TMT-A was the difference statistically significant. When eliminating the visual component by calculating the TMT-B – TMT-A, group scores were comparable (Figure 23).

Spearman's correlations were calculated between age and all visual and neuropsychological measures. Of all these measures, far VA ($rs = 0.34$, $p = .038$) and the TMT-A ($rs = 0.45$, $p = .004$) and TMT-B ($rs = 0.46$, $p = .003$) were significantly correlated with age. The scores on the TMT were percentile scores, already corrected for age.

The VTS reaction times are displayed as percentile scores for the reaction times, where higher scores indicate faster reaction times. There were no significant differences in reaction times on SRT-V, SRT-A and CRT between the groups. The UFOV1 subtest showed a ceiling effect, indicating that most participants were able to identify the item within the shortest display time possible. In the second and third subtest, where the visual field plays a role, the glaucoma group seemed to score lower, but this difference was not significant. The difference was larger the more cognitively and visually high demanding the UFOV subtest was from UFOV1 to UFOV2 and UFOV2 to UFOV3 (Figure 23).

Neuropsychological tests



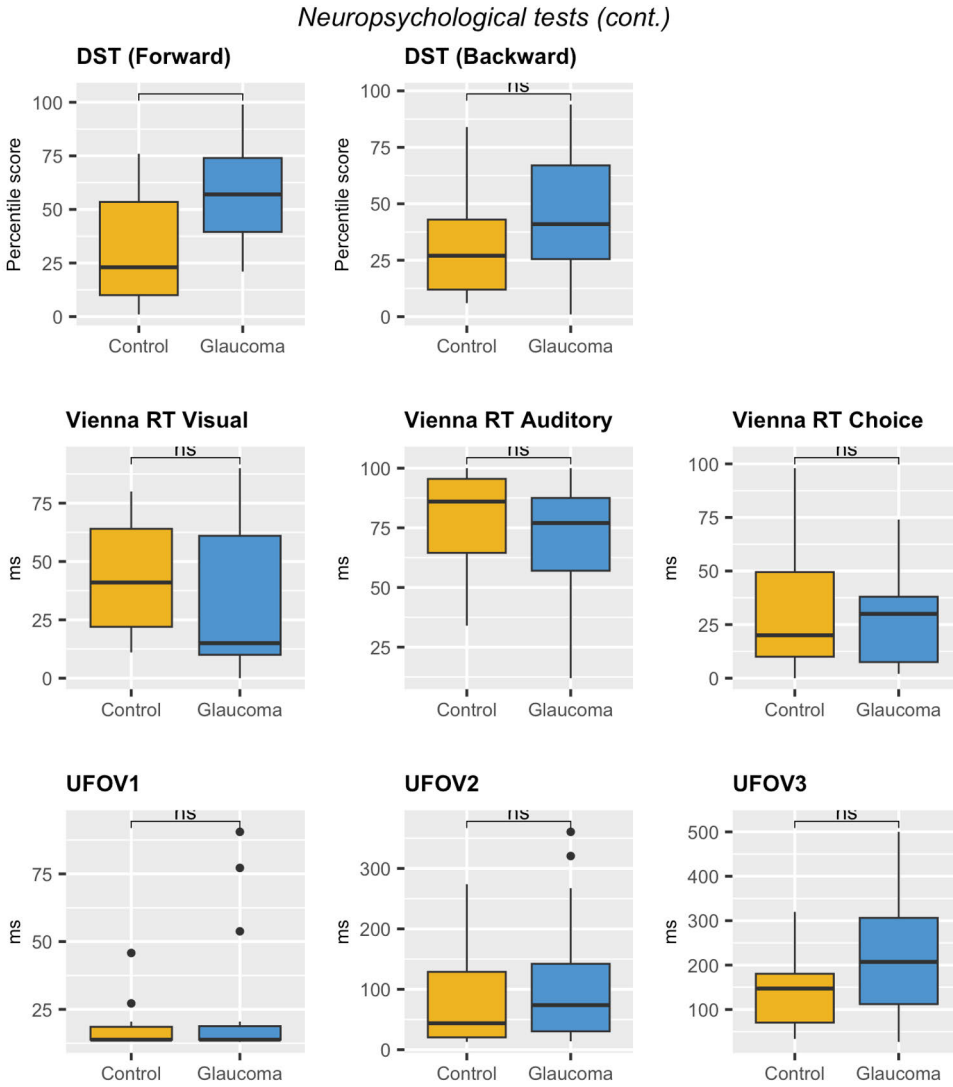


Figure 23. Neuropsychological test scores per group. For each neuropsychological test a boxplot was made showing the median scores and spread of the scores for both the control group and the glaucoma group. The difference was significant in the MoCA-Blind, the TMT-A and the DST (Forward).

Table 6. Neuropsychological test scores. Adapted from Original Publication II.

	Glaucoma Median (IQR)	Healthy Median (IQR)	p-value
MoCA	25 (22.5, 26.5)	26 (25.0, 27.5)	.21
MoCA-Blind	18 (17.0, 19.0)	20 (17.0, 19.0)	.046*
TMT-A	17 (6.5, 34.0)	46 (18.0, 60.0)	.038*
TMT-B	33 (19.0, 61.0)	51 (35.0, 60.0)	.21
TMT B-A	13 (-7.0, 26.0)	14 (2.5, 27.0)	.62
BVRT - correct	5 (4.0, 6.5)	6 (4.0, 7.0)	.32
BVRT - error	8 (6.0, 9.5)	7 (5.0, 9.5)	.67
SMT - seconds	36 (31.5, 39.5)	37 (32.0, 47.0)	.56
SMT- error	1 (0.0, 2.0)	0 (0.0, 0.5)	.078
DST - forward	57 (39.5, 74.0)	23 (10.0, 53.5)	.0015*
DS - backward	41 (25.5, 67.0)	27 (12.0, 43.0)	.30
VTS – SRT-V RT (percentile)	15 (10, 61)	41 (22, 64)	.096
VTS – SRT-A RT (percentile)	77 (57, 88)	86 (65, 96)	.23
VTS – CRT RT (percentile)	30 (8, 38)	20 (10, 50)	.43
UFOV1 (ms)	13.8 (13.8, 18.8)	13.8 (13.8, 18.6)	.86
UFOV2 (ms)	73.8 (30.4, 142.2)	43.8 (20.5, 128.9)	.22
UFOV3 (ms)	207.2 (112.2, 306.3)	147.2 (70.5, 180.5)	.09

* Indicates a significant result

Table 7. Correspondence between MoCA and MoCA-Blind. From Original Publication II.

	All	Glaucoma	Healthy
Agreement	68%	74%	63%
Disagreement MoCA impaired, MoCA-Blind not impaired	32%	26%	37%
Disagreement MoCA not impaired, MoCA-Blind impaired	0%	0%	0%

4.8.3.1 The effect of glaucoma severity

The glaucoma patients were stratified by disease severity of the better eye; early glaucoma was defined as an MD up to 6 dB and moderate/severe glaucoma was defined as MD of 6 dB or worse. The location of the visual field defects of the merged visual field of moderate/severe glaucoma patients varied among the group, but all of them had vision left in at least part of the central 10 degrees of the visual field, aiding compensational mechanisms to complete the task.

When comparing performance on the neuropsychological tests between early and moderate/severe glaucoma, glaucoma patients in the more advanced phase of the disease score similar on most items, except for the BVRT. Glaucoma patients

with moderate/severe visual field defects had – counterintuitively – a statistically significant higher number of drawings correct than those with early glaucoma and less errors (Table 8).

Table 8. Neuropsychological test scores in glaucoma patients stratified by disease severity of the better eye. From Original Publication II. The p-value indicates whether the glaucoma group and the healthy group scored significantly different to each other on each test.

	Early glaucoma (n=13) Median (IQR)	Moderate to severe glaucoma (n = 6) Median (IQR)	p-value
MoCA	25.0 (22.0, 26.0)	25.5 (23.3, 27.0)	.69
MoCA-Blind	18.0 (17.0, 19.0)	18.5 (17.3, 19.8)	.56
TMT-A	20.0 (11.0, 36.0)	9.0 (6.0, 23.3)	.54
TMT-B	30.0 (14.0, 59.0)	38.0 (32.3, 64.8)	.27
TMT B-A	10.0 (2.0, 19.0)	27.5 (19.5, 46.0)	.09
BVRT - correct	4.0 (3.0, 5.0)	7.0 (7.0, 7.0)	<.001*
BVRT - error	9.0 (8.0, 10.0)	5.5 (5.0, 6.8)	.008*
SMT - seconds	36.0 (32.0, 38.0)	36.0 (31.3, 47.5)	.86
SMT- error	1.0 (1.0, 2.0)	0.0 (0.0, 0.8)	.32
DST - forward	50.0 (41.0, 75.0)	57.0 (42.8, 60.0)	.83
DS - backward	37.0 (20.0, 56.0)	48.5 (41.3, 72.3)	.51

* Indicates a significant difference between early glaucoma and moderate to severe glaucoma.

4.8.4 Driving simulator performance

Because of simulator sickness, not all participants could complete all rides in the driving simulator. Of all 38 participants, 34 completed the Lane Tracking ride, 31 completed the Intersections ride, and 26 completed the Merging ride. There was no significant difference in simulator sickness between the glaucoma group and the healthy group in the Lane Tracking ride ($\chi^2(2, N=34)=2.53$, $p=0.28$), the Intersections ride (χ^2 $(2, N=31)=1.6$, $p=0.44$), or the Merging ride ($\chi^2(2, N=26)=1.3$, $p=0.51$).

There were no significant differences between the glaucoma group and the control group on any of the driving simulator variables, and on all three rides the group scores were close to each other (Table 9 and Figure 24). Both groups drove at a similar speed, which was often below the speed limit. No statistical analysis was done other than calculating the percentages for each type of response. However, it seems that both groups made more errors on intersections where the appropriate response was to keep driving, compared to intersections where the appropriate response was to stop (Table 10). Between intersection 5 and 6, participants had to react to an unexpected event, in which a car suddenly merged onto the road.

Participants could either evade the car by steering to the other side of the road, which includes the risk to run into oncoming traffic, brake to a full stop, or do nothing and crash into the car. Of those that steered without braking, there is no brake reaction time available. Lastly, overall performance scored by the test operator on a five-point Likert scale and the pass/fail numbers did not differ between the groups. In total, four participants crashed in the driving simulator, of which two were healthy participants and two were glaucoma patients.

Table 9. Driving simulator performance per group for all drives. From Study III.

Lane tracking (n=34)	**Glaucoma (n=16) Median (IQR)**	**Healthy (n=18) Median (IQR)**	**p-value**
Speed (km/u)	74.5 (69.1, 76.9)	75.4 (70.8, 77.0)	.93
Standard Deviation of the Lateral Position (SDLP) (cm)	29.3 (21.8, 31.6)	28.6 (26.7, 32.6)	.69
Line crossings (count)	1 (0.0, 3.0)	1 (1.0, 4.8)	.41
Collisions (count)	0	0	NA
Intersections (n=31)	**Glaucoma (n=14)**	**Healthy (n=17)**	**p-value**
Mean speed at 60 km/u (km/h)	58.1 (54.4, 60.2)	58.1 (56.3, 59.3)	.92
Mean speed at 80 km/u (km/h)	61.1 (49.8, 66.0)	65.4 (60.9, 68.0)	.20
Car response (%)			.71
Avoid	50%	53%	
Avoid + brake	33%	40%	
Brake	17%	7%	
Brake RT to merging car (s)	2.4 (2.2, 2.9)	2.2 (2.0, 2.4)	.35
Brake RT to traffic light (s)	1.9 (0.8, 2.6)	1.9 (0.9, 2.5)	.86
Head turns (count)	22 (15, 24)	20 (17, 28)	.68
Collisions (count)	1	2	.67
Merging (n=26)	**Glaucoma (n=13)**	**Healthy (n=13)**	**p-value**
Successful overtaking a car (%)	64%	50%	.51
Speed on ramp (km/u)	107.5 (91.1, 112.8)	97.8 (85.4, 104.4)	.34
Speed on right lane (km/u)	113.9 (110.5, 121.2))	118.9 (111.0, 123.8)	.92
Speed on left lane (km/u)	120.6 (114.9, 125.2)	118.0 (104.6, 125.2)	.47
Deceleration rear car (km/h)	0 (0.0, 4.2)	0 (0.0, 3.2)	.89
Minimal Time To Collision (s)	9.1 (6.6, 33.0)	10.8 (5.0, 30.0)	1.00
Head turns (count)	7 (6, 13)	8 (5, 9)	.50
Collisions (count)	1	0	.34
Total	**Glaucoma**	**Healthy**	**p-value**
Pass/fail (% pass)	80%	69%	.56

Table 10. Intersections overview. For each intersection the correct response and the percentage of participants that responded correctly at that intersection is shown. From Study III.

Intersection	Correct response	All % Correct	Glaucoma % Correct	Healthy % Correct
1 – Sign + shark teeth*	Drive	85%	100%	73%
2 – Right has priority	Stop	96%	100%	93%
3 – Right has priority	Drive	44%	42%	47%
4 – Traffic light	Stop	74%	58%	87%
5 – Sign + shark teeth*	Stop	100%	100%	100%
6 – Right had priority	Drive	48%	33%	67%

* Shark teeth are a priority indication on the pavement.

