Cognitive Predictors of Driving Ability in Older Adults

Nicole Amy Matas BPsych (Hons)

School of Psychology

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

THE UNIVERSITY of ADELAIDE

September, 2016

Contents

List of Tables

List of Figures

List of Acronyms

ABS Australian Bureau of Statistics

ACTIVE Advanced Cognitive Training for Independent and Vital Elderly

AIC Akaike Information Criterion

ATCQ Attitudes Towards Computers Questionnaire

BITRE Bureau of Infrastructure, Transport and Regional Economics

CALTEST California Test

CI Confidence Interval

df degrees of freedom

DMV Department of Motor Vehicles

GP General Practitioner

GRIMPS Gross Impairment Screening Battery of General Physical and Mental Abilities

HR Hazard Ratio

IADL Instrumental Activities of Daily Living

IT Inspection Time

logMAR logarithm of the Minimum Angle of Resolution

ms milliseconds

MAB Medical Advisory Board

MMSE Mini Mental State Examination

MOMSSE Mattis Organic Mental Syndrome Screening Examination

MSSQ Motion Sickness Susceptibility Questionnaire

MVPT Motor-Free Visual Perception Test

NHTSA National Highway Traffic Safety Administration

OECD Organisation for Economic Co-operation and Development

OR Odds Ratio

PASE Physical Activity Scale for the Elderly

ProPerVis Proficiency of Peripheral Visual Processing

RT Reaction Time

SDLP Standard Deviation of Lane Position

SS Simulator Sickness

SSQ Simulator Sickness Questionnaire

TMT-A Trail Making Test Part A

TMT-B Trail Making Test Part B

UFOV™ Useful Field of View Test™

UN United Nations

VIF Variance Inflation Factor

VMI Visualisation of Missing Information

Abstract

The main aim was to investigate functional predictors of driving ability in older adults. The principal focus was on cognitive predictors, but visual and physical function measures were also included. The cognitive assessments reflected domains identified as most relevant to driving outcomes, including visual attention, processing speed, and general cognitive functioning. The specific cognitive tests included the Useful Field of View Test™ (UFOV™), which is notable for its consistent relationship with a broad range of driving outcomes; and Inspection Time (IT) and ProPerVis, assessments of processing speed and crowding across the visual field, respectively, which have not previously been investigated in relation to driving outcomes but have potential as screening tests. A secondary aim of the thesis was to investigate methodological issues concerning use of driving simulators.

Five studies formed a sequential program of research. Study 1 examined factors contributing to performance on the UFOV™; although the UFOV™ has been extensively used in past research, its psychometric properties are not yet well understood. The results from Study 1 showed that UFOV™ Subtest 1 primarily reflected low-level visual function; UFOV™ Subtest 2 reflected change detection, processing speed (as assessed by IT), and general cognitive function; and UFOV™ Subtest 3 reflected crowding (as assessed by ProPerVis), processing speed (as assessed by IT), contrast sensitivity, and general cognitive function. These results suggested that IT and Crowding may be useful in predicting driving performance, based on their importance for UFOV™ Subtests 2 and 3, which have been consistently linked to important driving outcomes. Studies 2, 3 and 4 investigated methodological issues related to driving simulators, including simulator sickness, validity, reliability, and usability. The results from Study 2, which investigated risk factors for simulator sickness, showed that older adults in general are a high-risk group, as are females and those with a history of motion sickness. Studies 3 and 4 used a variety of methods to show that the simulator demonstrated reliability, face validity, content validity, and convergent validity, and was perceived by participants as providing an acceptable method of assessing driving skills.

Study 5 investigated functional predictors (cognitive, visual, and physical) of simulated driving performance on two tasks: a Brake Reaction Time (RT) task and a Traffic Participation Task. The results from Study 5 showed small but significant correlations between cognitive test performance (IT, Crowding, and UFOV™ Subtest 2) and Brake RT. For the Traffic Task, only Crowding was significantly correlated with driving performance. Physical activity and visual function were not associated with driving performance. These results have implications for current assessment procedures. They suggest that visual function measures are not generally useful for determining fitness-to-drive, a conclusion that has important implications for practices at present widespread in many jurisdictions responsible for driver licensing. Regarding the cognitive measures, it is suggested that the IT and Crowding measures may be useful as screening measures for older drivers, especially those who are most at-risk. Further research with a broader range of participants would be needed to establish appropriate test cut-points. Limitations and further implications of the results are discussed.

Signed Statement

I certify that this work, except where specified in the text, contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

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

Acknowledgements

I would like to thank my supervisors, Emeritus Professor Ted Nettelbeck and Professor Nick Burns. Your advice and support has been invaluable. Ted suggested the original idea for the project and provided me with encouragement and expert guidance throughout my candidature. I am grateful to Nick for his technical assistance and novel approach to statistical analysis. I am privileged to have had you both as advisors and mentors.

I gratefully acknowledge the support of Duncan Ward and the staff at Sydac Pty Ltd who donated their time to the development of the custom driving simulator used in this thesis and provided me with office space, support, and materials during the early stages of the project. Your generosity and expertise were sincerely appreciated.

I also acknowledge the staff at SimWorx for their advice and technical assistance during the latter stages of the project.

Thank you to my family, friends, and colleagues for your support throughout my candidature. The journey has been challenging at times but your encouragement and interest kept me motivated to achieve my goals.

Most of all, thank you to all of the participants who volunteered their time. I am particularly thankful to those participants who assisted in recruiting others to the project. I am grateful to have met many kind, generous and interesting people during this process. This thesis would not have been possible without your enthusiastic participation.

Dedication

This thesis is dedicated to my parents

Susan and Anthony

and my grandparents

Ivan and Audrey

Chapter 1.
Introduction

Background

Older drivers represent an increasing proportion of active drivers and, as the longevity in the general population increases, more older drivers are predicted to hold and maintain a driver's license than in any previous generation (Koppel & Berecki-Gisolf, 2015; Lyman, Ferguson, Braver, & Williams, 2002; Organisation for Economic Co-operation and Development [OECD], 2001). Older drivers have often been reported to be over-represented in serious crashes, especially when crash rate is considered per distance driven (Li, Braver, & Chen, 2003; OECD, 2001), and are more likely to be considered at-fault compared to younger drivers (L. E. Hakamies-Blomqvist, 1993; McGwin Jr & Brown, 1999). If they do become involved in a crash, older drivers are at increased risk of serious injury or death because of their greater fragility; more than half of the risk of involvement in a fatal crash can be attributed to fragility (Li et al., 2003). Older drivers are also susceptible to the "low mileage bias", whereby drivers travelling shorter distances are more likely to be involved in crashes (L. Hakamies-Blomqvist, Raitanen, & O'Neill, 2002; Langford, Methorst, & Hakamies-Blomqvist, 2006).

It is now recognised that mandatory age-based testing for older drivers, as defined to date by simple medical screening procedures, is not an effective strategy for reducing crash risk (Fildes et al., 2008). For example, a series of studies by Langford and colleagues (Langford, Bohensky, Koppel, & Newstead, 2008; Langford, Fitzharris, Newstead, & Koppel, 2004; Langford, Fitzharris, Koppel, & Newstead, 2004) compared crash involvement in Victoria, Australia, which does not require mandatory age-based driver testing, with other Australian states, which did have mandatory testing. The results from these studies, which primarily focused on a comparison between Victoria and New South Wales, indicated that older drivers in Victoria posed no more risk to themselves or others than older drivers in New South Wales, even after accounting for population, distance driven, time spent driving, and number of licenses held. In the United States, Grabowski, Campbell, and Morrisey (2004) reviewed the effectiveness of different driver re-assessment protocols on reducing crash fatalities, and found that the only policy measure related to reduced fatalities was in-person license renewal, while other policies including vision tests and mandatory on-road tests were not related to fatality rate.

That mandatory age-based testing has not been shown to reduce crash risk among older drivers has resulted in research directed towards identifying valid predictors of safe driving. The aim of this thesis was to investigate functional predictors of driving performance in older adults, particularly cognitive predictors. Because of the apparent risk faced by older drivers, much research has focussed on identifying those older drivers who may be most at-risk. Research has identified several functional correlates of driving ability in older drivers, including cognitive function, visual function, and physical function. Alternative licensing models have been proposed which feature several levels of screening and assessment for drivers deemed to be potentially at risk. Fundamental to the research program reported here is the conclusion from this review than cognitive abilities tests have potential applications for screening at-risk older drivers. The following sections contain a review of current and proposed licensing and assessment procedures, and the relationship of functional predictors to safe driving ability in older adults.

Licensing and assessment

Current Australian procedures

Licensing for older drivers varies by jurisdiction in Australia. Compulsory renewal procedures for each Australian State and Territory are summarised in Table 1.1. In 2015, the South Australian government abolished mandatory age-based testing for older drivers. Prior to this change, drivers aged 75 and older were required to report for annual medical testing, a process that focussed on the presence of medical conditions with potential to impair safe driving, visual acuity, and visual fields. However, under the new rules, drivers aged 75 and older who hold a "C (Car)" Class Driver's License are required to complete a self-assessment form annually[1]. This form is designed to help drivers think about how their physical and mental condition may affect their driving ability. If the driver reports the presence of certain medical conditions (e.g. heart problems, high blood pressure, severe arthritis, neurological conditions, mental health conditions, eye or vision conditions), they must take the form to a medical practitioner for further assessment. Regular compulsory medical assessments are still required for drivers aged 70 or older who hold a license with a class other than "C (Car)" (e.g. heavy vehicle or motorbike license), and for all drivers with a medical condition recorded against their license, which is subject to periodic review. Practical Driving Assessments are required where recommended by a medical practitioner, and are required annually for drivers aged 85 or older who hold a license with a class other than "C (Car)".

Several states have recently reviewed licensing procedures for older drivers and compulsory practical tests are no longer required in most jurisdictions. New South Wales is now the only jurisdiction that requires compulsory age-based practical driving tests for older drivers to hold a car license, and only when the driver wishes to maintain an unrestricted license. Drivers may forego this requirement if they hold a modified license, the conditions of which are determined in consultation with a medical practitioner and the licensing authority (conditions may include driving at certain times only, or driving only within certain areas). South Australia, Western Australia, and New South Wales have age-based compulsory practical testing for license classes other than "car". Compulsory medical assessments of some form are required in all jurisdictions except Victoria, Tasmania, or Northern Territory; however, drivers in all jurisdictions (regardless of age) are expected to self-assess their medical fitness to drive and self-report any medical concerns.

Vision requirements. All drivers in Australia must meet certain vision standards in order to hold a driver's license. The vision standards are specified in the Austroads Guidelines (Austroads, 2012)[2]. In most cases, drivers must demonstrate sufficient

---

[1]Self Assessment for C (Car) License Holders: `http://www.sa.gov.au/__data/assets/pdf_file/0004/168142/MR-1562-Medical-fitness-self-assessment-form-Nov-2015-.pdf`

[2]The Austroads guidelines described in this thesis were current at the time of thesis submission in September 2016. An updated version was released effective October 2016 (Austroads, 2016)

Table 1.1: License renewal procedures for older drivers in Australian States and Territories (current June 2016)

State/Territory	Medical Assessment	Practical Assessment
South Australia	Annual completion of medical self-assessment questionnaire from age 75	Annually for holders of license other than "car" from age 85
Victoria	No compulsory assessment	No compulsory assessment
Tasmania	No compulsory assessment	No compulsory assessment
Northern Territory	No compulsory assessment	No compulsory assessment
Western Australia	Annual medical assessment from age 80	Annually from age 85 for license other than "car"
Queensland	Required to hold certification of medical fitness to drive from age 75; must be reviewed annually	No compulsory assessment
Australian Capital Territory	Annual medical assessment from age 75	No compulsory assessment
New South Wales	Annual medical assessment from age 75	Required every two years from age 85 to maintain unrestricted license (practical not compulsory for modified license); Required annually from age 70 for multi-combination license; Required annually from age 80 for license other than car or rider

visual acuity and visual fields in order to hold a license. In the first instance, drivers will be assessed by a medical General Practitioner (GP) or other qualified health professional, and may be referred to an optometrist or ophthalmologist if further assessment is required.

For visual acuity, the requirement is a score of at least 6/12 (equivalent to 20/40, or logarithm of the Minimum Angle of Resolution (logMAR) 0.5) as tested on a standard visual acuity chart, such as a Snellen chart or logMAR chart. Vision correction is allowed. In cases where visual acuity is just below the required standard, private vehicle drivers may be referred to an optometrist or ophthalmologist who may recommend that a conditional license be granted. A driver's license will not be granted in any case where visual acuity is worse than 6/24 in the better eye. The guidelines also suggest that other visual assessments such as contrast sensitivity, or other specialised tests, may be used to help with assessment.

Visual fields are screened, in the first instance, by a method called "confrontation". This is an inexact measurement in which the health professional will sit opposite the patient and ask them to count the number of fingers held up in each of the four corners of the visual field while fixating on the non-occluded eye of the health professional. If a patient demonstrates visual field impairment or is suspected of having a visual field defect, they should be referred for assessment by an optometrist or ophthalmologist. Specialist assessment will involve automated perimetry using an automated static perimeter (Kinetic Goldman Visual Field, Humphrey Field Analyser, Medmont M700, Octopus, etc.). If the automated perimetry assessment suggests impairment, the Esterman binocular field test should be performed. The visual field requirement is a horizontal extent of at least 110 degrees within 10 degrees above and below the horizontal midline. Additionally, an unconditional driver's license will not be granted if there is any significant visual field loss within a central radius of 20 degrees of foveal fixation, or other significant visual field loss that is likely to impede driving performance. A conditional license may be considered based on information provided by the optometrist or ophthalmologist.

Vision requirements for driving are similar around the world. A review of vision requirements for licensure in Germany, Spain, Italy, France, the UK, and the US, indicated that all countries had visual acuity limits (Bron et al., 2010). Moreover, it was found that the limit was usually 0.5 (equivalent to the Australian limit of 6/12), with some minor variations; for example, in the UK drivers were required to pass a specialised "Number plate test", which is roughly equivalent to 6/10 (Bron et al., 2010; Kotecha, Spratt, & Viswanathan, 2008).

Health requirements. South Australian legislation requires mandatory reporting of medical conditions to the licensing authority by health professionals. This is also the case in the Northern Territory. In other states, reporting is not mandatory but professionals are able to take action without the consent of the patient if they believe the patient is unfit to drive and are protected from civil and criminal liability in such cases. All health professionals are expected to assess patients according to the standards specified in the Austroads guidelines. The Austroads guidelines contain medical standards for fitness to drive and management of conditions that can affect fitness to drive. Medical standards are described for a wide range of conditions, including blackouts, cardiovascular conditions, diabetes, hearing, musculoskeletal conditions, neurological conditions, psychiatric conditions, sleep disorders, substance misuse, and vision and eye disorders.

The Austroads guidelines describe how specific medical conditions can affect the ability to drive safely. Diagnosis-specific guidelines are provided for determining fitness-to-drive. Conditions are described in the categories of blackouts, cardiovascular conditions, diabetes, hearing, musculoskeletal conditions, neurological conditions, psychiatric conditions, sleep disorders, substance misuse, and vision and eye disorders. The medical guidelines provide clear advice to health professionals in regards to specific, individual conditions. When a driver has multiple conditions (including general age-related decline, or declines associated with degenerative disease), health practitioners are advised to exercise professional clinical judgement to determine the effects on safe driving ability. The health professional may consider the standards for each individual condition, but must also consider the possible compounding effect of the multiple conditions.

Cognitive ability. Currently, there are no specific standards relating to cognitive abilities that are required to hold a driver's license. However, the Austroads guidelines acknowledge that driving is a complex task that is dependent on many cognitive processes including visuospatial perception, attention, concentration, memory, and reaction time. Furthermore, it is stated that many conditions, as well as age-related change, can affect these cognitive abilities. Health professionals are advised to consider changes to cognitive abilities in their assessments. Some broad guidelines are given with respect to the cognitive demands of the driving task. For example, driver reaction time must be sufficient to be able to react quickly to unexpected hazards, and to perform basic tasks such as stopping, turning, and speeding up in a timely manner. Declines in problem solving skills and memory may contribute to problems with navigation. Problems with judgement and decision making may cause difficulties with appropriate gap selection, misjudging the speed of other vehicles, and failing to give way. Cognitive impairment is particularly relevant for patients diagnosed with degenerative neurological conditions such as dementia; drivers with dementia are not permitted to hold an unrestricted license, but may be considered for conditional licenses subject to annual review.

Although cognitive abilities are not currently routinely assessed or screened for driver licensing, research has demonstrated clear links between various cognitive abilities and driving safety (see section: Functional predictors of driving). Based on this, proposed models for driver assessment have featured functional screening tests as key components for making fitness to drive decisions. Two such models will be described in the next section.

Proposed models for older driver assessment and licensing

Australasian model. Austroads commissioned the development of a model license re-assessment procedure for older drivers. The development of the model has been described in a series of Austroads reports (Charlton et al., 2008; Fildes et al., 2004, 2000; Langford et al., 2009). An outline of the model is shown in Figure 1.1. The model is partly based on the American National Highway Traffic Safety Administration (NHTSA) model (Staplin, Lococo, Gish, & Decina, 2003) which will be briefly described in the next section. Both models are also consistent with the OECD recommendation that testing for older drivers will be most effective and equitable if only those drivers who pose a significant risk to others are targeted (OECD, 2001).

The model relies on at-risk drivers being referred for assessment by various community members, for example medical and health practitioners, family members, police, and the licensing authority; this is Level 1 of the Model. The model does not rely on mandatory age-based testing; instead, only those drivers considered to be high-risk are referred into the system. Level 2 of the model comprises a medical assessment and validated off-road screening tests. All drivers referred are required to undergo a comprehensive medical assessment before proceeding further through the system. Drivers are then required to undergo validated off-road screening tests intended to assess their functional ability and identify drivers who are at increased risk of crashing. Results on the tests would be used to make a decision of "fit to drive", "unfit to drive", or "unclear" (further assessment required). Drivers whose screening results indicate they are clearly fit to drive will be free to continue driving. Drivers whose screening results clearly indicate that they are unfit to drive will be invited to surrender their license, or may be eligible for a restricted license, or temporary suspension pending treatment or rehabilitation. Drivers with unclear results will progress to Level 3 of the model: a Specialist Assessment for drivers with medical concerns, and/or an on-road driving test.

The Specialist Assessment provides further information about the driver, which will be considered by the licensing authority. If the results indicate that the driver is unfit to drive, the driver's license may be cancelled or granted on a restricted basis. If the driver has a temporary medical condition affecting their driving, the driver's license may be temporarily suspended pending treatment, rehabilitation, and re-assessment. If the results are inconclusive regarding medical impairments, the driver will be referred for an On-Road Driving Test. The On-Road Driving Test can be conducted by the Licensing Authority or an Occupational Therapist. The results of the On-Road Driving Test will provide further information regarding fitness to drive.

A key factor of the model is that the Screening Tests at Level 2 must reliably and validly predict driving outcomes. Later stages of the Austroads project focussed on investigating potential screening tests for each level of the model, and particularly on identifying valid functional screening tests for Level 2 of the model.

Stage 2 of the model development project (Fildes et al., 2004) investigated three screening tests: the Gross Impairment Screening Battery of General Physical and Mental Abilities (GRIMPS), the California Test (CALTEST), and the DriveABLE test. These tests were evaluated for their ability to predict driving performance as assessed by an on-road driving test. The participants involved in the study were 1000 drivers in New Zealand aged 79 or older who were required to undertake an on-road driving test to renew their license. The target for a screening test to be deemed acceptable was that it must

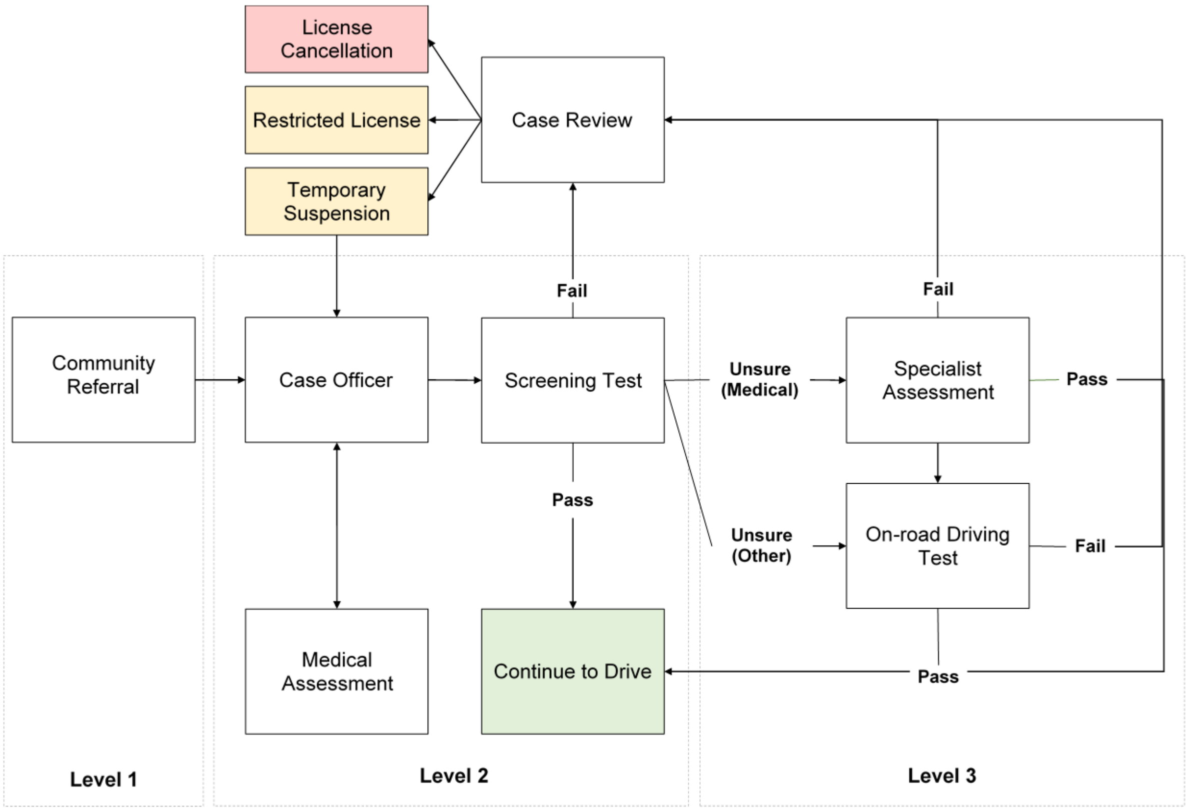

Figure 1.1: Proposed Australasian License Re-Assessment Procedure (Fildes et al., 2000)

identify “safe” drivers with at least 80% probability, “unsafe” drivers with at least 65% probability, and a maximum of 50% of drivers would be classified as “doubtful”. Both the GRIMPS and the CALTEST met these criteria with regards to predicting performance as rated on the on-road driving test.

Stage 3 of the model development project (Charlton et al., 2008) investigated three screening tests: the Health Screen for Drivers (overall score, and scores from Visualisation of Missing Information (VMI) and Months Backwards), the Useful Field of View Test™ (UFOV™), and AutoTrails. The UFOV™ and Autotrails are components of the CALTEST battery. A case-control study design was implemented. Cases (n = 10) were aged 80 or over and had been involved in an at-fault crash in the previous 6 months; matched controls (n = 10) had no recent crash involvement. The results indicated that the crash-involved cases were more likely than the controls to fail the VMI and the UFOV™, and some components of the Autotrails. Based on the small sample size, it was concluded that the tests had the potential to detect significant differences in a larger sample.

In Stage 4 of the model development project (Langford et al., 2009), a case-control study design was implemented. Cases (n = 62) were aged 75 or over and had been involved in an at-fault casualty crash in the past 12 months; matched controls (n = 62) had no recent crash involvement. The tests considered were a version of the VMI, Months Backwards, Health Screen for Drivers, Ascending Trails, Descending Trails, and UFOV™ (all subtests, and UFOV™ Total Score). The results indicated that when the tests were used to give a pass/fail result, there were no statistically significant ($p < .05$) differences between the cases and the controls on any of the tests. When the tests were used to give a pass/fail/unsure result, four of the tests failed a greater proportion of cases than controls (Months Backwards, Ascending Trails, UFOV™ Subtest 2, and UFOV™ Total Score). However, only Ascending Trails reached statistical significance ($p < .05$). The screening tests were also evaluated for their ability to predict occupational therapist ratings of driver health, and driving assessor ratings of performance on a structured on-road driving test. It was found that all tests showed indicative differences between cases and controls for occupational therapist ratings, although only UFOV™ Subtest 3 and Descending Trails showed statistically significant differences. For the driving assessor ratings, 10 of the 14 protocols showed differences in the expected direction, but none reached statistical significance ($p < .05$).