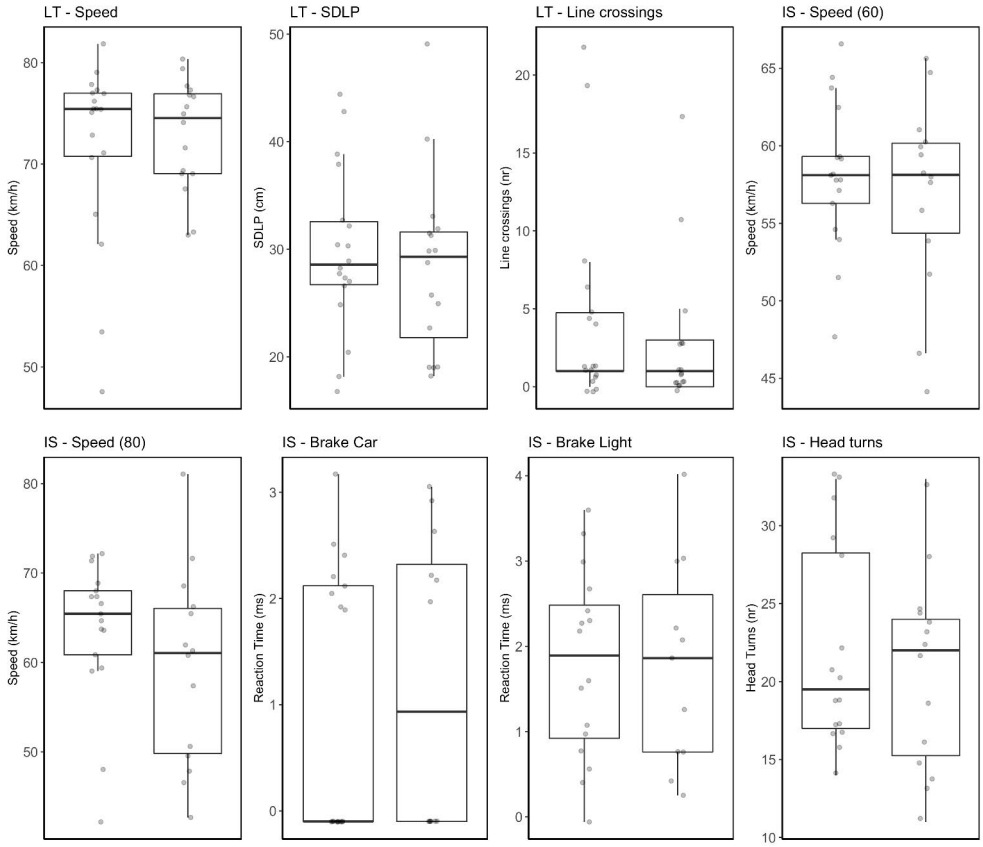


Figure 24. A Driving Simulator results for the glaucoma group and the healthy group for each variable shown in boxplots with median values.

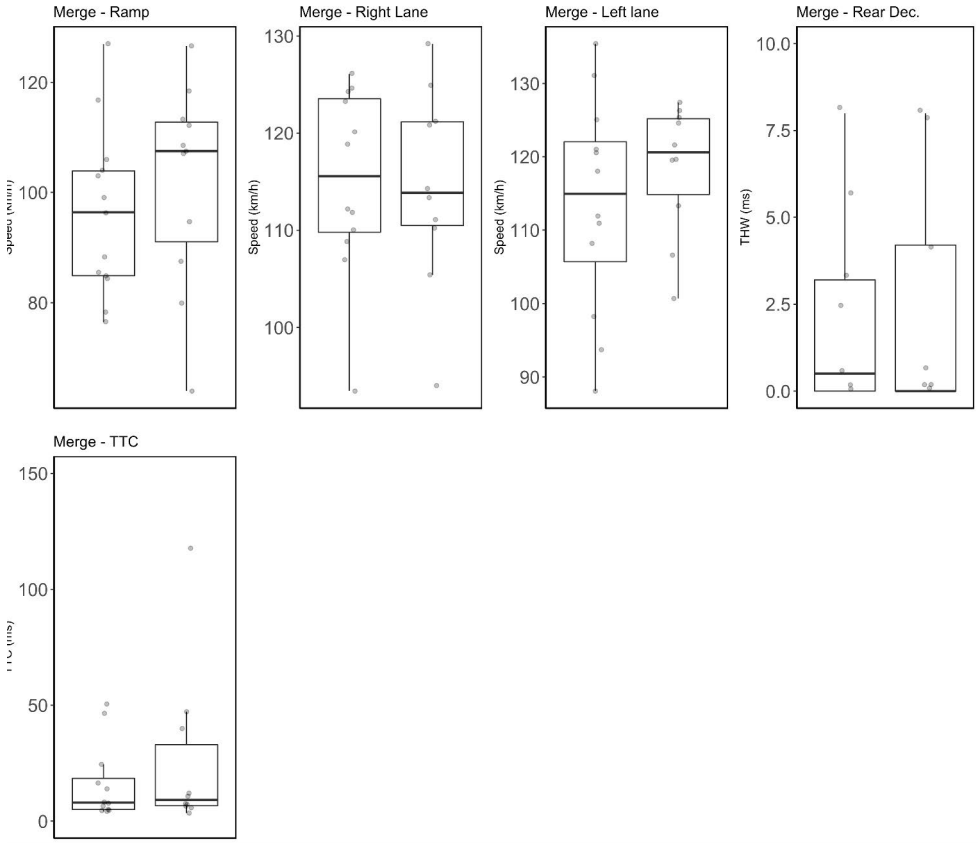


Figure 24. B Driving Simulator results for the glaucoma group and the healthy group for each variable shown in boxplots with median values [contin].

4.8.4.1 Visual functioning and driving simulator performance

Basic visual functioning, including visual acuity and contrast sensitivity, did not differ between the groups (see Methods - Table 3). Although the glaucoma patients seemed to have slower reaction times on OcuRT and the RTP test, these differences were not significant (Table 11). The Functional Ability Score, a composite measure of the RTP test, was significantly lower in glaucoma patients (p=0.03).

Table 11. Ocusweep test scores per group. Adapted from Study III.

	Glaucoma Median (IQR)	**Healthy Median (IQR)**	**p-value**
OcuRT (ms)	570 (522, 616)	541 (497, 585)	0.28
RTP (ms)	972 (939, 2023)	913 (810, 968)	0.13
RTP (Functional Ability Score)	55.0 (26.0, 72.0)	84.5 (50.0, 110.3)	0.03*

4.8.4.2 The effect of glaucoma severity and visual field defect location

The sample of glaucoma patients included glaucoma with both monocular and binocular visual field defects, as well as different stages of the disease. Visual field measurements of both eyes were merged, taking the better sensitivity value of both eyes in each location, and plotted onto a visual field map, similar to Crabb & Viswanathan (2004) (Figure 11). Six participants had relatively severe defects or defects in the central area in the merged visual fields (participant 4, 8, 44, 45, 49, & 51). Of these participants, three passed the driving simulator tests (no crashes), and three were unable to complete all drives due to simulator sickness.

4.8.4.3 Comparing those who passed or failed the driving simulator tests

Comparing the group of glaucoma patients to the group of healthy participants showed no significant differences in driving simulator performance. To find which factors, either demographics, visual, or cognitive, influenced driving performance, participants (from either group) who passed (no crashes) the driving simulator were compared to participants who did not pass the driving simulator test. For driving experience and self-reported driving behavior such as avoidance of situations and fines, there were no differences between the groups. It can, however, be noted that they all had comorbidities, including diabetes and carpal tunnel syndrome (100% versus 41%; p=.01) and slower reaction times in the OcuRT (617 ms versus 536 ms; p=.046), and the TMT-A (16th percentile versus 52nd percentile; p=.027) indicated lower percentile scores for the failed group (Table 12).

Table 12a. Showing the median scores and IQR for those that crashed in the driving simulator and those that did not. From Study III.

Patient characteristics	Pass (n = 17) Median (IQR)	Fail (n = 6) Median (IQR)	p-value
Age (years)	72.0 (70.0, 73.0)	74.5 (72.5, 75.8)	.20
Gender (% male)	71%	83%	.54
Education (years)	15 (15, 16)	15 (15, 15)	.23
Comorbidity (%)	41%	100%	.01
Near Visual Acuity (LogMAR)	0.16 (0.06, 0.22)	0.21 (0.17, 0.25)	.19
Far Visual Acuity (LogMAR)	-0.05 (-0.15, -0.02)	-0.02 (-0.04, -0.02)	.22
Contrast Sensitivity (LogMAR)	2.12 (2.08, 2.18)	2.08 (2.02, 2.22)	.81
Better eye MD (dB)	-3.05 (-6.29, -1.18)	-2.50 (-3.68, -1.31)	.71
Worse eye MD (dB)	-10.85 (-11.90, -10.09)	-18.02 (-22.67, -13.37)	NA
Binocular defects	41%	100%	.47
Driving experience (years)	52 (51, 54)	52.5 (49, 56.8)	.86
Minimal driving per week (days)	3 (1, 5)	3 (2, 3)	.61
Driving professionally (%)	12%	17%	.76
Avoidance of situations (%)	24%	33%	.64
Self-reported accidents (%)	12%	17%	.76
Self-reported fines (%)	12%	0%	.38

Table 12b. Showing the median scores and IQR for those that crashed in the driving simulator and those that did not. Adapted from Study III.

Ocusweep tests	Pass (n = 17) Median (IQR)	Fail (n = 6) Median (IQR)	p-value
OcuRT (ms)	536 (482, 572)	617 (565, 636)	.04595
RTP – Reaction Time RT	921 (855, 966)	959 (947, 976)	.34
RTP – Functional Ability Score	71 (44, 96)	53 (31, 71)	.40

Table 12c. Showing the median scores and IQR for those that crashed in the driving simulator and those that did not. From Study III.

Neuropsychological tests	Pass (n = 17) Median (IQR)	Fail (n = 6) Median (IQR)	p-value
MoCA	26.0 (23.0, 27.0)	25 (25.0, 25.0)	.80
MoCA-Blind	18.0 (18.0, 20.0)	18.5 (18.0, 19.8)	.74
TMT-A	52 (20, 63)	16 (9, 26)	.03
TMT-B	59 (47, 63)	39 (10, 53)	.16
TMT B-A	4 (-12, 27)	17 (2, 26)	.44
BVRT - correct	6 (4, 7)	5 (4, 6)	.69
BVRT - error	7 (5, 8)	9 (7, 10)	.32
SMT– seconds	32 (29, 36)	41 (33, 46)	.22
SMT – error	0.0 (0.0, 2.0)	0.5 (0.0, 1.0)	.11
DST - forward	41 (21, 59)	35 (25, 38)	.73
DST - backward	40 (18, 51)	35 (13, 55)	.75
SRT-V RT (percentile)	43 (22, 64)	24 (10, 61)	.12
SRT-V MT (percentile)	22 (8, 42)	30 (12, 53)	.58
SRT-A RT (percentile)	86 (65, 96)	73 (57, 88)	.70
SRT-A MT (percentile)	35 (14, 49)	53 (42, 57)	.19
CRT RT (percentile)	38 (9, 65)	13 (7, 19)	.26
CRT MT (percentile)	15 (12, 31)	23 (7, 28)	.23
UFOV 1 (ms)	13.9 (13.8, 20.4)	13.8 (13.8, 18.5)	.58
UFOV 2 (ms)	90.5 (20.4, 147.2)	32.2 (20.5, 48.8)	.50
UFOV 3 (ms)	170.5 (73.8, 260.5)	148.9 (82.2, 280.1)	.97

4.9 Introducing and evaluating two new tablet-based tests of real-world visual function for visual search and face recognition (Study IV)

4.9.1 Test refinement

Each test was intentionally run for longer than piloting indicated was necessary. To determine how many trials were actually required to obtain stable estimates of performance, we examined how test-retest variability varied as a function of test duration/number of trials (Figure 25). For Faces, this involved simply analysing the first N trials. For Search, data were analysed from progressively more sparse subsets of spatially-distributed locations, as shown in Figure 25a.

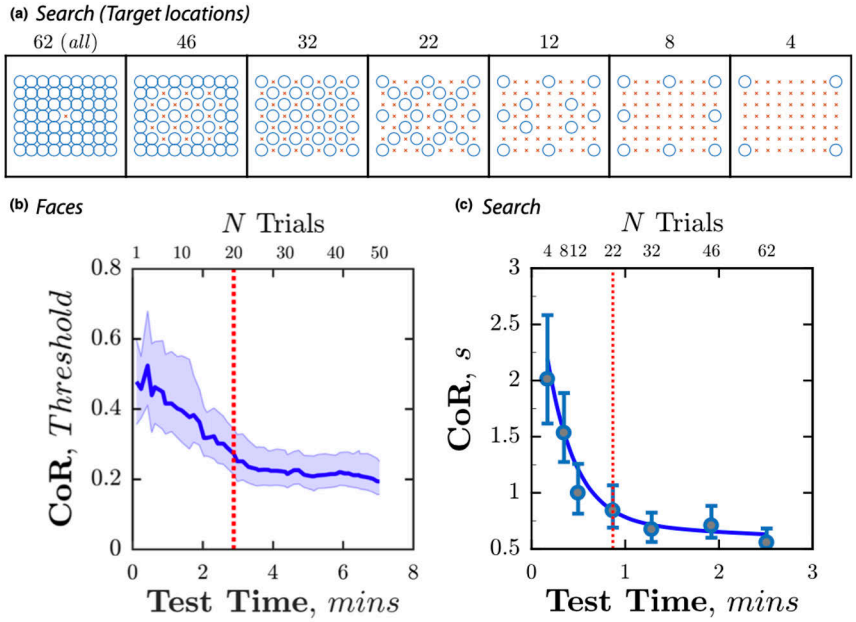


Figure 25. Test refinement. Within-subject (test-retest) measurement variability, as a function of N trials for the (**b**) Faces test and (**c**) Search Test. Coefficients of Repeatability were derived using Bland-Altman analysis, as detailed in Figure 28. Panel (**a**) shows the target locations (blue circles) associated with each Search grid.

Precision improved rapidly for the first 20 (Faces) or 22 trials (Search), and more gradually thereafter. We anticipate that ~20 trials will be sufficient for most clinical purposes. This corresponds to approximately 1–3 minutes of testing. We therefore report data only for these subsets of ~20 trials in the remainder of the manuscript.

4.9.2 Normative values

There was no systematic difference in performance between the first and second run, either for Faces (p = .11) or Search (p = .42). Accordingly, data from both runs were concatenated to produce the normative distributions shown in Figure 26. Appropriate probability distributions (black lines) were fitted to the raw data. These were used to determine the 99% upper-bound point (dashed vertical line): the cut-off point below which 99% of young, visually-normal participants would be expected to score. These values were 3.50 (Faces) and 3.1s (Search). Values greater than this may indicate abnormal test performance.