Overall, the conclusion from the Austroads project regarding the Level 2 screening tests was that none of the tests considered have so far demonstrated sufficient validity to be useful in the model. Although statistical associations were observed between some of the predictor tests and the driving measures, the results were not strong enough to be reliably predictive. Bédard, Weaver, Darzinš, and Porter (2008) have argued that a lack of predictive value is an issue for driver screening tests often considered in the literature, including the UFOV™, Mini Mental State Examination (MMSE), and Trail Making Test Part A (TMT-A). More recently, a large prospective cohort study, Candrive/Ozcandrive, has examined the predictive validity of tools for assessing fitness to drive using a sample of drivers aged over 70 in Australia, Canada, and New Zealand, and aims to develop an in-office screening tool for identifying unsafe drivers (Marshall et al., 2013). The study is ongoing.

NHTSA model. The NHTSA model (Staplin et al., 2003) has many similar features to the Austroads model. The model is shown in Figure 1.2. Like the Austroads model, it relies on at-risk drivers being referred into the system by various referral sources, including law enforcement, health professionals, social services, family or friends of the driver, other citizens, the driver themselves, or the driver's physician. Drivers are referred to a Department of Motor Vehicles (DMV)/Medical Advisory Board (MAB) Case Manager. The Case Manager is responsible for counselling the driver, and compiling a case file containing relevant information about the driver, including forms completed by the drivers, crash and conviction data, and medical history (completed by the driver and their physician). The driver will also be required to undergo functional screening tests, the results of which will be included in the case file. The case file will then be sent to an approved MAB (or equivalent) physician for review.

Based on the results of the functional screening tests and other information contained in the case file, the MAB physician will make a fitness-to-drive determination. The driver may be cleared to continue driving, or a license restriction or suspension may be recommended. All drivers at this stage should be provided with counselling. If the MAB physician is unable to make a fitness-to-drive determination, more extensive testing from various sources is required. Depending on the case, the driver may be required to attend an in-person interview with the MAB physician. Alternatively, or in addition to

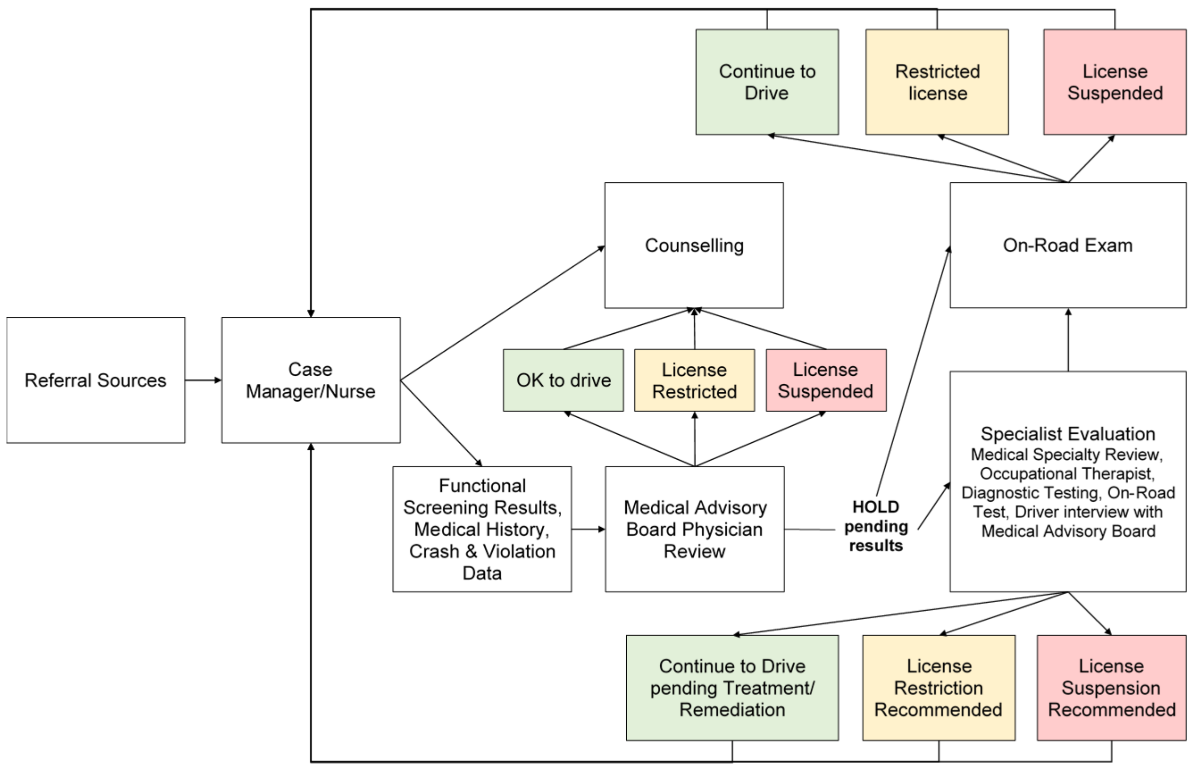

Figure 1.2: NHTSA Driver Screening and Evaluation Model (Staplin et al., 2003)

the MAB physician interview, the driver may be required to undergo an on-road driving assessment. The driver may also be required to undergo specialist assessment, for example evaluation by an occupational therapist or other medical specialist for the purpose of evaluating specific medical conditions, or a driver rehabilitation specialist. The results of the on-road test and/or the specialist evaluation will be used to make a fitness-to-drive determination, or recommendation for treatment or rehabilitation.

Like the Austroads model, the NHTSA model features functional screening tests at an early stage of the model as a tool for determining fitness-to-drive. The report recommends that results from screening tests provide two cut-points: a "prevention threshold" and an "intervention threshold". Scores below (i.e. worse than) the prevention threshold indicate that the driver needs further assessment, while scores below the intervention threshold indicate more severe impairment that should be assessed urgently. It is suggested that screening tests should assess three categories: visual function, physical ability, and mental function. The inclusion of screening tests is intended to reduce the number of drivers requiring more extensive testing to determine fitness-to-drive.

Several potential screening tests were assessed in a pilot study with $N = 2508$ drivers aged over 55 recruited from several sources: a Residential Community ($n = 266$), a Senior Centre ($n = 113$), Medical Referrals ($n = 366$) and drivers presenting for License Renewal ($n = 1876$). Six perceptual-cognitive tests were included (Motor-Free Visual Perception Test (MVPT) (Closure Subtest), Trail Making Test Part B (TMT-B), Dynamic Trails, UFOV™ Subtest 2, Delayed Recall, and Scan Test). Four physical abilities tests were also included (Rapid Pace Walk, Foot Tap Test, Head/Neck Rotation, and Arm Reach Test). The tests were evaluated for their ability to predict at-fault crash involvement retrospectively (one year prior to assessment) and prospectively (follow-up approximately 20 months following assessment). For at-fault crashes, the results showed that drivers who scored five or more errors on the MVPT were 4.96 times more likely to record a crash; drivers who scored 80 seconds or longer on TMT-B were 3.50 times more likely to record a crash; drivers who made three errors on Delayed Recall were 2.92 times more likely to record a crash; and drivers who scored 300 ms or longer on UFOV™ Subtest 2 were 2.48 times more likely to crash. Of the physical ability tests, Rapid Pace Walk and Head/Neck Rotation were significantly associated with crash risk.

The results were further analysed with a follow-up period of between 4.18 to 5.13

years (Ball et al., 2006). Only the License Renewal subsample ($n = 1910$)[3] was analysed. At-fault crashes were recorded for $n = 92$ drivers. Of the cognitive assessments, it was found that TMT-B, MVPT, and UFOV™ Subtest 2 were associated with at-fault crashes. Examining various cut points showed that drivers who made four or more errors on the MVPT were 2.10 times more likely to crash; drivers who took 147 seconds or more to complete TMT-B were 2.01 times more likely to crash; and drivers who scored 353 ms or longer on UFOV™ Subtest 2 were 2.02 times more likely to crash. When accounting for age, gender, and annual mileage, MVPT and UFOV™ Subtest 2 were still significantly associated with at-fault crashes (OR = 1.24 and 1.23 respectively), although TMT-B was not.

Functional predictors of driving

The factors contributing to Driving Behaviour are summarised in Anstey, Wood, Lord, and Walker's (2005), shown in Figure 1.3. Much research has focused on functional predictors of driving outcomes, but a multi-disciplinary approach has often been lacking. The Multi-Factorial Model of Driving Safety was formulated based on a literature review of sensory, cognitive, and physical/medical factors that are associated with driving outcomes (driving performance and/or crash risk) in older adults, and are prone to decline in normal aging. It should be noted that the driving literature includes a range of different driving outcome measures, including simulated and real-world (naturalistic or standardised assessment) driving performance assessed through various dependant measures such as Brake RT, steering variability, lane keeping, or timed headway; infringements (e.g. speeding); and crash involvement (at-fault vs not at-fault, self-reported vs official records).

In the category of cognitive function, Anstey et al. (2005) identified several factors that are relevant for driving, including attention, Reaction Time (RT), processing speed, and executive function. The most commonly assessed visual function measures were visual acuity and contrast sensitivity. Although adequate visual function is needed to process visual information from the driving scene accurately, the results suggested that

[3] For the License Renewal sample, $N = 2381$ drivers participated in screening activities; $n = 1876$ cases were analysed by Staplin et al. (2003), $n = 1910$ complete cases were analysed by Ball et al. (2006)

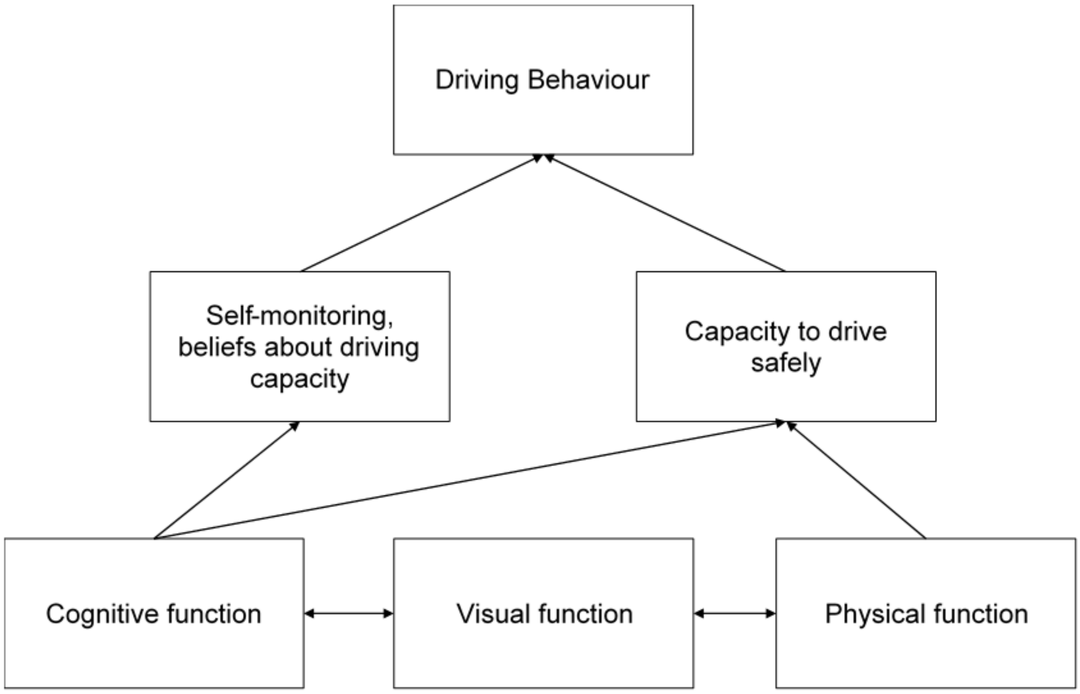

Figure 1.3: Multi-factorial model of driving safety (Anstey et al., 2005)

these tests in isolation were not good predictors of driving outcomes. In the category of physical function, several health, medical, and physical function measures were identified. Evidence relating to the physical function factors and most of the visual function factors was generally found to be weak and inconsistent, but several of the cognitive factors were consistently linked to driving performance.

The model separates the “capacity to drive safely” from actual “driving behaviour”. A person’s capacity to drive safely depends on inter-correlated cognitive, visual, and physical function factors and can be assessed using off-road functional screening tests. A person’s actual driving behaviour depends on both the capacity to drive safely, and beliefs about driving capacity, accuracy of self-monitoring, and insight into functional ability. Many older adults are able to regulate their driving appropriately as they age, for example by only driving at certain times, or avoiding challenging situations. Driving behaviour can be assessed with an on-road driving test.

The following section contains a summary of cognitive, vision, and physical function factors identified in the Multi-Factorial Model of Driving Safety, as well as other relevant factors identified in the literature.

Cognitive function

Visual attention. Visual attention is used to detect and process potentially important visual events in the environment, for example identifying the presence and location of other cars, pedestrians and potential hazards on the road (Ball, Owsley, Sloane, Roenker, & Bruni, 1993). Selective attention is required to ignore irrelevant information in the scene; and, divided attention allows the driver to attend to potential hazards in the environment while simultaneously focussing on the main driving task. Thus, attentional problems are often linked to poor driving performance and crash risk. Visual attention has often been measured by the UFOV™, which was developed specifically for the assessment of older drivers, and contains three subtests intended to assess processing speed, divided attention, and selective attention. Other measures of visual attention include the number cancellation task, tracking tasks, change detection tasks, and visual search tasks.

In a meta-analysis of cognitive predictors of unsafe driving in older adults, Mathias and Lucas (2009) reported that attention was the most commonly assessed cognitive domain. Tests featuring an attention component were associated with on-road driving performance, simulated driving performance, and crash risk. Similarly, Anstey et al. (2005) found that several different measures of attention, including tests of selective attention, divided attention, and visual attention, were consistently related to crash history and on-road driving performance. The UFOV™ in particular has been shown to be associated with a wide range of driving outcomes, including retrospective and prospective crash involvement, on-road driving performance, and driving simulator performance in older adults (Clay et al., 2005; Gentzler & Smither, 2012). Aspects of attention declinine as part of the normal ageing process, and general attention deficits are also a symptom of early dementia or mild cognitive impairment (Adler, Rottunda, & Dysken, 2005). Adler et al. (2005) found that in a battery or neuropsychological tests administered to older drivers with dementia, tests of attention, visuospatial skills, and RT were most related to driving performance.

Perceptual and visuospatial abilities. Perceptual and visuospatial skills are important for positioning the vehicle correctly on the road, executing common manoeuvres such as changing lanes and stopping at traffic lights, judging positions and distances of other items in the environment, and predicting the movement of traffic. Tests in this category include movement perception tests, the Paper Folding test, the Hooper organisation test, the Benton Line Orientation Task, and the Ergovision test. Many of these tests also involve a speed of processing component.

Anstey et al. (2005) reported that tests of visuospatial ability displayed significant low to moderate correlations with crash history and on-road test performance. Mathias and Lucas (2009) reported several moderate to large effect sizes for tests of perception and visuospatial abilities (e.g. Ergovision, Paper Folding Task, Benton Line Orientation Task). In a battery of cognitive, visual, and motor skills, Dawson, Uc, Anderson, Johnson, and Rizzo (2010) found that performance on an on-road test was best predicted by tests that reflected visuospatial and visuomotor abilities. For drivers with dementia, a meta-analysis indicated that, out of a battery of cognitive tests, measures of visuospatial skills were best related to driving outcomes including on-road assessment, driving simulator performance, and caregiver reports (Reger et al., 2004).

Processing speed and Reaction Time (RT). RT and processing speed have high face validity for the driving task. For example, drivers are required to process information in the environment quickly and efficiently under time constraints in order to assess hazards in the environment and act appropriately to avoid collisions. Common tests of RT and processing speed used in driving research include simple RT tasks, choice RT tasks, complex RT tasks, the TMT-A, UFOV™ Subtest 1, and the Digit Symbol Substitution test. The UFOV™ (all subtests) is also dependent on processing speed (J. D. Edwards, Vance, et al., 2005).

Associations between simple RT tasks and driving performance are generally small, but are larger for complex RT tasks (Anstey et al., 2005). Mathias and Lucas (2009) reported that several measures of processing speed and RT (including Complex RT tasks, Simple RT tasks, UFOV™ Subtest 1, and TMT-A) were associated with driving outcomes. Complex RT tasks were associated with on-road driving performance (effect size Cohen's $D = 1.32$, one study) and crash risk (effect size Cohen's $D = 0.64$, three studies). Most of the tasks were associated with small to moderate effect sizes.

Wood, Anstey, Kerr, Lacherez, and Lord (2008) reported that, in a battery of cognitive tests, faster choice RT was the best predictor of on-road driving performance.

Anstey, Horswill, Wood, and Hatherly (2012) conducted a factor analysis on a large battery of cognitive tests and found that a processing speed/executive functioning factor emerged, rather than a pure processing speed factor as expected. This complex factor was found to have the largest effect on the performance outcome, and the authors suggested that this reflects that the two constructs captured by this factor are inextricably linked and are both of great importance to driving.

Executive function. The theoretical construct "executive functioning" encompasses a range of higher-order cognitive processes, including initiation, planning, hypothesis generation, cognitive flexibility, decision making, regulation, judgement, feedback utilisation, and self-perception (Asimakopulos et al., 2012; Daigneault, Joly, & Frigon, 2002). These processes allow the driver to supervise and assess the environment, make decisions, and assess outcomes according to feedback from the environment. Drivers with an executive functioning problem may make inappropriate and dangerous decisions in high-risk situations, be unable to monitor and adjust their actions adequately, and may lack insight into their own abilities, decisions, and cognitive deficits. Examples of commonly used Executive Function tests are TMT-B, the Stroop Colour Word Test, the Wisconsin Card Sort test, the Paper Folding Task, and the Ray-Osterrieth Complex Figure Test.

There is overlap between executive function and other cognitive domains. For example, working memory has been proposed to be central to executive function (Baddeley & Hitch, 1974), and tests of general cognitive function such as the Mini Mental State Examination (MMSE) contain components reflecting executive functioning. TMT-B, which is frequently used in driving studies, reflects executive functioning, but also relies on processing speed and other aspects of attention.

Mathias and Lucas (2009) identified three studies that investigated the association between TMT-B and on-road driving performance. The meta-analysis indicated a moderate mean effect size (Cohen's D) of 0.79 for the difference favouring drivers who passed compared with those who failed an on-road assessment. Seven studies were identified that investigated the association between TMT-B and crash involvement. The meta-analysis reported a small mean effect size (Cohen's D) of 0.17 for the difference between crash-involved and non-crash involved drivers, favouring the latter. Similarly, Anstey et al. (2005) reported an inconsistent relationship between TMT-B and driving outcomes. Of the four included studies, two reported a significant association favouring better TMT-B performance, and two reported no association.

Daigneault et al. (2002) compared the executive functioning of older adults with a history of three or more crashes in the past three years to that in a crash-free control group. The study found that the crash-involved drivers performed significantly worse on several measures of executive functioning, including the Stroop Colour Word Test, the Wisconsin Card Sort test, and the Tower of London test.

Mental status. Tests of Mental Status assess aspects of general cognitive functioning including attention, concentration, memory, visuospatial reasoning, and executive function. These functions typically decline with age, and serious declines in these areas are characteristic of dementia or mild cognitive impairment. Common Mental Status assessments include the MMSE, the Mattis Organic Mental Syndrome Screening Examination (MOMSSE), and the Short Blessed Test.

Anstey et al. (2005) reported that tests of mental status showed inconsistent associations with driving outcomes. While some studies reported an association between mental status test scores and crash risk, other studies reported no association. It was noted that these tests have strong ceiling effects in normal samples. Mathias and Lucas (2009) reported that the MMSE (a test of general mental ability) showed moderate differences between pass/fail drivers on on-road and driving simulator assessments, but the effect size was smaller than the effect sizes for other cognitive functions (e.g. attention, perception, and reasoning).

In a review of older drivers with dementia, Adler et al. (2005) found a consistent relationship between mental status and driving, with most studies reporting significant, modest correlations between total MMSE scores and performance on on-road and simulator assessment. On the other hand, some studies have found no association between MMSE score and driving outcomes in drivers with dementia (Molnar, Patel, Marshall, Man-Son-Hing, & Wilson, 2006). Recent studies have also reported no association between MMSE scores and driving outcomes in non-impaired samples (Crizzle, Classen, Bédard, Lanford, & Winter, 2012; Joseph et al., 2014).

Visual function

Visual acuity. Despite the worldwide application of visual acuity requirements for driver licenscing, reviews tend to suggest a very small and inconsistent relationship between visual acuity and driving outcomes (Anstey et al., 2005; Owsley & McGwin Jr, 1999, 2010). Wood (2002) found that there was no clear evidence linking visual acuity to crash risk. Some studies have demonstrated an association between simulated visual acuity impairment and road sign recognition and hazard avoidance, and between visual acuity impairment in drivers with macular degeneration and simulated driving performance (Owsley & McGwin Jr, 2010). More recently, studies have reported no association between visual acuity and on-road driving test performance in non-impaired drivers with visual acuity above 20/40, the standard required by most licensing authorities (Koppel et al., 2016; Wood, Horswill, Lacherez, & Anstey, 2013).

Overall, it is now generally recognised that visual acuity is not an adequate predictor of driving safety because it does not reflect the complex visual and cognitive skills needed to operate a motor vehicle safely (Anstey et al., 2005; Owsley & McGwin Jr, 2010). It has also been argued that currently enforced visual acuity standards may be too high, and that drivers with lower visual acuity may be able to hold a license subject to conditions such as reduced speed (Charman, 1997).

Contrast sensitivity. In Australia, contrast sensitivity is not routinely assessed as part of driving license requirements, but may be taken into consideration by a vision specialist to determine whether a license should be granted in cases where other vision requirements have not been met. Contrast sensitivity is important for identifying and recognising important stimuli in the driving environment, and, like visual acuity, has high face validity for the driving task. Anstey et al. (2005) reported in their review that most studies indicated that contrast sensitivity had a low association with crashes and on-road test performance. However, Ball and Rebok (1994) found that contrast sensitivity was a slightly better predictor of driving safety than visual acuity. Other studies have reported that contrast sensitivity is linked to self-reported difficulties with both day and night driving (Charman, 1997). In a recent review of vision and driving, Owsley and McGwin Jr (2010) reported that contrast sensitivity was more related to driving performance (i.e. as assessed by a driving simulator or on-road test) than to crash outcomes, although some studies have found that contrast sensitivity predicted past and future crashes. Relationships between contrast sensitivity and driving outcomes were more likely to be observed in clinical samples (e.g. drivers with cataract or partial blindness) compared to population-based samples. Drivers with impaired contrast sensitivity tend to reduce their driving exposure, with lower annual mileage and decreased number of places and trips per week (Sandlin, McGwin, & Owsley, 2013); additionally, impaired contrast sensitivity is a strong predictor of driving cessation (Emerson et al., 2012).

Visual fields and other aspects of vision. The Austroads guidelines specify minimum standards for visual fields to hold a driver's license. The research relating to visual fields and driving outcomes is mixed and the definition of "visual field loss" is inconsistent between studies (Owsley & McGwin Jr, 2010). Several studies have reported an association between severe visual field impairment and crash risk. However, several studies have also reported no association (Owsley & McGwin Jr, 2010). More recently, in a non-impaired sample of older drivers, Wood et al. (2013) have reported that visual fields were not associated with on-road driving test performance.

It is generally concluded that low-level vision assessments (such as contrast sensitivity and visual acuity) do not adequately reflect the complexity of the driving task. Furthermore, Owsley and McGwin Jr (2010) noted that it is difficult to assess the relationship between vision problems and crash involvement, because people with severe vision difficulty are more likely to censor or cease their driving and are less likely to be relicensed. In summary, measures of low-level visual function have good face validity for the driving task, but generally show only small associations with crash risk and driving performance, and these associations are most often observed in clinical samples of drivers with eye disease or severe impairment. Individual tests of visual function are unlikely to be useful as screening tests; however, they may be useful as part of a comprehensive assessment including other measures of functional performance that more adequately reflect the complexity of the driving task.

Physical function

Physical function and health. Anstey et al. (2005) reviewed aspects of physical function and medical conditions related to driving ability and found that, while some studies report an association between certain medical conditions or symptoms (e.g. heart disease, stroke, arthritis, recent fall, orthostatic systolic blood pressure drop) and driving outcomes, few studies found associations between measures of physical function and driving outcomes. An association between neck rotation and crash risk was reported, but other studies found no association between other measures of physical function (e.g. grip strength, trunk rotation, shoulder abduction, disability status) and driving outcomes. It was noted that drivers with physical function impairments may self-regulate their driving, and may be more aware of their deficits than drivers with cognitive or sensory impairments.