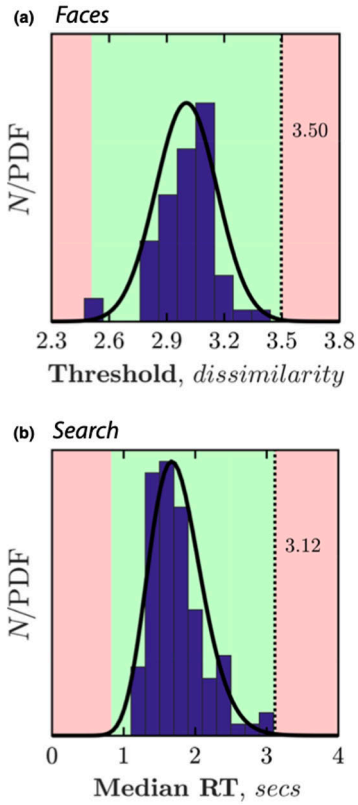


Figure 26. Normative data for the (**a**) Faces test, and (**b**) Search test. Curves show maximum likelihood fits of appropriate probability density functions (PDF) (Faces: Gaussian probability density function. Search: Gamma probability density function). Dashed vertical lines indicate the cut-off point, below which 99% of normally-sighted participants would be expected to score.

For Search it is also possible to consider performance for individual spatial locations. This could be important if, for example, attempting to detect localised visual field loss. Accordingly, Figure 27 shows normative median values, and the 99% upper cut-off value for each location). As highlighted in Figure 27b, there was a clear effect of eccentricity, with participants being slower to locate more peripheral targets.

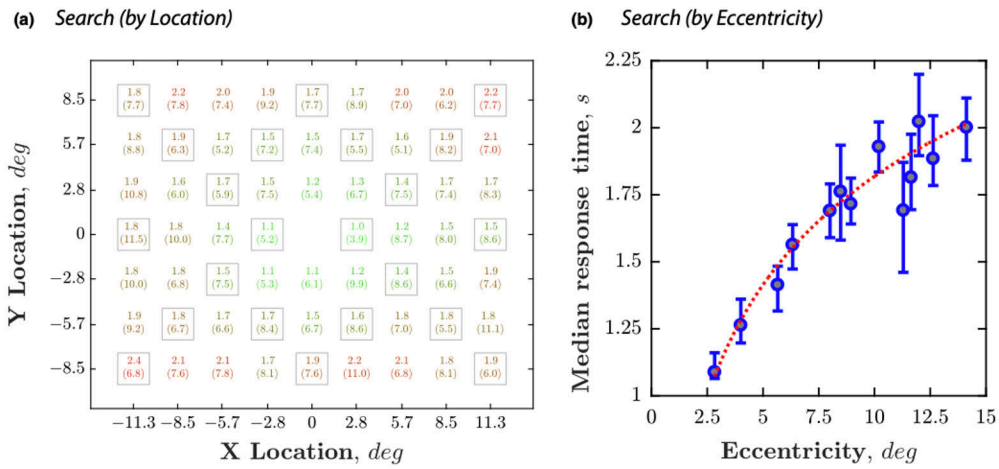


Figure 27. Pointwise normative data for the Search task, as a function of (**a**) Screen location; (**b**) Eccentricity from the centre. Grey boxes indicate the subset of 22 points that were used in all other figures and analyses. Data for the other 40 locations also given for completeness. Eccentricities computed assuming a viewing distance of 50 cm. Red dashed line indicates the best fitting power function [$y = -13.14x^{-0.05} + 13.56$; Adjusted $R^2 = 0.90$]. Individual markers indicate median response times with 95% confidence intervals.

4.9.3 Test-retest reliability

As shown in Figure 28, the Coefficient of Repeatability (± CI 95%) was 0.27 (0.22, 0.35) for Faces, and 0.84 (0.71, 1.07) for Search. There was no systematic effects of learning or fatigue. Measurement error tended to be approximately normally distributed, although on the Search task there was a tendency for variability to increase as a function of overall reaction time.

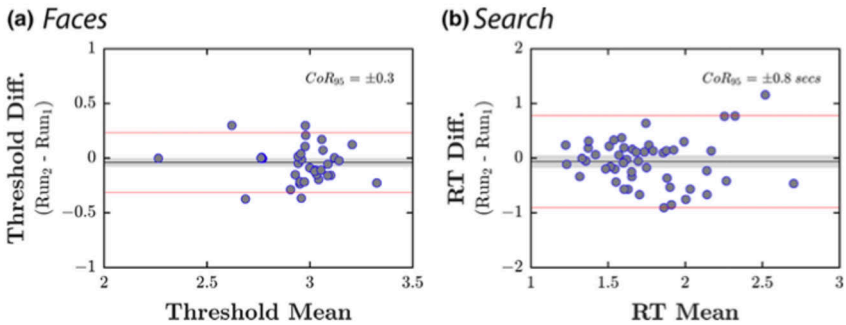


Figure 28. Bland-Altman analyses of retest repeatability for (**a**) Faces and (**b**) Search. Each marker represents a single participant. Grey shaded regions show 95% confidence intervals around the mean. Dashed red lines indicate the 95% limits of agreement. Text (top-right) gives the 95% Coefficient of Repeatability (CoR95).

4.9.4 Test duration

For Faces, median (CI95%) test duration was 191s (168, 228) for the first run and 155s (139, 186) for the second: a statistically significant difference (t-test: $p = .042$) of 19%. For Search, median (CI95%) test duration was 51s (46, 56) for the first run, and 47s (45, 50) for the second: a non-significant difference of 8% (t-test: $p = .16$).

4.9.5 Usability & completion rate

All participants (100%) completed both tests twice, with no participants exhibiting/reporting any difficulties. Participants were asked to rate how clearly they understood what to do on each test, on a scale from 0 (incomprehensible) to 10 (very understandable). Ratings of comprehension were high for both tests with a mean (CI95%) rating of 9.6 (9.1, 9.8) for Faces, and 9.7 (9.2, 9.8) for Search.

4.9.6 Relationships with cognition and basic vision

As shown in Figure 29, there was no significant association between performance on either test and with: (1) Digit Span general cognition; (2) logMAR letter acuity; or Pelli-Robson contrast sensitivity (see Figure 29 for p values) – although there was a trend towards an association between acuity and performance on the Search task ($p = .056$; r = 0.27). In those 30 participants who performed both tests, there was also no correlation between scores on the Faces and Search task ($r = 0.18$; $p = .16$).

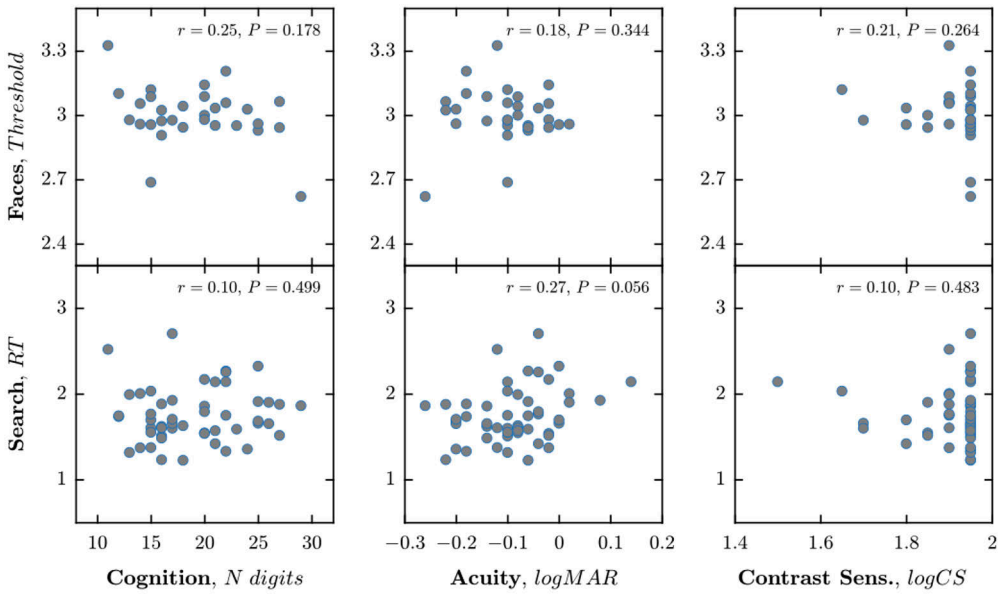


Figure 29. Scatter plots showing the relationships with cognition (Digit Span) and basic vision (Acuity, Contrast Sensitivity). Each marker indicates a single participant, with scores across the two runs of each test mean-averaged. Text (top-right) gives the results of independent Spearman's rank correlations.

4.10 Using a webcam to autonomously monitor compliance during visual field assessments (Study V)

4.10.1 Predicting Measurement Error

Results for healthy eyes are shown in Figures 30A to 30G, broken down by biomarker. Considered in isolation, three of the five biomarkers (head location variability, head rotation variability, and blink rate) were significantly associated with overall (test–retest) measurement error (p = .003–.007).

Put simply, the tests with lowest test–retest variability tended to be those during which individuals moved their eyes least (p = .226), moved their head least (p = .007, p = .003), blinked least (p = .007), and responded consistently quickly (p = .060). Furthermore, when all seven (including surprise and sadness) individual biomarkers were averaged together, a single composite variable was even more highly associated with measurement error than any single biomarker alone (Figure 30H; $p < .001$, r = 0.51).

We used analogous data for the 14 eyes from glaucoma patients to confirm the repeatability of these results and to ensure that they generalize to patients (Figure 31). As with normally sighted individuals (shown previously in Figure 30H), there

was a statistically significant positive association between the composite biomarker metric and test–retest variability (p = .011). The range of measurement errors observed was much greater in patients, however (Figure 31; note the difference in y-axis scale), consistent with previous reports of higher measurement variability in eyes with visual field loss (Gardiner & Mansberger 2016). Possibly owing to the small sample size, none of five individual biomarkers alone reached significance in patients (p > 0.05; data not shown). None of the biomarkers were correlated with false-positive rates (mean p = .206) or false-negative rates (mean p = .565) on the HFA.

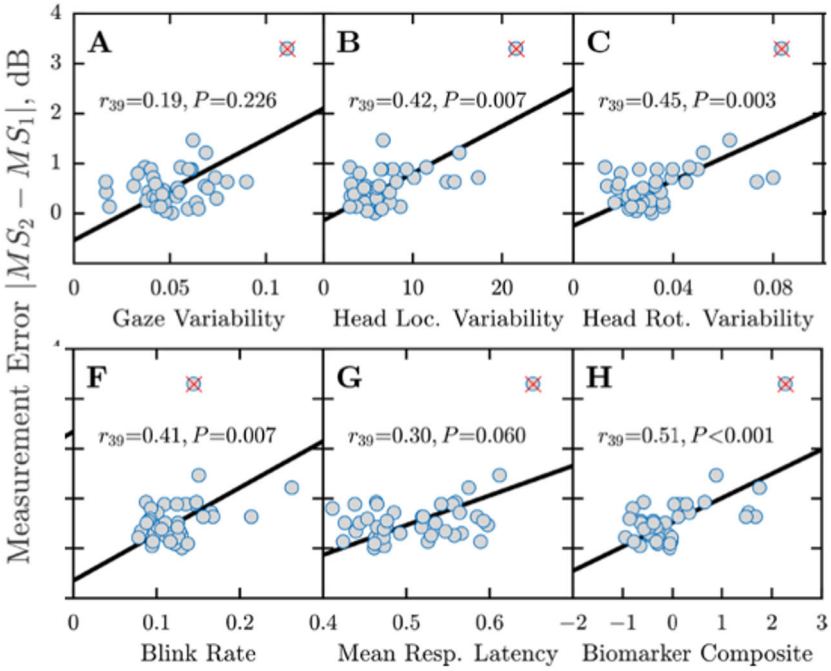


Figure 30. Overall test-retest data from healthy eyes. Each panel shows visual field measurement error (absolute test–retest difference in mean sensitivity) as a function of seven potential biomarkers of task compliance (**A**–**G**), as well as a function of a composite measure computed as the linear-weighted sum of all seven individual biomarkers (**H**). See Methods section for technical details on how each variable was computed. Panels **D** (mean sadness) and **E** (mean surprise) were omitted because they were not relevant in the scope of the thesis. Markers show the raw measurements for individual eyes. The marker with a *red cross* was excluded from all analyses as a possible statistical outlier. However, all *P* values were smaller if this point was included. Black lines show geometric mean regression slopes. Figures within each panel show correlation statistics.

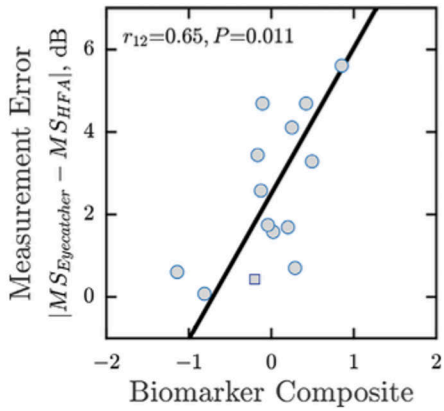


Figure 31. Overall test-retest data from 14 eyes from glaucoma patients; same format as Figure 30H. Note that in this instance, MS1 was measured using the HFA (not the screen perimeter). However, in practice the values from the two tests were robustly correlated (Pearson correlation, $r_{12} = 0.86$; $p < 0.001$), and any deviation between the two would likely only serve to minimize (add noise to) any of the effects reported in the present work. The *square marker* indicates the fellow eye from the one patient with unilateral secondary glaucoma for which the visual field was within normal limits.

4.10.2 Predicting Trial-by-Trial Lapses

The previous analyses suggest that the biomarkers are associated with the overall reliability of visual field assessment. Therefore, these biomarkers could be seen as new reliability metrics (Bengtsson & Heijl 2000, Birt et al. 1997). To examine whether it is possible to use these to detect lapses in concentration in real time, during the test, a trial-by-trial analysis was performed. The results are shown in Figure 32.

All the biomarkers were predictive of trial-by-trial lapses, with the hit rate, P(hit), for visible (>3 dB suprathreshold) stimuli decreasing progressively as a function of biomarker magnitude (logistic regression, $p < 0.001$).