In developing the NHTSA License Assessment model described above, Staplin et al. (2003) conducted an extensive literature review and consultation with expert panel members to identify functional abilities related to crash risk. Three physical function factors were identified: proprioception and somatosensory processes, coordination of visual and motor processes, and strength and range of motion. Four measures of physical function were then selected for inclusion in the pilot study to examine the validity of functional screening measures in the license assessment model. The four measures were Rapid Pace Walk, Foot-Tap, Head-Neck Rotation, and Arm Reach. The analysis showed that the Rapid Pace Walk and Head-Neck Rotation were significantly related to crash risk (1 year retrospective and mean 20 months prospective).

In a recent systematic review of screening and assessment tools used by occupational therapists to determine fitness to drive, Dickerson, Meuel, Ridenour, and Cooper (2014) identified 21 studies that used physical function measures. The studies included both healthy and clinical (e.g. dementia, Parkinson's disease, stroke) samples. Among the physical function assessments were grip strength, range of motion, simple RT, head-neck rotation, finger tapping test, co-ordination, balance, strength, and the Rapid Pace Walk test. Four studies reported an association between the Rapid Pace Walk test and driving outcomes; however, one study reported that the Rapid Pace Walk test was not associated with driving outcomes. Associations were also reported for other physical function assessments, including self-rated health, the 360 degree turn test, finger-tapping test (three studies), head-neck rotation, postural sway, and knee extension strength, but Dickerson et al. (2014) reported that the associations were generally weak. Several studies reported no association between measures of physical or motor function and driving outcomes. It was noted that measures of physical and motor function were most useful when included in a battery of other tests.

Physical fitness. Physical activity and fitness may also contribute to driving safety through an association with cognitive functions relevant for driving. A study by Marmeleira, Ferreira, Melo, and Godinho (2012) showed that self-reported physical activity, as measured by the International Physical Activity Questionnaire, was associated with better performance on tests of processing speed and divided attention. Similarly, Roth, Goode, Clay, and Ball (2003) reported that two measures of self-reported physical activity, the Exercise Participation Questionnaire and the Physical Activity Scale for the Elderly (PASE), were both significantly correlated with UFOV™ performance. Meta-analyses and reviews have generally indicated that physical activity is positively associated with brain function and cognitive ability in older adults (Bherer, Erickson, & Liu-Ambrose, 2013; Erickson, Hillman, & Kramer, 2015; Marmeleira, 2012; Stine-Morrow & Basak, 2011), and that physical activity interventions for older adults (particularly aerobic or combined aerobic/strength interventions) can have a positive effect on cognitive performance across several domains (Carvalho, Rea, Parimon, & Cusack, 2014; Colcombe & Kramer, 2003). However, in contrast, two recent reviews and meta-analyses of Randomised Control Trials have failed to find consistent evidence of cognitive benefit from physical exercise (Kelly et al., 2014; J. Young, Angevaren, Rusted, & Tabet, 2015).

Few studies have investigated the direct relationship between physical activity and driving. Marottoli et al. (2007) conducted a randomised control trial to investigate if participation in a physical conditioning program could improve driving performance in older drivers with physical impairments such as neck rotation, trunk rotation, or gait speed. The intervention targeted flexibility, coordination, and speed of movement. The results showed that, compared to the control group, the intervention group were able to maintain their driving performance on an on-road driving assessment, while the driving performance of the control group declined (Marottoli et al., 2007). In a pilot study

Caragata, Tuokko, and Damini (2009) found that older drivers who participated in a fitness program showed gains in physical function measures, and also reported that they felt their driving skills had improved and their driving confidence was higher. These results indicate that physical activity may be an important factor in predicting safe driving.

Assessment of older drivers

Assessment methods

The driving ability of older adults is most often measured in one of three ways: performance on an on-road driving assessment, performance on a simulated driving assessment, and history of crashes or violations. An on-road driving assessment is generally considered to be the "gold standard" because the driver's actual driving performance and competence can be observed (Shechtman, 2010). Crash history is an important indicator of driving safety; however, crashes are a relatively rare event, which can make it difficult to observe associations between risk factors and crash outcomes, particularly in small samples. Crash data may be obtained from law enforcement or motor vehicle administration authorities, or from self-report. Both sources are potentially inaccurate. Drivers may not accurately report crash involvement. State-recorded data include only those crashes that have been reported to authorities, and may not contain accurate information about severity or fault. Driving simulators have several advantages for assessing driving performance of older drivers. They are safer than on-road driving, allow dangerous and unusual situations to be assessed, and provide a consistent and repeatable test environment. They also avoid the cost, space, and personnel requirements of on-road testing (Allen, Rosenthal, & Cook, 2011; Classen, Bewernitz, & Shechtman, 2011; Classen & Brooks, 2014).

Driving simulators

Driving simulators are now widely available and frequently used for research, training, and assessment (Allen et al., 2011; Classen & Brooks, 2014; Crisler et al., 2011; Dickerson et al., 2014; Pollatsek, Vlakveld, Kappe, Pradhan, & Fisher, 2011). The availability of lower-cost options means that driving simulators are now increasingly accessible to researchers and results have indicated that, in certain situations, lower-fidelity simulators can produce results that are comparable to high cost, high fidelity simulators (Gibbons, Mullen, Weaver, Reguly, & Bédard, 2014; Lemieux, Stinchcombe, Gagnon, & Bédard, 2014).

Driving simulators have been successfully used for various applications including research, assessment and training. For example, simulators have been used for re-training older drivers and clinical patients (Casutt, Theill, Martin, Keller, & Jäncke, 2014; Pollatsek, Romoser, & Fisher, 2012; Unsworth & Baker, 2014) and for training novice drivers (Allen, Park, Cook, & Fiorentino, 2012; de Winter et al., 2009; Pollatsek et al., 2011). They have also been widely used to investigate the relationship between cognitive abilities and driving performance (Bélanger, Gagnon, & Yamin, 2010; Hoffman, Atchley, McDowd, & Dubinsky, 2005; Shanmugaratnam, Kass, & Arruda, 2010) and the effects of cognitive interventions on driving performance (Roenker, Cissell, Ball, Wadley, & Edwards, 2003). Driving simulators have been effectively used in different populations, including older drivers (Hoffman & McDowd, 2010; Horberry, Anderson, Regan, Triggs, & Brown, 2006; Lee, Cameron, & Lee, 2003; Martin et al., 2010; Stinchcombe & Gagnon, 2013), and clinical groups including patients with cognitive impairment (Devlin, McGillivray, Charlton, Lowndes, & Etienne, 2012; Frittelli et al., 2009), HIV (Vance, Fazeli, Ball, Slater, & Ross, 2014), diabetes (Cox, Gonder-Frederick, Kovatchev, Julian, & Clarke, 2000), sleep disorders (Smolensky, Di Milia, Ohayon, & Philip, 2011), and brain injury (Lew et al., 2005; Schultheis et al., 2006).

Simulator validity. For driving simulators to be a useful tool, they must be a valid measure of driving behaviour; that is, performance in the simulator should accurately reflect behaviour and performance in real, on-road driving. The gold standard for validation of a simulator is to compare measures of performance on a simulator with identical measures of performance on-road, where conditions in the simulator match those in the real environment (Shechtman, 2010). However, the time, cost, specialised equipment and expertise required for this method mean that it can be impractical in many cases. Fortunately, there are other methods that can be used to establish the validity of a simulator.

In general, driving simulators as a measure of on-road driving ability and driving safety have demonstrated acceptable validity across several modes of validation (Mullen, Charlton, Devlin, & Bédard, 2011; Shechtman, 2010). For example, driving simulator performance predicted at-fault or partially at-fault crashes for older drivers in the five years following assessment (Hoffman & McDowd, 2010) and, for learner drivers, performance on a driving simulator predicted performance on an on-road assessment six months later (de Winter et al., 2009). Discriminant validity has been demonstrated by statistically significant differences in the performance of non-drivers, novice drivers, and experienced drivers both on a simulator and during on-road driving (Mayhew et al., 2011) Measures of overall performance, when compared between simulator and on-road assessment, display concurrent validity across all age groups from young adults to the elderly (Engström, Johansson, & Östlund, 2005; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011). Specific aspects of driving are also related for simulated driving and on-road driving; for example, Shechtman, Classen, Awadzi, and Mann (2009) demonstrated relative validity for types of driving errors made; and Kaptein, Theeuwes, and van der Horst (1996) showed absolute validity for route choice behaviour and relative validity for speed and lateral control. Lee, Lee, and Cameron (2003) investigated the validity of their simulator by looking at the correlation between age and visual attention task performance in the simulator; they found that older age was associated with poorer visual attention as measured by the simulator via a secondary visual attention RT task, and Bédard, Parkkari, Weaver, Riendeau, and Dahlquist (2010) showed that simulated driving performance was related to cognitive ability, with measures of visual attention being related to simulator-recorded errors.

These results suggest that, in general, driving simulators are a useful tool for measuring and assessing driving performance. However, it must be noted that simulator validity is highly dependent on the specific simulator, task, and population under consideration (Kaptein et al., 1996; Mullen et al., 2011; Shechtman, 2010). The usefulness of the simulator also depends on its face validity, or acceptability to the user. A simulator with high usability enables accuracy of task performance, enjoyment of the experience, and acceptance of the technology (Schultheis, Rebimbas, Mourant, & Millis, 2007). It is important that users take the task seriously and perceive the tasks to be an accurate reflection of their ability. These issues will be further investigated throughout the thesis.

Simulator Sickness (SS). SS is a well-documented side effect of using a wide range of simulators and virtual reality technology (Brooks et al., 2010; Classen et al., 2011; D. M. Johnson, 2005; Kennedy, Lane, Berbaum, & Lilienthal, 1993; McCauley, 1984; Stoner, Fisher, & Mollenhauer Jr, 2011; Trick & Caird, 2011). Overall estimated prevalence of SS varies greatly: for example McCauley (1984) reported rates of $10-84\%$, D. M. Johnson (2005) reported rates of $0-90\%$. Of 3,691 trials on a flight simulator, 50% of all users experienced some SS (Kennedy et al., 1993). Experience of SS is related to high rates of participant dropout in driving simulator studies; Trick and Caird (2011) reported estimated dropout rates of between 35% and 75% from various institutions conducting driving simulation research with older drivers, with an average of around 40% attrition. This high dropout rate is a concern for users of driving simulators, but also poses an ethical challenge when seeking to recruit research participants due to SS being considered as a risk for potential harm, although minimal, to participants (Brooks et al., 2010).

Symptoms of SS are similar to those of motion sickness and may include general discomfort, fatigue, headache, eyestrain, difficulty focussing, increased salivation, sweating, nausea, difficulty concentrating, feelings of fullness or pressure in the head, blurred vision, dizziness, vertigo, stomach awareness, and burping (Kennedy et al., 1993).

SS is usually measured through specialised self-report questionnaires, such as the acfSSQ (SSQ; Kennedy et al., 1993). The Simulator Sickness Questionnaire (SSQ) has been called the "gold standard" for measuring SS (D. M. Johnson, 2005). On the SSQ, participants respond on a 4-point scale the extent to which they are experiencing each of 16 symptoms. The 16 symptoms form three factors: oculomotor symptoms (e.g. eyestrain), disorientation symptoms (e.g. dizziness), and nausea symptoms (e.g. nausea, stomach awareness; Kennedy et al., 1993).

A short form of the SSQ, the Mini-SSQ, has been tested (Mourant, Rengarajan, Cox, Lin, & Jaeger, 2007). This version was developed to avoid delays involved in repeated administration, and includes only six symptoms: general discomfort, headache, blurred vision, sweating, feeling faint, and stomach discomfort. As with the Kennedy SSQ, participants register on a four-point scale of symptom severity. Mourant et al. (2007) reported that the Mini-SSQ was sensitive to changes in driving environment, such as increased scene complexity and increased task demands.