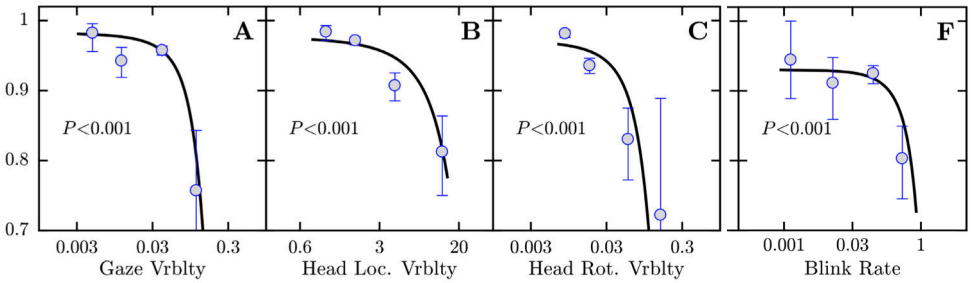


Figure 32. Trial-by-trial analyses examining the proportion of easily visible (−3 dB or brighter) stimuli that were correctly responded to as a function of biomarker magnitude. A failure to respond to these bright stimuli were considered a lapse in concentration. Markers represent mean hit rate [± 95% confidence intervals] for binned data, aggregated across all participants. Black lines and *P* values represent the result of logistic regressions fitted to the raw binary (hit/miss) data (not to the displayed markers). Note that these curves are plotted on a log *x*-axis, although tickmark values are shown in the original linear units, and all analyses were performed on the original, untransformed data. *P* values give the results of χ^2 tests, examining whether the logistic model fits the data significantly better than a constant model. Panels D (mean sadness) and E (mean surprise) were omitted because they were not relevant in the scope of the thesis.

5 Discussion

5.1 General discussion of studies

5.1.1 OcuDrive

Demographics were similar for glaucoma patients and healthy participants, including driving experience and driving exposure, except that glaucoma patients reported to self-regulate their driving behaviour by avoiding demanding driving situations more often. Situations that were most commonly avoided were driving in the dark, driving during rush hour and driving in bad weather. This was in line with literature, which showed that glaucoma patients avoid demanding situations more often than controls (Yamasaki et al. 2020). Both driving in the dark and driving in bad weather reduce contrast and require more cognitive effort from the driver. Since glaucoma patients already report a higher cognitive workload while driving (Engström, Johansson & Östlund 2005, Gangeddula et al. 2017, Prado Vega et al. 2013), this could explain why they choose to avoid situations that are adding cognitive load to the task of driving.

The first aim of this study was to evaluate whether it is needed to use vision fair testing in glaucoma patients, specifically in a population of current drivers, as cognitive decline and glaucoma can co-exist. In line with literature and our hypothesis that neuropsychological tests with a large visual field component are most affected by glaucomatous visual field defects, glaucoma patients scored lower percentile scores than the healthy group on the TMT-A and TMT-B. However, when removing the visual component by calculating the TMT B-A index, the difference between both groups was very small. The TMT B-A index seems more robust for visual field defects in glaucoma patients, but more studies are needed to evaluate these results in larger groups. Glaucoma severity, based on MD, was not related to performance on neuropsychological tests.

Research showed that some drivers with glaucoma can be safe traffic participants, but that glaucoma can be a risk factor for MVAs in real-world situations (Bhorade et al. 2016, Bowers et al. 2005, Kwon et al. 2016, McGwin, Owsley & Ball 1998) as well as in driving simulators. In our driving simulator, glaucoma patients performed similarly to the controls on all measures. Those who crashed differed from

those who did not regarding the presence of comorbidity, reaction time (OcuRT), and TMT-A score. This could indicate that other factors than the presence of glaucomatous visual field defects alone play are role, including neuropsychological functioning. However, the current work was limited in group size due to simulator sickness and therefore statistical analysis cannot confirm these differences. One observation that can be made in the Intersections ride is that drivers from both groups are relatively careful, as they often stop at intersections where they have priority and could have continued driving. Therefore, behavioural factors should also be taken into account.

Regarding the Ocusweep tests, the most interesting difference between the groups can be seen in the FAS of the Reaction Time Perimetry (RTP) test. These results were expected since this test measures functional vision across the visual field.

5.1.2 New vision tests

OcuRT was shown to be a robust measurement of reaction time, as it only showed a significant learning effect between OcuRT1 and OcuRT3 (Figure 16), consistency measures show good agreement between all three repetitions and comparison to the cloud dataset shows that our sample of reaction times is similar to the population sample. Therefore, it is likely that conclusions based on these reaction times are representative for future references to these data. Comparing OcuRT to another simple reaction time test, the VTS, shows that it is faster to complete and correlates significantly to other neuropsychological tests. The reaction times in the cognitive load condition are significantly longer than in any of the normal repetitions of OcuRT. Through these longer reaction times and by the calculated validity index, diminished attention due to cognitive load was detected. When the OcuRT measurement is invalid, this might be an indication of inattention.

The Visual Search test and the Face Recognition test were completed by normally sighted young adults and provide normative control data. The young participants indicated that these tests feel like games, which helps engagement if they would be completed autonomously by patients, without the need for costly technicians (Aslam et al. 2016). There was little or no association with more basic tests of vision, suggesting that these novel tests measure something else than visual acuity and contrast sensitivity alone and is particularly well suited to identify those who are experiencing everyday difficulties not captured by more basic measures of visual function (Drum, Calogero & Rorer 2007, Gonder et al. 2014). Additionally, these measures are easy to administer since there will be no need to calibrate the luminance or chromaticity or display or to precisely control the viewing distance or ambient lighting of the observer.

Data from both normally sighted young adults and older glaucoma patients collected during the laptop-based visual field test indicated that autonomous biomarkers of task compliance are associated with measurement error during visual field assessments. The association was greatest when multiple biomarkers were considered in combination and was true in terms of the overall reliability of a test as well as with individual trials. The biomarkers can be computed in real time using an ordinary webcam and laptop and represent "free" information: measurements that can be made in the background without requiring existing perimetric protocols to be altered in any way and without extending assessment durations, or the demands placed on patients. The biomarker estimates could also be used to trigger automated feedback, encouraging the patient to keep going and remain vigilant or to give each response a reliability score. This principle is also used in most modern perimeters already to estimate sensitivity (Turpin et al. 2003). This information can be used to determine the next stimulus to present (Watson 2017) and when a certain level of reliability has been attained (McKendrick & Turpin 2005). This would mean that more compliant participants would be required to complete fewer trials, whereas non-compliant participants may be asked to complete additional trials in order to reach a given level of data quality. More generally, the present work can also be viewed in the context attempting to replace or augment human technicians. This way, it is possible to facilitate home monitoring (Anderson et al. 2017) or free up manpower in clinics (McKendrick et al. 2019). Using biomarkers to assess reliability of measurements could also improve comfort and accessibility for patients, as with using head and eye tracking there is no need for chin rests and fixation targets (Jones et al. 2019). It is also interesting to note that some of the biomarkers considered here (e.g., blinking, unsteady fixation) have also been shown to provide direct indications of the presence and magnitude of various ophthalmic pathologies (Long et al. 2019).

5.2 Strengths and limitations

Because the studies in this thesis were mostly of an exploratory nature, analysis in this thesis could benefit from a larger, more heterogenous group of participants. Participants of different age ranges with different severities of visual impairment, cognitive impairment, and a combination of both should be studied. This would allow the OcuDrive project to do multivariable analysis and interaction analysis and the new tests could benefit from heterogenous normative data, as well as more information on how different aetiologies influence performance on these specific tests. Next to the already limited sample size, not all participants were able to complete all the tests, such as the Ocusweep Neural and Reaction Time Perimetry tests. And, as with most driving simulator studies specifically in older age groups, there was significant drop-out due to driving simulator sickness. Therefore, the

ability to draw conclusions based on the low number of participants that completed all tests is limited. Another possible limitation related to the driving simulator used in this study was the tendency to stop at intersections, even if the participant had right of way. It is unclear if that behaviour is specific to this driving simulator, or if they also do this in real world situations to give themselves more time to safely cross an intersection.

New tests were introduced with the goal to investigate the efficacy and validity of the tests. The primary limitation of our studies introducing new tests is that we only assessed young people with healthy vison with the goal of performing a preliminary assessment of feasibility and to provide limits on what constitutes normal performance. A final limitation is that testing in the present work took place in a controlled environment, often different from a busy clinic.

In the compliance study, it is likely that even more sensitive biomarkers could be obtained in the future through improved hardware or more sophisticated computer vision algorithms. It is noteworthy, for example, that the biomarkers in the present study were derived from an integrated laptop camera, sampled at only 5 Hz and 640 × 480-pixel resolution. The glaucoma patients in particular appeared highly motivated and were likely more compliant than the typical individual seen in a busy glaucoma clinic. It is well established that test variability increases with eccentricity (Heijl, Lindgren & Lindgren 1989, Phu, Kalloniatis & Khuu 2016, Phu, Kalloniatis & Khuu 2018, Wall, Kutzko & Chauhan 2002), whereas testing in the present study was limited to the central 15°. Future studies might use a larger screen to consider how well biomarkers correlate with a greater dynamic range of measurement variability at more peripheral test locations.

The most important strength of this study is that it assesses more than just simple vision tests. The OcuDrive project conducted an extensive neuropsychological examination, where the tests were selected because they are related to fitness to drive, next to several visual measurements. The combination of both is not often seen in literature on fitness to drive. The combination of the carefully selected neuropsychological test battery and visual measurements allows for comparison within the domains of neuropsychology and ophthalmology, but also between both domains.

The tests in this thesis go beyond static measures that are currently used in clinics, and which are used by the European Union to evaluate fitness to drive. Research has shown that those do not relate well to fitness to drive, and this is where real-world related tasks might help to explain the relationship between impairment and driving performance.

This thesis also showed that tests can be made less mentally straining for the patients by having portable devices that do not use a chin rest and using tasks that are described as fun and game-like.

6 Conclusions

6.1 Aim 1 – current regulations and research

The first overarching aim of this thesis was to establish whether the current regulations for fitness to drive in older glaucoma patients agree with what has been found in research. The literature review described that the regulations are based on relatively simple visual measures such as visual acuity and the extent of the visual field, but that those cannot completely describe or predict the risk an individual has while driving. This risk also depends on environmental factors, for example type of road or weather conditions, self-regulations, cognitive resources, and compensational abilities. Therefore, we can conclude that the current regulations don't agree with what has been found in research on fitness to drive in glaucoma.

6.2 Aim 2 – the effect of visual field defects on neuropsychological tests

The second aim of the thesis was to investigate the effect of visual field defects on neuropsychological tests that are commonly used in assessments of fitness to drive of those with cognitive impairment. In the second and third study, the OcuDrive project, the first aim was to evaluate whether it is warranted to use vision-fair testing in glaucoma patients. We hypothesized that visual field defects interfere with scores on neuropsychological tests that use visual items that are spread across the visual field and that more severe visual field defects are related to lower scores than mild visual field defects on these tests. The results in this study indicate that the TMT might be affected by glaucomatous visual field defects, but other tests in the current study provide little evidence for the need for vision fair testing. Glaucoma severity did not influence performance on the neuropsychological tests in this group of current drivers, which might be explained by the location of the visual field defects. Non-visual neuropsychological tests are scarce and can be less reliable. In the case of glaucoma, the benefits do not outweigh the downsides. However, one must be careful when interpretating the results of tests that require a larger intact visual field.

6.3 Aim 3 – driving simulator performance of glaucoma patients

The third aim was to evaluate driving performance in a group of older glaucoma patients compared to an age-similar healthy group. In this study, no significant differences between the two groups were identified in driving simulator performance. In the literature a higher crash risk related to reduced driving skills in glaucoma patients has been described, but this was not seen in the current study. Glaucoma patients self-regulate their driving behaviour by avoiding circumstances that increase cognitive load and lower visibility. Additionally, the role of either neuropsychological or vision-based tests related to driving performance were analysed. Those who crashed all had some form of comorbidity (p=.01), slower reaction time (OcuRT; p=0=.046), and lower percentile TMT-A score (p=.027), but the sample sizes after drop-out due to simulator sickness were small. The current study does not provide evidence that glaucoma patients that are current drivers as a group are more at risk for crashes in a driving simulator. However, more research is needed to identify which glaucoma patients are at higher risk for unsafe driving behaviour as there were only three participants in this thesis with glaucoma that completed all testing. Future research should focus on high demanding situations, such as driving in low luminance conditions, and should be done using larger groups, to allow for multivariable analysis involving not only visual performance but also cognition and comorbidity.

6.4 Aim 4 – new tests for visual functioning

The fourth aim of the thesis was to evaluate new vision tests that are related to real-world functioning in a healthy population, including a reaction time test, a test for visual search, a test for face recognition and how to monitor compliance during tablet-based visual field testing. In the first study, OcuRT, it can be concluded that OcuRT can be used to assess reaction time to a basic and easy to see stimulus and could be used to indicate validity of responses. Further studies are however necessary to evaluate these results in patient groups with visual and cognitive impairment.

The fourth study with new tablet-based tests concluded that within a few minutes, reliable measures could be taken on face recognition and visual search. The healthy young participants enjoyed the tests as they felt like games. The final fifth study concluded that using only an ordinary webcam, it is possible to derive real-time measures of task compliance during visual field assessment, and these can be used to identify unreliable assessments and/or unreliable responses within an assessment. In the long term, such autonomous measures could facilitate the creation of more intelligent and accessible forms of vision assessment: assessments in which "compliant" individuals can be processed even more rapidly than at present, but

wherein individuals who might otherwise struggle to complete an automated test will be given the additional time, care, and attention required to ensure robust, clinically useful data.

6.5 To conclude

Current regulations do not align with what is seen in real-world situations. To bridge the gap from research to practice, new tests are needed. This thesis identifies several new robust measures that are fast, easy, and more fun than traditional testing, but more research is needed in larger, heterogenous groups. Additionally, simply more tests will most likely not solve this issue. We should focus on the whole person, their circumstances, compensation strategies, and self-regulation. Awareness on what can influence driving performance could therefore increase safety on the roads.

7 Acknowledgements

This thesis would not exist without the personal support of countless people. From teachers supporting my curiosity about the world, even if I showed it by digging up worms and keeping them in my pocket at school, to those continuing to explain the same thing over and over when I had a harder time studying in high school and to those who believed in me when I wanted to explore new skills and subjects with little experience at first. Additionally, this thesis is only possible because it is built on previous scientific work and interesting ideas, for which I am grateful. During this PhD project there are a several persons in specific who I would like to thank for their contribution to this thesis.

I want to thank everyone at Ocusweep for welcoming me in their company and teach me about their culture. Some things – like hot lunches - I could not get on board with, but perjantai pulla was a hit. Specifically, I wish to express my gratitude to Markku. Thank you for giving me the chance to develop my own ideas and supporting my research in all ways possible. I enjoyed our discussions and your creative insights that benefitted my work. I want to thank Eija for her help and feedback on my writing.

I wish to thank Nomdo and Frans for their input and guidance and teaching me about the academic world and making me feel included and at home in a new workplace. I want to thank Dick for working together and our interesting talks about traffic psychology. I enjoyed our collaboration a lot. I want to thank everyone at the Crabb Lab for our fruitful collaboration.

I also want to thank my EGRET(+) colleagues. Being in a consortium with a group of peers, learning together and travelling to various places in Europe together has enriched the experience of doing a PhD. Allie, you have thought me so much about following what you believe in and excelling at what you want to do. Kosta, nobody laughs harder at your jokes than you do, but I think I am in second place. I am in awe of your intelligence and believe you'll do great things in your life. Hinke, you managed the consortium in an excellent way, but I value our discussions at the end of our PhDs - whether it was about spacing text, brain analysis or motherhood – the most.

Aafke, Anouk, Jildou, Mieke, Suzanne and Vera. I thank you all for your support, for our amazing trip to Stockholm, and your interest in my studies. It does not matter whether we all live in Leeuwarden, Groningen or at times all in a different country, I believe the Nitro's will be together forever.

Thank you to Joop and Marian, for supporting both me and Niels to go pursue this dream of mine. I will never forget our fun times during the 27-hour drive from Leeuwarden to Turku in the car, and the best nap ever after that.

I want to thank my parents Douwe-Henk and Mieke for their support throughout the years. Finland seemed far in the beginning, but we've made some memories that we would not otherwise have. Thank you for your support for my studies over all the years, from primary school to the PhD and beyond.

Niels, I want to thank you for following me when I was following my dreams. Thank you for your trust in me. Even though glaucoma and driving were not your areas of interest, you have listened to all my stories and invested your time and energy into helping me. You are the best partner anyone could wish for.

Hugo and Lasse – you are both the best things that have ever happened to me. I love you to infinity and beyond. This work is dedicated to you.

This thesis was financially supported by European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 675033 and the Turku University Foundation.

Leeuwarden, September 18th, 2025
Iris Tigchelaar

References

Adrian J, Le Brun J, Miller NR, Sahel JA, Saillant G & Bodaghi B (2019): Implications of monocular vision for racing drivers. PLoS One **14**: 1–16.

Anderson AJ, Bedggood PA, George Kong YX, Martin KR & Vingrys AJ (2017): Can Home Monitoring Allow Earlier Detection of Rapid Visual Field Progression in Glaucoma? Ophthalmology **124**: 1735–1742.

Andriuta D, Diouf M, Roussel M & Godefroy O (2019): Is Reaction Time Slowing an Early Sign of Alzheimer's Disease? A Meta-Analysis. Dement Geriatr Cogn Disord **47**: 281–288.

Anstey KJ, Eramudugolla R, Chopra S, Price J & Wood JM (2017): Assessment of Driving Safety in Older Adults with Mild Cognitive Impairment. J Alzheimer's Dis **57**: 1197–1205.

Anstey KJ, Wood J, Lord S & Walker JG (2005): Cognitive, sensory and physical factors enabling driving safety in older adults. Clin Psychol Rev. 45–65.

Aslam TM, Parry NRA, Murray IJ, Salleh M, Dal Col C, Mirza N, Czanner G & Tahir HJ (2016): Development and testing of an automated computer tablet-based method for self-testing of high and low contrast near visual acuity in ophthalmic patients. Graefe's Arch Clin Exp Ophthalmol **254**: 891–899.

Atan T & Akyol P (2014): Reaction Times of Different Branch Athletes and Correlation between Reaction Time Parameters. Procedia - Soc Behav Sci **116**: 2886–2889.

Aulhorn E & Karmeyer H (1977): Frequency distribution in early glaucomatous visual field defects. Doc Ophthalmol Proc Ser **14**: 75–83.

Australian Bureau of Statistics (2005): *Population Projections*.

Bailon O, Roussel M, Boucart M, Krystkowiak P & Godefroy O (2010): Psychomotor slowing in mild cognitive impairment, Alzheimer's disease and lewy body dementia: Mechanisms and diagnostic value. Dement Geriatr Cogn Disord **29**: 388–396.

Ball K & Owsley C (1993): The useful field of view test: a new technique for evaluating age-related declines in visual function. J Am Optom Assoc **64**: 71–9.

Ball K, Owsley C, Sloane ME, Roenker DL & Bruni JR (1993): Visual attention problems as a predictor of vehicle crashes in older drivers. Investig Ophthalmol Vis Sci **34**: 3110–3123.

Ball K, Owsley C, Stalvey B, Roenker DL, Sloane ME & Graves M (1998): Driving avoidance and functional impairment in older drivers. Accid Anal Prev **30**: 313–322.

Baltrusaitis T, Zadeh A, Lim YC & Morency LP (2018): OpenFace 2.0: Facial behavior analysis toolkit. Proc - 13th IEEE Int Conf Autom Face Gesture Recognition, FG 2018 59–66.

Bao S & Boyle LN (2008): Driver performance at two-way stop-controlled intersections on divided highways. Transp Res Rec **2069**: 26–32.

Bao S & Boyle LN (2009): Age-related differences in visual scanning at median-divided highway intersections in rural areas. Accid Anal Prev **41**: 146–152.

Bastawrous A, Rono HK, Livingstone IAT, Weiss HA, Jordan S, Kuper H & Burton MJ (2015): Development and validation of a smartphone-based visual acuity test (peek acuity) for clinical practice and Community-Based Fieldwork. JAMA Ophthalmol **133**: 930–937.

Bédard M, Parkkari M, Weaver B, Riendeau J & Dahlquist M (2010): Assessment of driving performance using a simulator protocol: Validity and reproducibility. Am J Occup Ther **64**: 336–340.

Bengtsson B & Heijl A (2000): False-negative responses in glaucoma perimetry: Indicators of patient performance or test reliability? Investig Ophthalmol Vis Sci **41**: 2201–2204.

Bennett JM, Chekaluk E & Batchelor J (2016, September 1): Cognitive Tests and Determining Fitness to Drive in Dementia: A Systematic Review. J Am Geriatr Soc. Wiley/Blackwell (10.1111) 1904–1917.

Bentley SA, LeBlanc RP, Nicolela MT & Chauhan BC (2012): Validity, reliability, and repeatability of the useful field of view test in persons with normal vision and patients with glaucoma. Investig Ophthalmol Vis Sci **53**: 6763–6769.

Benton AL (1945): A Visual Retention Test for Clinical Use. Arch NeurPsych **54**: 212–216.

Bhorade AM, Yom VH, Barco P, Wilson B, Gordon M & Carr D (2016): On-road Driving Performance of Patients with Bilateral Moderate and Advanced Glaucoma. Am J Ophthalmol **166**: 43–51.

Birt CM, Shin DH, Samudrala V, Hughes BA, Kim C & Lee D (1997): Analysis of reliability indices from Humphrey visual field tests in an urban glaucoma population. Ophthalmology **104**: 1126–1130.

Black JM, Jacobs RJ, Phillips G, Chen L, Tan E, Tran A & Thompson B (2013): An assessment of the iPad as a testing platform for distance visual acuity in adults. BMJ Open **3**: 5–7.

Blana E & Golias J (2002): Differences between vehicle lateral displacement on the road and in a fixed-base simulator. Hum Factors **44**: 303–313.

Blane A (2016): Through the looking glass: A review of the literature investigating the impact of glaucoma on crash risk, driving performance, and driver self-regulation in older drivers. J Glaucoma **25**: 113–121.

Bodduluri L, Boon MY, Ryan M & Dain SJ (2018): Normative values for a tablet computer-based application to assess chromatic contrast sensitivity. Behav Res Methods **50**: 673–683.

Bosten J, Álvaro L, Alvarez J, Meyer B, Tang T, Maule J & Franklin A (2019): Tablet-based app for screening for CVD in young children. J Vis **19**: 71.

Boufous S, Finch C, Hayen A & Williamson A (2008): The impact of environmental, vehicle and driver characteristics on injury severity in older drivers hospitalized as a result of a traffic crash. J Safety Res **39**: 65–72.

Bowers A, Peli E, Elgin J, McGwin G & Owsley C (2005): On-road driving with moderate visual field loss. Optom Vis Sci **82**: 657–667.

Box E, Gandolfi J & Mitchell K (2010): Maintaining Safe Mobility for the Ageing Population. The Role of the Private Car. RAC Foundation.

Braaten, E. B. (2018). Wechsler Adult Intelligence Scale. In J. Matson (Ed.), *SAGE encyclopedia of intellectual and developmental disorders*. SAGE Publications.

Brady TF, Konkle T, Alvarez GA & Oliva A (20008): Visual long-term memory has a massive storage capacity for object details. PNAS **105**:.

Bredewoud A (2008): Cognitieve stoornissen en het rijbewijs 1–6.

Brinker TJ, Brieske CM, Esser S, Klode J, Mons U, Batra A, Rüther T, Seeger W, Enk AH, von Kalle C, Berking C, Heppt MV, Gatzka MV, Bernardes-Souza B, Schlenk RF & Schadendorf D. (2018): A face-aging app for smoking cessation in a waiting room setting: Pilot study in an HIV outpatient clinic. J Med Internet Res **20**:.

Brooks JO, Goodenough RR, Crisler MC, Klein ND, Alley RL, Koon BL, Logan WC Jr, Ogle JH, Tyrrell RA & Wills RF (2010): Simulator sickness during driving simulation studies. Accid Anal Prev **42**: 788–796.

Brusini P & Johnson CA (2007): Staging Functional Damage in Glaucoma: Review of Different Classification Methods. Surv Ophthalmol **52**: 156–179.

Carr DB, Barco PP, Wallendorf MJ, Snellgrove CA & Ott BR (2011): Predicting Road Test Performance in Drivers with Dementia. J Am Geriatr Soc **59**: 2112–2117.

CDC (2020): WISQARS (Web-based Injury Statistics Query and Reporting System)|Injury Center|CDC. Centers Dis Control Prev Natl Cent Inj Prev Control.

Charman WN (1997): Vision and driving - A literature review and commentary. Ophthalmic Physiol Opt **17**: 371–391.

Chihuri S, Mielenz TJ, Dimaggio CJ, Betz ME, Diguiseppi C, Jones VC & Li G (2016): Driving cessation and health outcomes in older adults. J Am Geriatr Soc **64**: 332–341.

Clarke DD, Ward P, Bartle C & Truman W (2010): Older drivers' road traffic crashes in the UK. Accid Anal Prev **42**: 1018–1024.

Classen S, Wang Y, Crizzle AM, Winter SM & Lanford DN (2013): Predicting older driver on-road performance by means of the useful field of view and trail making test part b. Am J Occup Ther **67**: 574–582.

Clay OJ, Wadley VG, Edwards JD, Roth DL, Roenker DL & Ball KK (2005): Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: Current and future implications. Optom Vis Sci **82**: 724–731.

Coeckelbergh TRM, Brouwer WH, Cornelissen FW, Van Wolffelaar P & Kooijman AC (2002): The Effect of Visual Field Defects on Driving Performance. Arch Ophthalmol **120**: 1509.

Coleman AL & Miglior S (2008): Risk Factors for Glaucoma Onset and Progression. Surv Ophthalmol **53**: 2–10.

Correa PC, Medeiros FA, Abe RY, Diniz-Filho A & Gracitelli CPB (2019): Assessing driving risk in patients with glaucoma. Arq Bras Oftalmol. Conselho Brasileiro de Oftalmologia 245–252.

Crabb DP, Smith ND, Glen FC, Burton R & Garway-Heath DF (2013): How does glaucoma look?: Patient perception of visual field loss. Ophthalmology **120**: 1120–1126.

Crabb DP, Smith ND, Rauscher FG, Chisholm CM, Barbur JL, Edgar DF & Garway-Heath DF (2010): Exploring eye movements in patients with glaucoma when viewing a driving scene. (He S, Ed.)PLoS One **5**: e9710.

Crabb DP & Viswanathan AC (2005): Integrated visual fields: A new approach to measuring the binocular field of view and visual disability. Graefe's Arch Clin Exp Ophthalmol **243**: 210–216.

Cronin-Golomb A, Corkin S & Growdon JH (1995): Visual dysfunction predicts cognitive deficits in alzheimer's disease. Optom Vis Sci **72**: 168–176.

Cross J, McGwin G, Rubin G, Ball K, West SK, Roenker DL & Owsley C (2009): Visual and Medical Risk Factors for Motor Vehicle Collision Involvement among Older Drivers. Br J Ophthalmol **93**: 400–404.

Dabasia PL, Edgar DF, Garway-Heath DF & Lawrenson JG (2014): A survey of current and anticipated use of standard and specialist equipment by UK optometrists. Ophthalmic Physiol Opt **34**: 592–613.

Dalziel JR & Job RFS (1997): Motor vehicle accidents, fatigue and optimism bias in taxi drivers. Accid Anal Prev **29**: 489–494.

Danesh-Meyer H V. & Levin LA (2015): Glaucoma as a Neurodegenerative Disease. J Neuro-Ophthalmology **35**: S22–S28.

Dawson JD, Anderson SW, Uc EY, Dastrup E & Rizzo M (2009): Predictors of driving safety in early Alzheimer disease. Neurology **72**: 521–527.

De Bondt WFM & Thaler RH (1995): Financial Decision-Making in Markets and Firms: A Behavioral Perspective. Handbooks Oper Res Manag Sci. 385–410.

de Fez D, Luque MJ, Matea L, Piñero DP & Camps VJ (2018): New iPAD-based test for the detection of color vision deficiencies. Graefe's Arch Clin Exp Ophthalmol **256**: 2349–2360.

de Haan, GA, Tucha, O, & Heutink, J (2020). Effects of low visual acuity on neuropsychological test scores: A simulation study. *The Clinical neuropsychologist*, ***34*(**1), 140–157. https://doi.org/10.1080/13854046.2019.1596315

Decina LE & Staplin L (1993): Retrospective evaluation of alternative vision screening criteria for older and younger drivers. Accid Anal Prev **25**: 267–275.

Dhamija S & Boult TE (2017): Exploring Contextual Engagement for Trauma Recovery. IEEE Comput Soc Conf Comput Vis Pattern Recognit Work **2017-July**: 2267–2277.