Factors contributing to SS can be located within three categories: Factors related to the individual, factors related to the simulator, and factors related to the simulated task (Cassavaugh, Domeyer, & Backs, 2011; Kolasinski, 1995). Of these, the simulator and task specifications can be controlled to an extent, for example by using a motion base simulator which replicates the pitch and roll movements of a real car (Stoner et al., 2011), using shorter scenarios (Cassavaugh et al., 2011), avoiding turns (Mourant et al., 2007; Stoner et al., 2011), and reducing the field of view (D. M. Johnson, 2005; Kolasinski, 1995). Pertinent individual risk factors for SS include age, gender, health, prior experience with simulators, and prior experience of motion sickness (D. M. Johnson, 2005; Stoner et al., 2011). Issues relating to SS will be further investigated throughout the thesis.

Conclusion

The foregoing review has argued that driving is a complex task requiring a combination of physical, sensory, and cognitive abilities. The literature has demonstrated associations between functional performance measures and driving outcomes, including crash risk and on-road driving performance. The most consistent results have come from cognitive function measures, particularly those reflecting visual attention, speed of processing, and visuospatial abilities. Results relating to sensory (visual) function measures and physical function measures have been less consistent. When drawing these conclusions it should be noted that the quality of studies has varied; however, despite small sample sizes and a focus on clinical groups, the overall quality of the work has generally been good. Taken together, the conclusion presented here represents a good account of the literature at this point in time.

Licensing procedures in Australia and around the world are moving away from mandatory age-based testing for older drivers. The focus is now on identifying drivers who are displaying signs of impairment and are at a heightened risk for crashes. Such drivers need to be appropriately assessed using valid and accurate screening measures. Tests of cognitive abilities appear to be most relevant for this purpose. Therefore, the main focus of this thesis was on the relationship between tests of cognitive abilities and driving performance in older adults. The specific cognitive tests of interest included the UFOV™, which is prevalent in the driving literature and is notable for its consistent relationship with a broad range of driving outcomes; and Inspection Time (IT) and Proficiency of Peripheral Visual Processing (ProPerVis) Crowding, assessments of processing speed and crowding across the visual field, respectively, which have not previously been investigated in relation to driving outcomes but have the potential to be used as screening tests. These measures will be discussed in the next Chapter.

The foregoing review also described the relationship between physical function and driving performance. Assessments of physical function have generally shown weak or inconsistent relationships with driving outcomes. However, physical activity and fitness may contribute to driving safety through an association with cognitive functions relevant for driving; higher self-reported physical activity has been associated with better cognitive performance, and physical activity interventions have been reported to improve cognitive performance. Few studies have directly investigated the link between physical activity and driving performance, but recent research has suggested that physical activity and exercise interventions may be associated with improved driving outcomes. These results suggest that physical activity interventions may be an effective strategy for improving or maintaining driving ability in older drivers. Therefore, this thesis also investigated the relationship between self-reported physical activity and driving performance in older adults.

Chapter 2.
Method and aims

Materials

This section provides an overview of the main materials to be used in the thesis and a statement of the aims of the thesis.

Cognitive measures

Useful Field of View Test™ (UFOV™). The UFOV™ is a computerised test of visual attention and processing speed involving detection and localisation of briefly presented targets throughout the visual field (Ball & Owsley, 1993). There are three subtests: processing speed, divided attention, and selective attention. Two stimuli are used; a silhouette of a car and a truck, sized 2×1.5 cm. In the processing speed subtest (Subtest 1), one of these stimuli is briefly presented within a central fixation box of 3×3 cm at varying exposure durations across trials in central vision, and the observer indicates which stimulus they saw by clicking the appropriate symbol on the screen. For the divided attention subtest (Subtest 2), one of the two stimuli appears briefly in central vision as before, and the car appears simultaneously in the periphery at one of the eight cardinal or intercardinal points, 12.5 cm from the centre of the display. Participants indicate which object was presented in central vision, and the location of the car in the periphery. An example trial from Subtest 2 (divided attention) is shown in Figure 2.1. The task for the selective attention subtest (Subtest 3) is the same but with the addition of 47 distractors (triangles of the same size and luminance as the target stimuli) distributed evenly throughout the visual field. For each subtest, the score is exposure time (ms) for which 75% of responses are correct. UFOV™ total score is calculated by summing the score from each subtest.

The UFOV™ was originally developed to assess everyday visual difficulties encountered by older adults and has since been interpreted as assessing visual attention and

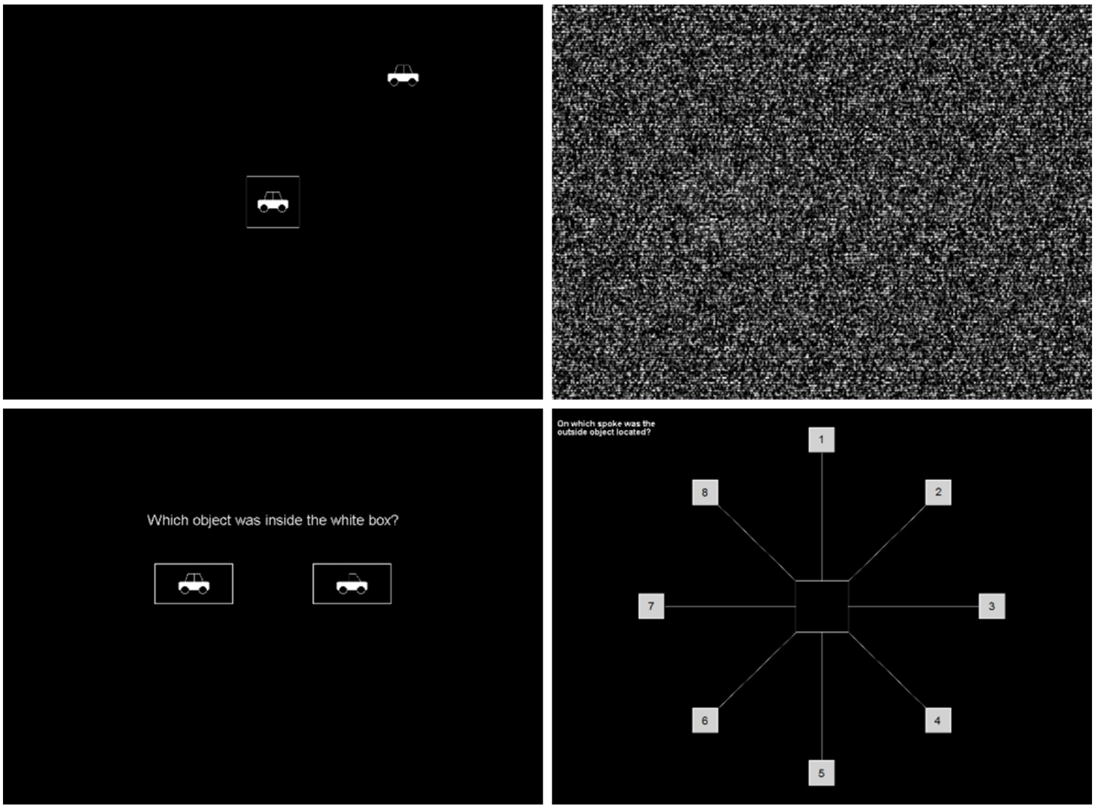

Figure 2.1: Example trial from UFOV™ Subtest 2 (Divided Attention). First, either a car or a truck appeared briefly inside a white box in central vision. At the same time, another car appeared at one of eight intercardinal points in peripheral vision. The presented stimuli were then masked with a full-screen random noise pattern. The viewer then responded to the question "Which object was inside the white box?" by using the mouse to click on either the car or the truck. Finally, the viewer responded to the question "On which spoke was the outside object located" by using the mouse to click on one of the eight numbered boxes corresponding to the eight possible intercardinal presentation points. In this example, the car was presented inside the white box, and the outside object was located at point two.

processing speed (Ball, Beard, Roenker, Miller, & Griggs, 1988) (Ball & Owsley, 1993; Ball, Roenker, & Bruni, 1990; Owsley, 1994). The UFOV™ involves detection, localisation, and identification of stimuli located throughout the visual field and comprises three subtests of increasing difficulty that require identification of a central stimulus and localisation of a peripheral stimulus under different conditions (J. D. Edwards, Vance, et al., 2005). All subtests involve visual processing under limited time, justifying the conclusion that UFOV™ measures "an individual's speed of processing across increasingly complex visual displays" (J. D. Edwards, Vance, et al., 2005, p. 530).

Regarding driving performance, UFOV™ performance has been shown to predict retrospective and prospective crash involvement, on-road driving performance, and driving simulator performance in older adults (Clay et al., 2005; Gentzler & Smither, 2012; Mathias & Lucas, 2009). UFOV™ Subtest 2 has been found to be particularly sensitive to driving outcomes. Subtest 2, designed to assess processing speed for a divided attention task, has been reported to be highly correlated with the Digit Symbol Substitution task which assesses visual processing speed (J. D. Edwards et al., 2006). Of the three subtests, Subtest 2 has been most related to crashes (Ball et al., 2006; Owsley et al., 1998; Oxley, Charlton, Koppel, Scully, & Fildes, 2005), on-road driving (Bowers et al., 2013) and simulated driving performance (Molnar et al., 2007). Subtest 2 has therefore often been considered on its own for purposes of brevity (Ball et al., 2006). The pattern of results reported for the UFOV™ suggest that it is a valid index of driving safety in older adults; however, it should also be recognised that there may be studies with non-statistically significant findings that have not been published.

Research has indicated that UFOV™ performance can be improved with training, and that these training improvements were associated with positive functional and driving outcomes. Of particular note is the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study, which was conducted with 2832 participants aged 65 years or over (Jobe et al., 2001). The study investigated the effects of cognitive training interventions on measures of daily functioning. Participants were assigned to one of three cognitive intervention groups (memory, reasoning, or speed of processing) or a no-contact control group. The speed of processing training task was based on the UFOV™, but features variations in stimuli and task difficulty not found in the UFOV™. Results from the ACTIVE study showed that speed of processing training was associated with significant improvements in UFOV™ reported for up to five years after initial intervention (Ball et al., 2002; Ball, Ross, Roth, & Edwards, 2013).

Speed of processing training has also been reported to have positive effects on driving outcomes. Ball, Edwards, Ross, and McGwin Jr (2010) reported that drivers in the ACTIVE speed of processing intervention group (and the reasoning training group) had a 50% lower at-fault crash risk than drivers in the control group, over a 6-year period after training commencement. The speed of processing intervention was also associated with improvements in self-reported driving mobility five years after initial training, with the results most evident in those who had lower baseline processing speed (Ross et al., 2016).

The association between speed of processing training and improved UFOV™ performance has been consistently reported in several other studies (J. D. Edwards et al., 2015, 2002; J. D. Edwards, Wadley, et al., 2005; Roenker et al., 2003; Vance et al., 2007). However, training benefits appear to be specific to speed of processing and do not transfer to other domains of cognitive ability. Speed of processing training is reportedly associated with improvements in Timed Instrumental Activities of Daily Living (Timed IADL; J. D. Edwards et al., 2002; J. D. Edwards, Wadley, et al., 2005). The Timed IADL task measures the time taken to complete tasks resembling everyday activities, such as finding a telephone number, finding and reading medicine instructions, and finding and counting correct change. Studies have also shown that speed of processing training was associated with improvements in certain measures of simulated and on-road driving performance (Roenker et al., 2003) and with self-reported driving mobility (Edwards et al., 2009).

The UFOV™ was originally developed with specialised testing equipment. Later, a commercial PC version of the test was developed, with the option of mouse input or touchscreen input. The Mouse PC version was used for this thesis. A study on the reliability of the UFOV™ reported that test-retest reliability of the Mouse PC version of the UFOV™ over 10 days was 0.68 for Subtest 1, 0.81 for Subtest 2, 0.85 for Subtest 3, and 0.88 for UFOV™ Total Score. Scores on the PC Mouse version corresponded well with scores on the PC Touchscreen version and the Original Version. Between the PC Mouse version and the original version, the validity coefficients were reportedly 0.49 for Subtest 1, 0.74 for Subtest 2, 0.76 for Subtest 3, and 0.72 for UFOV™ Total Score. Between the PC Mouse version and the PCS touchscreen version, the validity coefficients were reportedly 0.44 for Subtest 1, 0.89 for Subtest 2, 0.99 for Subtest 3, and 0.92 for UFOV™ Total Score (J. D. Edwards, Vance, et al., 2005). These results suggest excellent validity and reliability for the PC version of the UFOV™, particularly Subtests 2 and 3.