Dobbs BM & Shergill SS (2013): How effective is the trail making test (parts a and b) in identifying cognitively impaired drivers? Age Ageing **42**: 577–581.

Dorr M, Lesmes LA, Elze T, Wang H, Lu ZL & Bex PJ (2017): Evaluation of the precision of contrast sensitivity function assessment on a tablet device. Sci Rep **7**: 1–11.

Dorr M, Lesmes LA, Lu ZL & Bex PJ (2013): Rapid and reliable assessment of the contrast sensitivity function on an iPad. Investig Ophthalmol Vis Sci **54**: 7266–7273.

Drum B, Calogero D & Rorer E (2007): Assessment of visual performance in the evaluation of new medical products. Drug Discov Today Technol **4**: 55–61.

Dupuis K, Pichora-Fuller MK, Chasteen AL, Marchuk V, Singh G & Smith SL (2015): Effects of hearing and vision impairments on the Montreal Cognitive Assessment. Aging, Neuropsychol Cogn **22**: 413–437.

Eck RW & Winn G (2002): Older-driver perception of problems at unsignalized intersections on divided highways. Transp Res Rec **1818**: 70–77.

Economou, A, Pavlou, D, Beratis, I, Andronas, N, Papadimitriou, E, Papageorgiou, SG, & Yannis, G. (2020). Predictors of accidents in people with mild cognitive impairment, mild dementia due to Alzheimer's disease and healthy controls in simulated driving. *International journal of geriatric psychiatry*, ***35***(8), 859–869. https://doi.org/10.1002/gps.5306

Ekman P & Rosenberg EL (2012): What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). What Face Reveal Basic Appl Stud Spontaneous Expr Using Facial Action Coding Syst. Oxford University Press.

Emerson J, Johnson A, Dawson J, Ergun Y, Anderson SW & Rizzo M (2012): Predictors of driving outcomes in older age. Psychol Aging **27**: 550–559.

Emerson JL, Johnson AM, Dawson JD, Uc EY, Anderson SW & Rizzo M (2012): Predictors of driving outcomes in advancing age. Psychol Aging **27**: 550–559.

Engström J, Johansson E & Östlund J (2005): Effects of visual and cognitive load in real and simulated motorway driving. Transp Res Part F Traffic Psychol Behav **8**: 97–120.

Enoch J, McDonald L, Jones L, Jones PR & Crabb DP (2019). Evaluating Whether Sight Is the Most Valued Sense. *JAMA ophthalmology*, ***137***(11), 1317–1320. https://doi.org/10.1001/jamaophthalmol.2019.3537

Eramudugolla R, Price J, Chopra S, Li X & Anstey KJ (2016): Comparison of a Virtual Older Driver Assessment with an On-Road Driving Test. J Am Geriatr Soc **64**: e253–e258.

European Council of Optometry and Optics (2017): *Visual standards for driving in Europe*.

Fonda SJ, Wallace RB & Herzog AR (2001): Changes in driving patterns and worsening depressive symptoms among older adults. Journals Gerontol - Ser B Psychol Sci Soc Sci **56**: S343–S351.

Fox GK, Bowden SC, Bashford GM & Smith DS (1997): Alzheimer's disease and driving: Prediction and assessment of driving performance. J Am Geriatr Soc **45**: 949–953.

Friedman C, McGwin G, Ball KK & Owsley C (2013): Association between higher order visual processing abilities and a history of motor vehicle collision involvement by drivers ages 70 and over. Investig Ophthalmol Vis Sci **54**: 778–782.

Fukuoka H, Nagaya M & Toba K (2015): The occurrence of visual and cognitive impairment, and eye diseases in the super-elderly in Japan: A cross-sectional single-center study. BMC Res Notes **8**:.

Gangeddula V, Ranchet M, Akinwuntan AE, Bollinger K & Devos H (2017): Effect of cognitive demand on functional visual field performance in senior drivers with glaucoma. Front Aging Neurosci **9**: 286.

Garbutt S, Matlin A, Hellmuth J, Schenk AK, Johnson JK, Rosen H, Dean D, Kramer J, Neuhaus J, Miller BL, Lisberger SG & Boxer AL (2008): Oculomotor function in frontotemporal lobar degeneration, related disorders and Alzheimer's disease. Brain **131**: 1268–1281.

Gardiner SK & Mansberger SL (2016): Effect of restricting perimetry testing algorithms to reliable sensitivities on test-retest variability. Investig Ophthalmol Vis Sci **57**: 5631–5636.

Gierczuk D & Ljach W (2012): Evaluating the coordination of motor abilities in Greco-Roman wrestlers by computer testing. Hum Mov **13**: 323–329.
Glen FC, Baker H & Crabb DP (2014): A qualitative investigation into patients' views on visual field testing for glaucoma monitoring. BMJ Open **4**:.
Glen FC, Baker H & Crabb DP (2014): A qualitative investigation into patients' views on visual field testing for glaucoma monitoring. BMJ Open **4**:.
Glen FC, Crabb DP, Smith ND, Burton R & Garway-Heath DF (2012): Do patients with glaucoma have difficulty recognizing faces? Investig Ophthalmol Vis Sci **53**: 3629–3637.
Glen FC, Crabb DP, Smith ND, Burton R & Garway-Heath DF (2012): Do patients with glaucoma have difficulty recognizing faces? Investig Ophthalmol Vis Sci **53**: 3629–3637.
Glen FC, Smith ND & Crabb DP (2013): Saccadic eye movements and face recognition performance in patients with central glaucomatous visual field defects. Vision Res **82**: 42–51.
Glen FC, Smith ND & Crabb DP (2015): Impact of superior and inferior visual field loss on hazard detection in a computer-based driving test. Br J Ophthalmol **99**: 613–617.
Gonder JR, Walker VM, Barbeau M, Zaour N, Zachau BH, Hartje JR & Li R (2014): Costs and quality of life in diabetic macular edema: Canadian burden of diabetic macular edema observational study (C-REALITY). J Ophthalmol **2014**:.
Grace J, Amick MM, D'abreu A, Festa EK, Heidel WC & Ott BR (2005): Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. J Int Neuropsychol Soc **11**: 766–775.
Gracitelli CPBB, Tatham AJ, Boer ER, Abe RY, Diniz-Filho A, Rosen PN, FAM & Medeiros FA (2015): Predicting Risk of Motor Vehicle Collisions in Patients with Glaucoma: A Longitudinal Study. PLoS One **10**: e0138288.
Hagger-Johnson G, Deary IJ, Davies CA, Weiss A & Batty GD (2014): Reaction time and mortality from the major causes of death: The NHANES-III study. PLoS One **9**: 3–9.
Harrabi H, Kergoat M-J, Rousseau J, Boisjoly H, Schmaltz H, Moghadaszadeh S, Roy-Gagnon M-H & Freeman EE (2015): Age-Related Eye Disease and Cognitive Function. Invest Ophthalmol Vis Sci **56**: 1217–1221.
Hassan R, Twyman N, Fui-Hoon Nah F & Siau KL (2016): Patient Engagement in the Medical Facility Waiting Room Using Gamified Healthcare Information Delivery. Int Conf HCI Business, Gov Organ.
Hawkins AS, Szlyk JP, Ardickas Z, Alexander KR & Wilensky JT (2003): Comparison of contrast sensitivity, visual acuity, and Humphrey visual field testing in patients with glaucoma. J Glaucoma **12**: 134–138.
Haymes SA, LeBlanc RP, Nicolela MT, Chiasson LA & Chauhan BC (2007): Risk of falls and motor vehicle collisions in glaucoma. Investig Ophthalmol Vis Sci **48**: 1149–1155.
Haymes SA, LeBlanc RP, Nicolela MT, Chiasson LA & Chauhan BC (2008): Glaucoma and on-road driving performance. Investig Ophthalmol Vis Sci **49**: 3035–3041.
Heijl A, Lindgren A & Lindgren G (1989): Test-retest variability in glaucomatous visual fields. Am J Ophthalmol **108**: 130–135.
Heijl A, Patella VM, Chong LX, Iwase A, Leung CK, Tuulonen A, Lee GC, Callan T & Bengtsson B (2019): A New SITA Perimetric Threshold Testing Algorithm: Construction and a Multicenter Clinical Study. Am J Ophthalmol **198**: 154–165.
Henson DB & Emuh T (2010): Monitoring vigilance during perimetry by using pupillography. Investig Ophthalmol Vis Sci **51**: 3540–3543.
Ho WL, Leung Y, Tsang AWT, So KF, Chiu K & Chang RCC (2012): Review: Tauopathy in the retina and optic nerve: Does it shadow pathological changes in the brain? Mol Vis **18**: 2700–2710.
Horswill MS, Marrington SA, McCullough CM, Wood J, Pachana NA, McWilliam J & Raikos MK (2008): The hazard perception ability of older drivers. Journals Gerontol - Ser B Psychol Sci Soc Sci **63**: 212–218.

Huang P, Shi Y, Wang X, Liu M & Zhang C (2014): Interocular asymmetry of the visual field defects in newly diagnosed normal-tension glaucoma, primary open-angle glaucoma, and chronic angle-closure glaucoma. J Glaucoma **23**: 455–460.

Huisingh C, McGwin G & Owsley C (2016): Association of visual sensory function and higher-order visual processing skills with incident driving cessation. Clin Exp Optom **99**: 441–448.

Hunt L, Morris JC, Edwards D & Wilson BS (1993): Driving Performance in Persons with Mild Senile Dementia of the Alzheimer Type. J Am Geriatr Soc **41**: 747–753.

Johnson CA & Keltner JL (1983): Incidence of Visual Field Loss in 20,000 Eyes and Its Relationship to Driving Performance. Arch Ophthalmol **101**: 371–375.

Johnson CA, Thapa S, George Kong YX & Robin AL (2017): Performance of an iPad Application to Detect Moderate and Advanced Visual Field Loss in Nepal. Am J Ophthalmol **182**: 147–154.

Jonas JB, Wei W Bin, Zhu LP, Xu L & Wang YX (2018): Cognitive Function and Ophthalmological Diseases: The Beijing Eye Study. Sci Rep **8**:.

Jones PR, Smith ND, Bi W & Crabb DP (2019): Portable perimetry using eye-tracking on a tablet computer—A feasibility assessment. Transl Vis Sci Technol **8**:.

Kasneci E, Sippel K, Aehling K, Heister M, Rosenstiel W, Schiefer U & Papageorgiou E (2014): Driving with binocular visual field loss? A study on a supervised on-road parcours with simultaneous eye and head tracking. PLoS One **9**:.

Kaur A, Mustafa A, Mehta L & Dhall A (2019): Prediction and Localization of Student Engagement in the Wild. 2018 Int Conf Digit Image Comput Tech Appl DICTA 2018. Institute of Electrical and Electronics Engineers Inc. 637–644.

Kawano N, Iwamoto K, Ebe K, Suzuki Y, Hasegawa J, Ukai K, Umegaki H, Iidaka T & Ozaki N (2011): Effects of mild cognitive impairment on driving performance in older drivers. Neurosci Res 71(Supplement): e389.

Keeffe JE, Jin CF, Weih LM, McCarty CA & Taylor HR (2002): Vision impairment and older drivers: Who's driving? Br J Ophthalmol **86**: 1118–1121.

Kelly EA, Stadler ME, Nelson S, Runge CL & Friedland DR (2018): Tablet-based Screening for Hearing Loss: Feasibility of Testing in Nonspecialty Locations. Otol Neurotol **39**: 410–416.

Kempen JH, Mark K & Feldman ST (1994): Effect of Visual Impairment on Neuropsychological Test Performance. J Clin Exp Neuropsychol **16**: 223–231.

Kleiner M, Brainard D, Pelli D, Ingling A, Murray R & Broussard C (2007): What's new in psychtoolbox-3. Perception **36**: 1–16.

Kong YXG, He M, Crowston JG & Vingrys AJ (2016): A comparison of perimetric results from a tablet perimeter and humphrey field analyzer in glaucoma patients. Transl Vis Sci Technol **5**:.

Kübler TC, Kasneci E, Rosenstiel W, Heister M, Aehling K, Nagel K, Schiefer U & Papageorgiou E (2015): Driving with Glaucoma: Task Performance and Gaze Movements. Optom Vis Sci **92**: 1037–1046.

Kwon M, Huisingh C, Rhodes LA, McGwin G, Wood JM & Owsley C (2016): Association between Glaucoma and At-fault Motor Vehicle Collision Involvement among Older Drivers: A Population-based Study. Ophthalmology **123**: 109–116.

Lahav K, Levkovitch-Verbin H, Belkin M, Glovinsky Y & Polat U (2011): Reduced Mesopic and Photopic Foveal Contrast Sensitivity in Glaucoma. Arch Ophthalmol **129**: 16.

Laksana E, Baltrusaitis T, Morency LP & Pestian JP (2017): Investigating Facial Behavior Indicators of Suicidal Ideation. Proc - 12th IEEE Int Conf Autom Face Gesture Recognition, FG 2017 - 1st Int Work Adapt Shot Learn Gesture Underst Prod ASL4GUP 2017, Biometrics Wild, Bwild 2017, Heteroge 770–777.

Langford J & Koppel S (2006): Epidemiology of older driver crashes - Identifying older driver risk factors and exposure patterns. Transp Res Part F Traffic Psychol Behav **9**: 309–321.

Lee HC, Cameron D & Lee AH (2003): Assessing the driving performance of older adult drivers: On-road versus simulated driving. Accid Anal Prev **35**: 797–803.

Lee M, Zulauf M & Caprioli J (1994): The influence of patient reliability on visual field outcome. Am J Ophthalmol **117**: 756–761.

Lee S, Black AA & Wood JM (2017): Effect of glaucoma on eye movement patterns and laboratory-based hazard detection ability. PLoS One **12**:.

Lee SS-Y, Wood JM & Black AA (2020): Impact of glaucoma on executive function and visual search. Ophthalmic Physiol Opt.

Lee SSY, Black AA & Wood JM (2018): Scanning Behavior and Daytime Driving Performance of Older Adults with Glaucoma. J Glaucoma **27**: 558–565.

Leinonen M & Mäntysalo T (2018): How do the Finns really see – Vision judged by visual fields and contrast sensitivity. Poster UK Eire Glaucoma Soc Meet 2018.

Lemay S, Bédard MA, Rouleau I & Tremblay PLG (2004): Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. Clin Neuropsychol **18**: 284–302.

Levecq L, De Potter P & Jamart J (2013): Visual acuity and factors influencing automobile driving status in 1,000 patients age 60 and older. Graefe's Arch Clin Exp Ophthalmol **251**: 881–887.

Lincoln NB, Radford KA, Lee E & Reay AC (2006): The assessment of fitness to drive in people with dementia. Int J Geriatr Psychiatry **21**: 1044–1051.

Perttu L (2018): *Octopus- ja Ocusweep-näkökenttien vertailu glaukoomapotilailla [Comparison of Octopus and Ocusweep visual fields in glaucoma patients]. Master's thesis, Turku University, Faculty of Medicine.*

Long E, Liu Z, Xiang Y, Xu A, Huang J, Huang X, Li X, Lin Z, Li J, Chen J, Zhang Y, Zhu Y, Chen C, Zhou Z, Ding X, Wu X, Li W, Chen H, Li R, Yang Y, Zheng W, Chen W, Lin H, Liu Y (2019): Discrimination of the behavioural dynamics of visually impaired infants via deep learning. Nat Biomed Eng **3**: 860–869.

Mancino R, Martucci A, Cesareo M, Giannini C, Corasaniti MT, Bagetta G & Nucci C (2017): Glaucoma and Alzheimer Disease: One Age-Related Neurodegenerative Disease of the Brain. Curr Neuropharmacol **16**: 971–977.

Marottoli RA, Mendes De Leon CF, Glass TA, Williams CS, Cooney LM & Berkman LF (2000): *Consequences of Driving Cessation: Decreased Out-of-Home Activity Levels*. J Gerontol (Vol. 55).

Marottoli RA, Mendes de Leon CF, Glass TA, Williams CS, Cooney LM, Berkman LF & Tinetti ME (1997): Driving cessation and increased depressive symptoms: prospective evidence from the New Haven EPESE. Established Populations for Epidemiologic Studies of the Elderly. J Am Geriatr Soc **45**: 202–206.

Marottoli RA, Richardson ED, Stowe MH, Miller EG, Brass LM, Cooney LM & Tinetti ME (1998): Development of a Test Battery to Identify Older Drivers at Risk for Self-Reported Adverse Driving Events. J Am Geriatr Soc **46**: 562–568.

Martin L (2017). The effects of task-relevant saccadic eye movements performed during the encoding of a serial sequence on visuospatial memory performance. Experimental Brain Research.Mathias JL & Lucas LK (2009): Cognitive predictors of unsafe driving in older drivers: A meta-analysis. Int Psychogeriatrics **21**: 637–653.

Mathias JL & Lucas LK (2009): Cognitive predictors of unsafe driving in older drivers: A meta-analysis. Int Psychogeriatrics. 637–653.

Mathôt S, Schreij D & Theeuwes J (2012): OpenSesame: An open-source, graphical experiment builder for the social sciences. Behav Res Methods **44**: 314–324.

Mayhew DR, Simpson HM, Wood KM, Lonero L, Clinton KM & Johnson AG (2011): On-road and simulated driving: Concurrent and discriminant validation. J Safety Res **42**: 267–275.

Mazumdar D, Pel J, Panday M, Asokan R, Vijaya L, Shantha B, George R & Van Der Steen J (2014): Comparison of saccadic reaction time between normal and glaucoma using an eye movement perimeter. Indian J Ophthalmol **62**: 55.

Mazzoli LS, Urata CN & Kasahara N (2019): Face memory deficits in subjects with eye diseases: a comparative analysis between glaucoma and age-related macular degeneration patients from a developing country. Graefe's Arch Clin Exp Ophthalmol **257**: 1941–1946.

Mccloskey LW, Koepsell TD, Wolf ME & Buchner DM (1994): Motor vehicle collision injuries and sensory impairments of older drivers. Age Ageing **23**: 267–273.

McCoskey M, Addis V, Goodyear K, Sankar PS, Ying G-S, Yu Y, Salowe R, Cui QN, Miller-Ellis E, Maguire M, O'Brien J (2018): Association between Primary Open-Angle Glaucoma and Cognitive Impairment as Measured by the Montreal Cognitive Assessment. Neurodegener Dis **18**: 315–321.

McDonald JH (2014): Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland.

McDowell JE, Dyckman KA, Austin BP & Clementz BA (2008): Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans. Brain Cogn **68**: 255–270.

McGwin G, Huisingh C, Jain SG, Girkin CA & Owsley C (2015): Binocular visual field impairment in glaucoma and at-fault motor vehicle collisions. J Glaucoma **24**: 138–143.

McGwin G, Mays A, Joiner W, DeCarlo DK, McNeal S & Owsley C (2004): Is glaucoma associated with motor vehicle collision involvement and driving avoidance? Investig Ophthalmol Vis Sci **45**: 3934–3939.

McGwin G, Owsley C & Ball K (1998): Identifying crash involvement among older drivers: Agreement between self-report and state records. Accid Anal Prev **30**: 781–791.

McGwin G, Owsley C & Ball K (1998): Identifying crash involvement among older drivers: Agreement between self-report and state records. Accid Anal Prev **30**: 781–791.

McGwin G, Xie A, Mays A, Joiner W, DeCarlo DK, Hall TA & Owsley C (2005): Visual field defects and the risk of motor vehicle collisions among patients with glaucoma. Investig Ophthalmol Vis Sci **46**: 4437–4441.

McKendrick AM & Turpin A (2005): Advantages of terminating Zippy Estimation by Sequential Testing (ZEST) with dynamic criteria for white-on-white perimetry. Optom Vis Sci **82**: 981–987.

McKendrick AM, Zeman A, Liu P, Aktepe D, Aden I, Bhagat D, Do K, Nguyen HD & Turpin A (2019): Robot assistants for perimetry: A study of patient experience and performance. Transl Vis Sci Technol **8**:59.

McKnight AJ & McKnight AS (1999): Multivariate analysis of age-related driver ability and performance deficits. Accid Anal Prev **31**: 445–454.

Medeiros AF, Weinreb RN, Boer E & Rosen PN (2012): Driving Simulation as a Performance-based Test of Visual Impairment in Glaucoma. J Glaucoma **21**: 221–227.

Merickel J, High R, Dawson J & Rizzo M (2019): Real-world risk exposure in older drivers with cognitive and visual dysfunction. Traffic Inj Prev **20**: S110–S115.

Mills RP, Budenz DL, Lee PP, Noecker RJ, Walt JG, Siegartel LR, Evans SJ & Doyle JJ (2006): Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. Am J Ophthalmol **141**: 24–30.

Montolio FGJ, Wesselink C, Gordijn M & Jansonius NM (2012): Factors that influence standard automated perimetry test results in glaucoma: Test reliability, technician experience, time of day, and season. Investig Ophthalmol Vis Sci **53**: 7010–7017.

Mulcahey MJ, Haley SM, Duffy T, Ni P & Betz RR (2008): Measuring Physicial Functioning In Children with Spinal Impairments with Computerized Adaptive Testing. J Pediatr Orthop **28**: 330–335.

Musselwhite C & Haddad H (2018): Older people's travel and mobility needs: a reflection of a hierarchical model 10 years on. Qual Ageing Older Adults **19**: 87–105.

Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whiteheard V, Collin I, Cummings JL & Chertkow H (2005): The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. J Am Geriatr Soc **53**: 695–699.

Nesaratnam N, Thomas PBM, Kirollos R, Vingrys AJ, Kong GYX & Martin KR (2017): Tablets at the bedside - IPad-based visual field test used in the diagnosis of Intrasellar Haemangiopericytoma: A case report. BMC Ophthalmol **17**: 1–5.

Ocusweep Oy. (2015): VAL1006 Clinical evaluation of Ocusweep, v. 2.0. Data on file.

Ogata NG, Daga FB, Jammal AA, Boer ER, Hill LL, Stringham JM, Susanna R & Medeiros FA (2019): Mobile Telephone Use and Reaction Time in Drivers With Glaucoma. JAMA Netw Open **2**: e192169.

Ono T, Yuki K, Asaoka R, Kouyama K, Abe T, Tanabe S, Fukagawa K, Uchino M, Shimoyama M, Ozawa Y, Ozeki N, Shiba D, Tsubota K (2015): Glaucomatous visual field defect severity and the prevalence of motor vehicle collisions in Japanese: A hospital/clinic-based cross-sectional study. J Ophthalmol **2015**:2015:497067

Ott BR, Heindel WC, Whelihan WM, Caron MD, Piatt AL & DiCarlo MA (2003): Maze Test Performance and Reported Driving Ability in Early Dementia. J Geriatr Psychiatry Neurol **16**: 151–155.

Owsley C, McGwin Jr G & Ball K (1998): Vision impairment, eye disease, and injurious motor vehicle crashes in the elderly. Ophthalmic Epidemiol **5**: 101–113.

Owsley C, Stalvey B, Wells J & Sloane ME (1999): Older drivers and cataract: Driving habits and crash risk. Journals Gerontol - Ser A Biol Sci Med Sci **54**: 203–211.

Owsley C, Stalvey BT, Wells J, Sloane ME & McGwin G (2001): Visual risk factors for crash involvement in older drivers with cataract. Arch Ophthalmol **119**: 881–887.

Owsley C, Wood JM & McGwin G (2015): A roadmap for interpreting the literature on vision and driving. Surv Ophthalmol **60**: 250–262.

Pallier G, Wilkinson R, Danthiir V, Kleitman S, Knezevic G, Stankov L & Roberts RD (2002): The role of individual differences in the accuracy of confidence judgments. J Gen Psychol **129**: 257–299.

Park S-W, Choi ES, Lim MH, Kim EJ, Hwang SI, Choi K-I, Yoo H-C, Lee KJ, Jung H-E (2011): Association Between Unsafe Driving Performance and Cognitive-Perceptual Dysfunction in Older Drivers. PM R **3**: 198–203.

Patel V, Hale TM, Palakodeti S, Kvedar JC & Jethwani K (2015): Prescription Tablets in the Digital Age: A Cross-Sectional Study Exploring Patient and Physician Attitudes Toward the Use of Tablets for Clinic-Based Personalized Health Care Information Exchange. JMIR Res Protoc **4**: e116.

Payne M, Janzen S, Earl E, Deathe B & Viana R (2017): Feasibility testing of smart tablet questionnaires compared to paper questionnaires in an amputee rehabilitation clinic. Prosthet Orthot Int **41**: 420–425.

Phu J, Kalloniatis M & Khuu SK (2016): The effect of attentional cueing and spatial uncertainty in visual field testing. PLoS One **11**: 1–18.

Phu J, Kalloniatis M & Khuu SK (2018): Reducing spatial uncertainty through attentional cueing improves contrast sensitivity in regions of the visual field with glaucomatous defects. Transl Vis Sci Technol **7**:.

Piersma D, Fuermaier ABM, de Waard D, De Deyn PP, Davidse RJ, de Groot J, Doumen MJA, Bredewoud RA, Claesen R, Lemstra AW, Vermeeren A, Ponds R, Verhey F, Brouwer WH, Tucha O (2018): The MMSE should not be the sole indicator of fitness to drive in mild Alzheimer's dementia. Acta Neurol Belg **0**: 3.

Piersma D, Fuermaier ABM, de Waard D, Davidse RJ, de Groot J, Doumen MJA, Bredewoud RA, Claesen R, Lemstra AW, Vermeeren A, Ponds R, Verhey F, Brouwer WH, Tucha O (2016): Prediction of Fitness to Drive in Patients with Alzheimer's Dementia. PLoS One.

Pollatsek A, Romoser MRE & Fisher DL (2012): Identifying and Remediating Failures of Selective Attention in Older Drivers. Curr Dir Psychol Sci **21**: 3–7.

Prado Vega R, van Leeuwen PM, Rendón Vélez E, Lemij HG & de Winter JCF (2013): Obstacle Avoidance, Visual Detection Performance, and Eye-Scanning Behavior of Glaucoma Patients in a Driving Simulator: A Preliminary Study. PLoS One **8**:.

Prea SM, Kong YXG, Mehta A, He M, Crowston JG, Gupta V, Martin KR & Vingrys AJ (2018): Six-month Longitudinal Comparison of a Portable Tablet Perimeter With the Humphrey Field Analyzer. Am J Ophthalmol **190**: 9–16.

Prieler J (2008): Manual RT: Reaction test. Version 31. Schuhfried GmbH, Mödling.

Pyun J, Kang M, Kim S, Baek M, Wang M & Kim S (2018): Driving Cessation and Cognitive Dysfunction in Patients with Mild Cognitive Impairment. J Clin Med **7**: 545.

Quigley H & Broman AT (2006): The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 262–267.

Racette L & Casson EJ (2005): The impact of visual field loss on driving performance: evidence from on-road driving assessments. Optom Vis Sci **82**: 668–74.

Ramulu PY, West SK, Munoz B, Jampel HD & Friedman DS (2009): Driving Cessation and Driving Limitation in Glaucoma. The Salisbury Eye Evaluation Project. Ophthalmology **116**: 1846–1853.

Rao HL, Yadav RK, Begum VU, Addepalli UK, Choudhari NS, Senthil S & Garudadri CS (2015): Role of visual field reliability indices in ruling out glaucoma. JAMA Ophthalmol **133**: 40–44.

Rashbass C (1961): The relationship between saccadic and smooth tracking eye movements. J Physiol **159**: 326–338.

Reigal RE, Barrero S, Martín I, Morales-Sánchez V, Juárez-Ruiz de Mier R & Hernández-Mendo A (2019): Relationships Between Reaction Time, Selective Attention, Physical Activity, and Physical Fitness in Children. Front Psychol **10**: 2278.

Reitan RM (1958): Validity of the Trail Making Test as an Indicator of Organic Brain Damage. Percept Mot Skills **8**: 271–276.

Richman J, Lorenzana LL, Lankaranian D, Dugar J, Mayer J, Wizov SS & Spaeth GL (2010): Importance of visual acuity and contrast sensitivity in patients with glaucoma. Arch Ophthalmol **128**: 1576–1582.

Rodríguez-Vallejo M, Ferrando V, Montagud D, Monsoriu JA & Furlan WD (2017): Stereopsis assessment at multiple distances with an iPad application. Displays **50**: 35–40.