Psychometric assessment of the UFOV™ has been limited, but it appears to be a valid measure of processing speed and aspects of visual attention. UFOV™ is reportedly significantly correlated with measures of processing speed and visual attention, including WAIS-R digit symbol substitution (J. D. Edwards et al., 2006) and the Road Sign Test (J. D. Edwards, Vance, et al., 2005). Additionally, UFOV™ and the Digit Symbol Substitution test, a well-known neuropsychological assessment with a speed of processing component, showed a similar trajectory of decline accompanying older age (Lunsman et al., 2008).

Inspection Time (IT). IT is a computerised assessment of speed of information processing (Burns & Nettelbeck, 2003; Nettelbeck, 2001) Two high-contrast lines, one markedly shorter than the other, appear for limited variable time as a target on a computer screen. Participants indicate whether the shorter line is located left or right of a focal point. Time available for processing is limited by a backward masking procedure and reduced or extended using an adaptive staircase algorithm according to response accuracy. Targets are preceded by a warning cue ("+" in the centre of the screen, 370ms) and are immediately followed by a mask figure shaped like two lightning bolts.

The IT task in the studies in this thesis was administered as follows. First, participants completed three sets of 10 practice trials with decreasing target presentation time for each set (835ms, 420ms, and 250ms) that required answering 10/10 items correctly for the first and second set, and 9/10 items for the third set. After successful completion of the practice trials, the test trials were administered according to the adaptive staircase algorithm. IT was measured in ms as the duration between target onset and mask onset at which the viewer achieved 79% response accuracy. Lower scores indicated faster speed of visual processing. A representation of the IT task is shown in Figure 2.2. Test-retest reliability for adults has been reported as .81 (Grudnik & Kranzler, 2001).

IT was included rather than a measure of simple or complex RT because it measures speed of processing and accuracy under time constraints, but is not dependent on speed of motor performance (Nettelbeck, 2001). IT is notable for its moderately high cor-

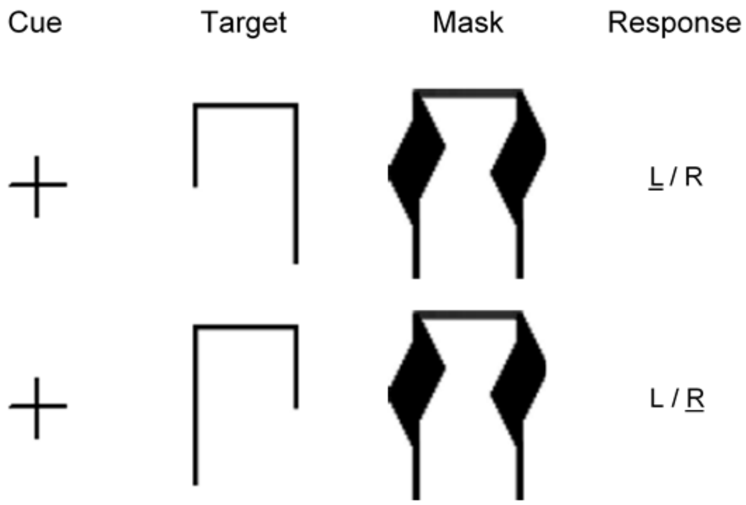

Figure 2.2: Representation of the Inspection Time task. First, the + cue appeared briefly in central vision. Next, the target figure appeared briefly in central vision for a duration determined by the adaptive staircase algorithm, followed by the mask figure. The viewer then responded whether the shorter line was located on the left or the right of the target figure using the buttons on the mouse (i.e. the left mouse button was clicked if the the left line was shorter, the right mouse button was clicked if the right line was shorter).

relation with IQ (around 25% shared variance), and its relationship with other cognitive abilities including fluid reasoning and short-term memory (Grudnik & Kranzler, 2001). Furthermore, there is evidence that IT may be useful as a biomarker for future general cognitive decline. To be considered a biomarker, an indicator must be able to predict future changes in cognitive and functional outcomes better than chronological age; additionally, short term rate of change in a biomarker should be able to predict future declines in important outcomes (Baker & Sprott, 1988). IT is argued to meet these conditions; for example, although IT generally lengthens with normal aging, it has also been associated with poorer future performance on a range of outcomes, including future cognitive ability and performance of everyday tasks (Deary, Johnson, & Starr, 2010; Gregory, Nettelbeck, Howard, & Wilson, 2008; Gregory, Callaghan, Nettelbeck, & Wilson, 2009). Deary et al. (2010) compared a number of processing speed measures as potential biomarkers and concluded that IT was the most promising, as IT performance in old age was least dependent on cognitive abilities measured earlier in life. Although IT has not previously been investigated in relation to driving outcomes for older adults, these results and IT's argued status as a biomarker suggest that IT may be useful as a screening measure to detect drivers who are currently experiencing slowed processing speed, and who may be at risk for future cognitive decline and increased driving problems in the future.

ProPerVis. ProPerVis is a computerised assessment of "Proficiency of Peripheral Visual Processing" and contains two subtests: a crowding subtest and an IT subtest. Only the crowding subtest was included in this thesis. The crowding subtest assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns, Kremer, & Baldock, 2005). The stimuli are a four-square parent figure and six figures derived from it, resembling stylised characters M, E, W, 3, 5, 2. On each trial, one of the six figures is presented, flanked on either side by the parent figure. The target and flankers appear randomly in one of five lateral positions on the screen. The positions are "central", "parafoveal" (left and right), and "peripheral" (left and right). Participants attempt to identify which of the six figures was presented. The outcome was total errors made across the five positions from 40 trials; lower scores indicated better performance (i.e. fewer errors). The ProPerVis Crowding Task and Stimuli are shown in Figure 2.3.

Visual crowding is a phenomenon whereby visual interference occurs when a visually off-centre primary target is flanked by similar materials, resulting in difficulty identifying a crowded stimulus (Bouma, 1970; Levi, 2008; Pelli, Palomares, & Majaj, 2004). Crowding has generally been considered a visual phenomenon but recent work suggests it may also involve limits to attentional resolution (He, Cavanagh, & Intriligator, 1996; Levi, 2008). Crowding influences the successful execution of everyday tasks, including reading, interacting with the environment, and driving, because the visual scene is often cluttered (Whitney & Levi, 2011). In a cluttered scene, crowding does not affect the detection of an object; rather, it affects the identification of an object, making it appear indistinct or jumbled with surrounding objects (Levi, 2008). The effect of crowding increases with age, such that older adults find it much more difficult to extract information from a crowded visual field, especially when the stimuli of interest are presented in peripheral vision (Burns et al., 2005; Corlett & White, 2002; Werner, 2008). In regards to change blindness in driving scenes, it has been found that drivers are able to detect changes to relevant targets (driving-related items such as road signs, hazards, and pedestrians) more quickly than irrelevant targets (Galpin, Underwood, & Crundall, 2009).

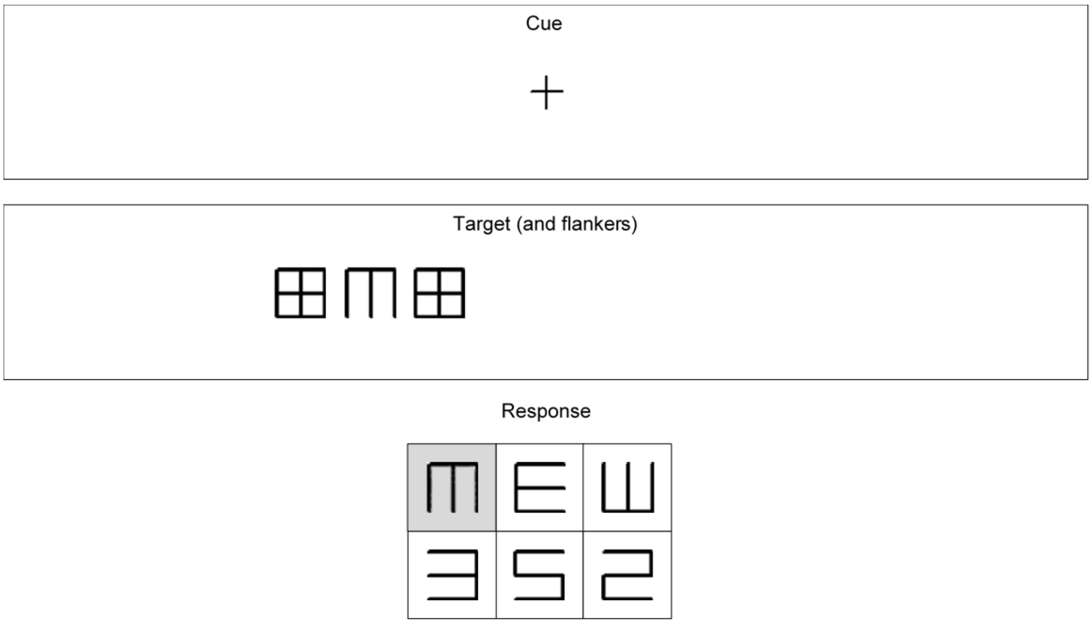

Figure 2.3: Example trial from ProPerVis Crowding. First, the + cue appeared briefly in central vision. Then, the target figure appeared at one of five lateral positions on the screen, flanked on both sides by the four-square parent figure. The participant then responded by using the mouse to click on the presented figure. In this example trial, the “M” figure was presented in the left parafoveal position.

ProPerVis was developed as a potential alternative to the UFOV™ as a driver screening tool. Cognitive processes contributing to UFOV™ performance are unclear, although it was thought to reflect aspects of processing speed and peripheral vision processing, as reflected in the Crowding and IT subtests of ProPerVis. Results from the development of ProPerVis indicated that ProPerVis Crowding, ProPerVis IT, and UFOV™ Subtest 3 (Selective Attention) were highly correlated, although correlations between the ProPerVis measures and UFOV™ Subtests 1 and 2 were low. Test-retest reliability over one week for ProPerVis Crowding was reportedly .40 for the central position, .89 for the parafoveal position, and .89 for the peripheral position (Burns et al., 2005).

DriverScan. DriverScan is a change detection task designed to assess visual attention skills needed for safe driving (Hoffman, Yang, Bovaird, & Embretson, 2006) The test uses the change detection "flicker" paradigm described by Rensink, O'Regan, and Clark (1997). Real world images of driving scenes (A) presented on a computer screen are alternated with a blank screen and an altered version of the image (A') to which a small change has been made, creating a "blinking" effect. Changes include deletion of objects (e.g. cars), colour/lettering changes (e.g. road signs), and signal changes (e.g. traffic lights, brake lights). The pattern of presentation is A, blank, A, blank, A', blank, A', blank, A, blank, A, blank... etc., with the images (A, A') being presented for 280 ms and the blank screen being presented for 80ms.

The DriverScan task was administered as follows. First, there were four practice trials displaying different types of changes. If participants did not answer correctly or were not confident, the practice trials were repeated. This was followed by 25 test trials where each trial was presented for 45 seconds or until a response was recorded. Participants viewed the images and attempted to detect the change between A and A', responding by clicking the mouse as soon as they detected a change and then verbally describing the change to the experimenter. The response time for each trial was recorded, and these times were used to generate an ability estimate for each participant using a constrained Graded Response Model (an application of Item Response Theory for polytomous data; Hoffman et al., 2006; Samejima, 1970). Responses were classed as immediate response (< 8 s), delayed response ($8 - 45$s), no response, or incorrect response. Higher ability scores represented better performance on the test (i.e. fast, correct responses). Viewing distance was approximately 60cm.

During the development of the instrument, items and responses were concluded to be sufficiently unidimensional and reliable according to principles of Item Response Theory (Hoffman et al., 2006). In a validation study of DriverScan, a sample of older adults completed DriverScan, UFOV™, and a driving simulation assessment (Hoffman et al., 2005). DriverScan scores correlated moderately strongly with the UFOV™ Subtests 2 and 3, providing evidence that DriverScan reflects aspects of visual attention. Both DriverScan and UFOV™ significantly predicted driving simulator performance.

The DriverScan task involves detecting changes in photographs of real world driving scenes and on this basis it appears to have ecological validity as a test for real-world driving skills, where drivers are required to be constantly vigilant for important changes in the environment like changes to traffic lights, position and velocity of other vehicles on the road, and the sudden presence of potential hazards. Along with sustained vigilance, successful detection of change is also argued to involve executive functioning, working memory, and processing speed (Hoffman et al., 2006; Pringle, Irwin, Kramer, & Atchley, 2001; Pringle, Kramer, & Irwin, 2004; Rizzo et al., 2009).