Rodríguez-Vallejo M, Remón L, Monsoriu JA & Furlan WD (2015): Designing a new test for contrast sensitivity function measurement with iPad. J Optom **8**: 101–108.

Romoser M & Fisher DL (2009): The Effect of Active Versus Passive Training Strategies on Improving Older Drivers' Scanning in Intersections. Hum Factors **51**: 652–668.

Rosen P, Akinwuntan A, Wachtel J, Boer E, Weinreb R & Medeiros F (2018): Driver Fitness in Patients with Cognitive Impairment and Glaucoma. Proc Sixth Int Driv Symp Hum Factors Driv Assessment, Train Veh Des Driv.

Rottach KG, Zivotofsky AZ, Das VE, Averbuch-Heller L, Discenna AO, Poonyathalang A, Leigh RJ (1996): Comparison of horizontal, vertical and diagonal smooth pursuit eye movements in normal human subjects. Vision Res **36**: 2189–2195.

Roux-Sibilon A, Rutgé F, Aptel F, Attye A, Guyader N, Boucart M, Chiquet C & Peyrin C (2018): Scene and human face recognition in the central vision of patients with glaucoma. PLoS One **13**:.

Roy M & Liersch M (2013): I Am a Better Driver Than You Think **43**: 1–20.

Rubin GS, Ng ESW, Bandeen-Roche K, Keyl PM, Freeman EE & West SK (2007): A prospective, population-based study of the role of visual impairment in motor vehicle crashes among older drivers: The SEE study. Investig Ophthalmol Vis Sci **48**: 1483–1491.

Rubin GS, Roche KB, Prasada-Rao P & Fried LP (1994): Visual impairment and disability in older adults. Optom Vis Sci **71**: 750–760.

Rudman DL, Friedland J, Chipman M & Sciortino P (2006): Holding On and Letting Go: The Perspectives of Pre-seniors and Seniors on Driving Self-Regulation in Later Life. Can J Aging / La Rev Can du Vieil **25**: 65.

Sakong J, Kang P-S, Kim C-Y, Hwang T-Y, Jeon M-J, Park S-Y, Lee S-J, Won K-C, Lee S-B & Chung J-H (2007): Evaluation of reliability of traditional and computerized neurobehavioral tests. Neurotoxicology **28**: 235–239.

Sant'Ana J, Franchini E, da Silva V & Diefenthaeler F (2017): Effect of fatigue on reaction time, response time, performance time, and kick impact in taekwondo roundhouse kick. Sport Biomech **16**: 201–209.

Schuhfried G (1992): Vienna test system. Test Manag Program Modling, Austria 360.

Schulz AM, Graham EC, You YY, Klistorner A & Graham SL (2018): Performance of iPad-based threshold perimetry in glaucoma and controls. Clin Exp Ophthalmol **46**: 346–355.

Selander H, Lee HC, Johansson K & Falkmer T (2011): Older drivers: On-road and off-road test results. Accid Anal Prev **43**: 1348–1354.

Silva MT, Laks J & Engelhardt E (2009): Neuropsychological tests and driving in dementia: a review of the recent literature. Rev Assoc Med Bras **55**: 484–488.

Singletary BA, Do AN, Donnelly JP, Huisingh C, Mefford MT, Modi R, Mondesir FL, Ye Y, Owsley C & McGwin G (2017): Self-reported vs state-recorded motor vehicle collisions among older community dwelling individuals. Accid Anal Prev **101**: 22–27.

Sinnott C, Foley T, Forsyth J, Mcloughlin K, Horgan L & Bradley CP (2018): Consultations on driving in people with cognitive impairment in primary care: A scoping review of the evidence. PLoS One.

Siren A & Meng A (2011): Cognitive screening of older drivers does not produce safety benefits. Accid Anal Prev **45**: 634–638.

Sivan AB (1992): Benton Visual Retention Test (5th ed.). San Antonio, TX: The Psychological Corporation.

Smith ND, Crabb DP & Garway-Heath DF (2011): An exploratory study of visual search performance in glaucoma. Ophthalmic Physiol Opt **31**: 225–232.

Smith ND, Glen FC & Crabb DP (2012): Eye movements during visual search in patients with glaucoma. BMC Ophthalmol **12**: 1.

Snellgrove CA (2005): Cognitive screening for the safe driving competence of older people with mild cognitive impairment or early dementia 8-25–2009.

Song S, Shen L & Valstar M (2018): Human behaviour-based automatic depression analysis using hand-crafted statistics and deep learned spectral features. Proc - 13th IEEE Int Conf Autom Face Gesture Recognition, FG 2018. Institute of Electrical and Electronics Engineers Inc. 158–165.

Sonntag D (2018): Interactive Cognitive Assessment Tools: A Case Study on Digital Pens for the Clinical Assessment of Dementia. arXiv. https://arxiv.org/abs/1810.04943

Staplin L, Gish KW, Lococo KH, Joyce JJ & Sifrit KJ (2013): The Maze Test: A significant predictor of older driver crash risk. Accid Anal Prev **50**: 483–489.

Steinkuller PG (2010): Legal Vision Requirements for Drivers in the United States. AMA J Ethics **12**: 938–940.

Sterpu G, Saam C & Harte N (2018): Attention-based audio-visual fusion for robust automatic speech recognition. arXiv 111–115.

Stribling JC & Richardson JE (2016): Placing wireless tablets in clinical settings for patient education. J Med Libr Assoc **104**: 159–164.

Strutton R, Du Chemin A, Stratton IM & Forster AS (2016): System-level and patient-level explanations for non-attendance at diabetic retinopathy screening in Sutton and Merton (London, UK): A qualitative analysis of a service evaluation. BMJ Open **6**: 1–6.

Sweller J (1988): Cognitive Load During Problem Solving: Effects on Learning. Cogn Sci **12**: 257–285.

Szlyk JP, Mahler CL, Seiple W, Edward DP & Wilensky JT (2005): Driving performance of glaucoma patients correlates with peripheral visual field loss. J Glaucoma **14**: 145–150.

Takahashi A, Yuki K, Awano-Tanabe S, Ono T, Shiba D & Tsubota K (2018): Association between glaucoma severity and driving cessation in subjects with primary open-angle glaucoma. BMC Ophthalmol **18**:.

Tatham AJ, Boer ER, Gracitelli CPB, Rosen PN & Medeiros FA (2015): Relationship Between Motor Vehicle Collisions and Results of Perimetry, Useful Field of View, and Driving Simulation in Drivers With Glaucoma. Transl Vis Sci Technol **4**: 5.

Tham YC, Li X, Wong TY, Quigley HA, Aung T & Cheng CY (2014): Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. Ophthalmology **121**: 2081–2090.

Tiddeman BP, Stirrat MR & Perrett DI (2005): Towards realism in facial image transformation: Results of a wavelet MRF method. Comput Graph Forum **24**: 449–456.

Tilastokeskus (2020): Killed and injured in road traffic accidents by Area, Road user, Gender, Age, Year and Information. PxWeb.

Toner CK, Reese BE, Neargarder S, Riedel TM, Gilmore GC & Cronin-Golomb A (2012): Vision-fair neuropsychological assessment in normal aging, Parkinson's disease and alzheimer's disease. Psychol Aging **27**: 785–790.

Törnros J (1998): Driving behaviour in a real and a simulated road tunnel - A validation study. Accid Anal Prev **30**: 497–503.

Turpin A, McKendrick AM, Johnson CA & Vingrys AJ (2003): Properties of Perimetric Threshold Estimates from Full Threshold, ZEST, and SITA-like Strategies, as Determined by Computer Simulation. Investig Ophthalmol Vis Sci **44**: 4787–4795.

Tyler CW, Chan H, Liu L, McBride B, &
Kontsevich LL. (1992). Bit stealing: How to get 1786 or more gray levels from an 8-bit color monitor. Proceedings of SPIE - The International Society for Optical Engineering, **1666:** 351–364. https://doi.org/10.1117/12.135981

UNDESA (2019): World Population Prospects 2019: Highlights. 2–3.

Ungewiss J, Kübler T, Sippel K, Aehling K, Heister M, Rosenstiel W, Kasneci E & Papageorgiou E (2018): Agreement of driving simulator and on-road driving performance in patients with binocular visual field loss. Graefe's Arch Clin Exp Ophthalmol.

Van Landingham SW, Hochberg C, Massof RW, Chan E, Friedman DS & Ramulu PY (2013): Driving patterns in older adults with glaucoma. BMC Ophthalmol **13**: 14–16.

Vaucher P, Herzig D, Cardoso I, Herzog MH, Mangin P & Favrat B (2014): The trail making test as a screening instrument for driving performance in older drivers; A translational research. BMC Geriatr **14**:.

Vesti E, Perttu L & Niinistö M (2017): Glaucomatous visual field defects using Octopus and Ocusweep preimters. Poster P-WT-359 7th World Glaucoma Congr 2017, Helsinki.

Vingrys AJ, Healey JK, Liew S, Saharinen V, Tran M, Wu W & Kong GYX (2016): Validation of a tablet as a tangent perimeter. Transl Vis Sci Technol **5**:.

Vora JP, Varghese R, Weisenbach SL & Bhatt T (2016): Test-retest reliability and validity of a custom-designed computerized neuropsychological cognitive test battery in young healthy adults HHS Public Access. J Psychol Cogn **1**: 11–19.

Wall M, Kutzko KE & Chauhan BC (2002): The relationship of visual threshold and reaction time to visual field eccentricity with conventional automated perimetry. Vision Res **42**: 781–787.

Wang Y, Toor SS, Gautam R & Henson DB (2011): Blink frequency and duration during perimetry and their relationship to test-retest threshold variability. Investig Ophthalmol Vis Sci **52**: 4546–4550.

Watson AB (2017): QUEST: A general multidimensional bayesian adaptive psychometric method. J Vis **17**: 1–27.

Watson AB & Pelli DG (1983): Quest: A Bayesian adaptive psychometric method. Percept Psychophys **33**: 113–120.

Weinreb RN & Khaw PT (2004): Primary open-angle glaucoma. Lancet.

Whelihan WM, DiCarlo MA & Paul RH (2005): The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. Arch Clin Neuropsychol **20**: 217–228.

Wiecek E, Pasquale LR, Fiser J, Dakin S & Bex PJ (2012): Effects of peripheral visual field loss on eye movements during visual search. Front Psychol **3**: 1–13.
Windsor TD, Anstey KJ, Butterworth P, Luszcz MA & Andrews GR (2007): The role of perceived control in explaining depressive symptoms associated with driving cessation in a longitudinal study. Gerontologist **47**: 215–223.
Wittich W, Phillips N, Nasreddine ZS & Chertkow H (2010): Sensitivity and specificity of the Montreal cognitive assessment modified for individuals who are visually impaired. J Vis Impair Blind **104**: 360–368.
Wong AY, Dodge RM & Remington LA (1995): Comparing threshold visual fields between the Dicon TKS 4000 automated perimeter and the Humphrey Field Analyzer. J Am Optom Assoc **66**: 706–711.
Wood J, Chaparro A, Anstey K, Lacherez P, Chidgey A, Eisemann J, Gaynor A & La P (2010): Simulated Visual Impairment Leads to Cognitive Slowing in Older Adults. Optom Vis Sci **87**: 1037–1043.
Wood JM (1999): How do visual status and age impact on driving performance as measured on a closed circuit driving track? Ophthalmic Physiol Opt **19**: 34–40.
Wood JM, Anstey KJ, Kerr GK, Lacherez PF & Lord S (2008): A multidomain approach for predicting older driver safety under in-traffic road conditions. J Am Geriatr Soc **56**: 986–993.
Wood JM, Black AA, Mallon K, Thomas R & Owsley C (2016): Glaucoma and Driving: On-Road Driving Characteristics. PLoS One **11**: e0158318.
Wood JM, Horswill MS, Lacherez PF & Anstey KJ (2013): Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. Accid Anal Prev **50**: 1161–1168.
Wood JM, McGwin G, Elgin J, Vaphiades MS, Braswell RA, de Carlo DK, Kline LB & Owsley C (2011): Hemianopic and quadrantanopic field loss, eye and head movements, and driving. Investig Ophthalmol Vis Sci **52**: 1220–1225.
Wood JM & Owsley C (2014): Useful field of view test. Gerontology. 315–318.
Wood JM & Troutbeck R (1994): Effect of visual impairment on driving. Hum Factors **36**: 476–487.
Wood JM & Troutbeck R (1995): Elderly drivers and simulated visual impairment. Optom Vis Sci **72**: 115–124.
World Health Organization (2019): *World report on vision*. World Heal Organ (Vol. 214).
Wu Z, Guymer RH, Jung CJ, Goh JK, Ayton LN, Luu CD, Lawson DJ, Turpin A & McKendrick AM (2015): Measurement of retinal sensitivity on tablet devices in age-related macular degeneration. Transl Vis Sci Technol **4**: 1–8.
Wu Z, Yao T, Fu Y & Jiang Y-G (2017): Deep learning for video classification and captioning. Front Multimed Res 3–29.
Yamani Y, Horrey WJ, Liang Y & Fisher DL (2016): Age-related differences in vehicle control and eye movement patterns at intersections: Older and middle-aged drivers. PLoS One **11**:.
Yamasaki T, Yuki K, Awano-Tanabe S, Ono T, Murata H, Tsubota K & Asaoka R (2020): Binocular superior visual field areas associated with driving self-regulation in patients with primary open-angle glaucoma. Br J Ophthalmol **0**: 1–6.
Yan X, Abdel-Aty M, Radwan E, Wang X & Chilakapati P (2008): Validating a driving simulator using surrogate safety measures. Accid Anal Prev **40**: 274–288.
Yochim BP, Mueller AE, Kane KD & Kahook MY (2012): Prevalence of cognitive impairment, depression, and anxiety symptoms among older adults with glaucoma. J Glaucoma **21**: 250–254.
Nasreddine Z (2004): Montreal Cognitive Assessment (MoCA) Administration and Scoring Instructions The 1–4.
Zhang ZT, Zhang SC, Huang XG & Liang LY (2013): A pilot trial of the iPad tablet computer as a portable device for visual acuity testing. J Telemed Telecare **19**: 55–59.

List of Tables and Figures

Tables

Figures

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2025

ISBN 978-952-02-0332-0 (PRINT)
ISBN 978-952-02-0333-7 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)