Working memory. A sentence span task was used to assess working memory. The task was administered according to specifications described by Lewandowsky, Oberauer, Yang, and Ecker (2010). Participants were presented with a series of sentences and to-be-remembered letters. Each trial consisted of between 4 and 8 sentence/letter pairs. Participants were required to answer TRUE or FALSE to each question (e.g. "All trees are plants"); after answering, a single letter was briefly presented on screen. At the end of each trial, participants were required to enter the remembered letters in order of presentation. The outcome was the overall proportion of correctly remembered letters. The working memory task is reported to have good reliability, and high internal and external validity (Lewandowsky et al., 2010).

Mini Mental State Examination (MMSE). The MMSE is a short test of global cognitive function used to estimate the extent of any cognitive impairment (Folstein, Folstein, & McHugh, 1975). Participants answer questions assessing short-term memory, orientation in space and time, constructional ability, executive functioning, and ability to follow instructions. The MMSE was administered according to standardised instructions and scoring. The score was reported out of a maximum of 30, with lower scores indicating more impairment. A score of 23 or less is generally taken to indicate the presence of cognitive impairment, and impairment status can be interpreted as follows; 24 − 30, no cognitive impairment 18 − 23, mild cognitive impairment, 0 − 17, severe cognitive impairment (Tombaugh & McIntyre, 1992). The MMSE has been included in this thesis as a screening measure for study inclusion; participants were required to be free from cognitive impairment. A minimum score of 24 out of 30 on the MMSE was required for participation in the studies reported in this thesis[1]. The MMSE also provided an index of global cognitive function, which was assessed in relation to driving performance. The MMSE is widely used as a quick and convenient measure of global cognition; given time constraints on participants, the MMSE was selected for this purpose.

In a review focussing on older drivers with dementia, it was found that MMSE was moderately correlated with on-road and simulated driving performance (Adler et al., 2005). On the other hand, some studies have found no association between MMSE score and driving outcomes in both impaired (Molnar et al., 2006) and non-impaired samples (Crizzle et al., 2012; Joseph et al., 2014). It is likely that the MMSE is not sensitive enough as a predictor of driving outcomes in non-impaired older drivers (Crizzle et al., 2012).

The test-retest reliability of the MMSE over periods up to two months has been reported to be between .66 and .99 for periods between 1 day and 2 months. Internal consistency reliability has been reported to be between .54 and .96. MMSE scores correlate highly with scores from other cognitive screening tests, and other tests of intelligence, memory, and activities of daily living (Tombaugh & McIntyre, 1992).

Physical activity

Physical Activity Scale for the Elderly (PASE). The PASE (Washburn, Smith, Jette, & Janney, 1993) was developed to measure the physical activity of older people. The self-report questionnaire, which takes around five minutes to complete, asks older adults to recall their physical activity over the past seven days, in categories of

[1] It should be noted that no participants were in fact screened out due to this requirement

leisure time activity, household activity, and work-related activity. Respondents report if they participated in the listed activities over the past seven days. For leisure time activities and work, participants report how many hours per day they participated in the activity. For household activities, participants respond either “yes” or “no” to having engaged in that activity over the past seven days.

For this thesis, the PASE was preferred to a functional assessment of physical capabilities given constraints on time available to participants. The PASE was completed via computer. Items were scored according to validated item weights, multiplied by the number of hours per week spent engaging in the activity. The item weights were developed in the original sample and were dependent on their contribution to an overall physical activity score extracted from motion-sensor counts, physical activity diary, and global activity self-assessment (Washburn et al., 1993). PASE scores in this original sample were significantly correlated with grip strength, static balance, leg strength, resting heart rate, age, perceived health status, and overall Sickness Impact profile (Washburn et al., 1993). Further studies have validated the PASE using the Doubly-Labeled Water Method, a “gold standard” method for measuring energy expenditure (Schuit, Schouten, Westerterp, & Saris, 1997), activity monitoring devices (Dinger, Oman, Taylor, Vesely, & Able, 2004; Harada, Chiu, King, & Stewart, 2001; Washburn & Ficker, 1999), and associations with age, health, and physical functioning measures including self-reported health and functioning, lower body functioning (lower body strength, balance, and walking), 6-minute walk score, peak oxygen uptake, systolic blood pressure, balance, and chronic health conditions (Harada et al., 2001; Washburn, McAuley, Katula, Mihalko, & Boileau, 1999). PASE has also been associated with executive functioning measures including the Clock Drawing Test, Animal Naming Test, and Trail Making Test (Eggermont, Milberg, Lipsitz, Scherder, & Leveille, 2009).

Test-retest reliability of PASE over $3-7$ weeks in the original sample was reportedly $r = .75$ (Washburn et al., 1993), and in a different sample the intraclass correlation coefficient over 3 days was $r = .91$ (Dinger et al., 2004). PASE has been reported to be significantly correlated with scores on the Community Health Activities Model Program for Seniors (CHAMPS) questionnaire (Stewart et al., 2001), $r = .58$; CHAMPS is a more detailed activity questionnaire for older adults.

Visual function measures

Pelli-Robson Contrast Sensitivity Chart. The Pelli-Robson Contrast Sensitivity Chart (Pelli, Robson, & Wilkins, 1988) consists of eight rows of six letters, arranged into groups of three letters. The contrast of the letters reduces for each subsequent triplet. For the studies reported in this thesis, participants viewed the chart binocularly at eye level from a distance of approximately one metre while wearing normal vision correction. Participants were instructed to read the letters on the chart as far down as they could and were encouraged to guess if they were uncertain. The outcome was log contrast sensitivity, which was recorded according to the faintest triplet of letters for which the participant correctly identified at least two of the three letters. Possible scores ranged from 0 to 2.25, with higher scores indicating better contrast sensitivity.

The coefficient of repeatability for Pelli-Robson Contrast Sensitivity Chart scores over two weeks has been reported at 0.15 log units (Elliott, Sanderson, & Conkey, 1990). The coefficient of repeatability represents the 95% confidence interval for the difference between two sets of results.

Freiburg Visual Acuity Test. The Freiburg Visual Acuity Test (Bach, 1996) is a computerised test of visual acuity. The test is quick (approximately two minutes) and easy to administer. For this thesis, the Freiburg Visual Acuity Test was administered as follows. Participants viewed the computer monitor from a set distance and indicated the orientation of a "Tumbling E" optotype (Figure 2.4) using the arrows on the computer keyboard. The size of the optotype varied on each trial depending on the current estimated threshold of the participant, calculated via the best Parameter Estimate by Sequential Testing procedure (Bach, 1996). There were 30 trials, with every sixth trial being an "easy" trial where the optotype size was significantly larger than the current estimated threshold. Acuity was recorded as the logarithm of the Minimum Angle of Resolution (logMAR), with lower scores representing better visual acuity.

Scores on the Freiburg Visual Acuity Test have been reported to correspond within one percent with scores on a conventional forced-choice chart visual acuity test (the DIN 58220; Bach, 1996), and to correspond well with other validated measures of visual acuity (Schulze-Bonsel, Feltgen, Burau, Hansen, & Bach, 2006). The Freiburg Visual Acuity Test has been reported to have excellent test-retest reliability (Bach, 2007; Schulze-Bonsel et al., 2006).

Figure 2.4: Freiburg Visual Acuity Test Tumbling E Optotype

Aims

The psychometric properties of theUseful Field of View Test™ test are unclear. Given the prolificacy of this test in the driving literature, an initial aim of this thesis was to determine what each of the three UFOV™ subtests measures, in terms of lower-level cognitive and sensory factors.

A driving simulator was used to assess the driving performance of older drivers. Driving simulator validity is highly dependent on the specific simulator, driving tasks, and population. One aim was therefore to establish the validity, reliability, and usability (including rates of simulator sickness and user acceptance) of the simulator and tasks used in the thesis.

The main aim was to investigate the contribution of cognitive, visual, and physical activity factors to simulated driving task performance in healthy older adults. The UFOV™ was considered, as were two cognitive abilities test that have not previously been assessed in relation to driving performance: Inspection Time and Crowding Across the Visual Field. The contribution of self-reported physical activity to driving performance was also investigated, along with visual function measures. The intention was to determine the best combination of these cognitive, visual, and physical function variables for predicting driving performance; and to consider the potential application of these tests as screening assessments for determining fitness-to-drive in older adults.

Chapter 3.
Exegesis

The aim of the thesis was to identify functional predictors of driving performance in older adults, with a focus on cognitive predictors. Several potential predictors of particular interest were identified for inclusion in the thesis, and have been described in the foregoing review in Chapter 1. Two of these predictors were Inspection Time (IT; a measure of visual processing speed and a potential biomarker for cognitive aging) and ProPerVis Crowding (a measure of crowding across the visual field). These two measures have not previously been investigated in relation to driving performance. Other predictors of interest were the Useful Field of View Test™ (UFOV™), working memory, change detection, visual acuity, contrast sensitivity, and physical activity. While the main focus of the thesis was on the cognitive predictors, the inclusion of visual and physical function measures was also considered to be relevant. The foregoing review indicated that, despite the literature suggesting inconsistent links with driving outcomes, these constructs and measures of them could be considered to be relevant for the driving task and different versions have been assessed by licensing authorities and medical practitioners to determine fitness to drive.

Also of interest were the characteristics of the UFOV™ and its relationship to driving ability. The UFOV™ has been widely used as a screening test for fitness-to-drive in older adults and has shown consistent relationships to various driving outcomes. However, the psychometric properties of the UFOV™ remain unclear.

During the course of the thesis, the focus also widened to include methodological issues relating to use of driving simulators to assess driving performance in older drivers. These issues included the fact that some participants experience simulator sickness when performing simulator tasks; and assessment of the validity, reliability, and usability of the driving simulator and the tasks designed to test driving competencies.

Five studies have contributed to the thesis. These studies form a program of sequential investigations, whereby there has been an attempt to address questions generated during the course of the research. As submitted, this thesis is consistent with the option permitted under current rules of submission of a thesis by publication. As a consequence of preparing the papers for publication, there are small differences in the formatting of the papers as included here. Moreover, some information, particularly relating to methodology and materials, has necessarily been duplicated in each of the papers. The first three studies have been published during candidature, and the publication details can be found at the beginning of each relevant chapter. The final two studies are as yet unpublished, but have been prepared in the format of a manuscript suitable for submission to a journal for publication.

The following sections contain an overview of the rationale, aim, and methodology for each study. The methodology and test selection for each study evolved throughout the program of research and was informed by the results of each previous study. Briefly, the first study aimed to identify the contribution of a set of cognitive variables to performance on the UFOV™, in order to determine cognitive factors that contribute to UFOV™ performance. The second and third studies investigated methodological issues related to driving simulator assessment for older drivers; these two studies focussed on factors associated with simulator sickness and on establishing the reliability, validity, and usability of the simulator. The fourth study further investigated these methodological concerns including both younger and older drivers. The final study investigated functional predictors (cognitive, visual, and physical activity) of simulated driving performance in healthy older adults.

Study 1

The first study focussed on the UFOV™. The UFOV™ has frequently been used in driving research and has been consistently linked to a wide range of driving outcomes. Despite its frequent use and established relationships with driving outcomes, the psychometric properties of the UFOV™ are not well defined, although it is thought to measure visual processing speed and aspects of visual attention. Approximately half of the data collection for this study was completed by me in 2012 as part of my Honours research project. The remaining data collection occurred after the commencement of my candidature, and all analyses presented in Study 1 are new and were completed during my candidature.

The aim of the study was to investigate whether performance on the three UFOV™ subtests could be explained in terms of combinations of other functional screening tests of cognitive abilities theoretically linked to driving performance in older adults. The study aimed to explore underlying psychometric properties of the three UFOV™ subtests and the theoretical and practical implications for driving. A battery of six cognitive and visual predictor variables was included. The variables were selected to represent a range of cognitive and visual functions identified in the foregoing review as relevant for driving. The cognitive predictors were general mental functioning, assessed by the Mini Mental State Examination (MMSE); visual processing speed, assessed by IT; Crowding across the visual field, assessed by ProPerVis Crowding; and change detection, assessed by DriverScan. The visual function predictors were contrast sensitivity and visual acuity.

As described in Chapter 4, the results contributed to an improved understanding of the underlying cognitive factors that contribute to performance on the three UFOV™ subtests.

Study 2

Following Study 1, the contributions of the UFOV™ and other cognitive variables of interest to driving performance in older adults were investigated. The intention was to use a simulated driving assessment as the outcome measure. A custom-built, low cost driving simulator and two driving tasks were developed in-house for this purpose. The driving tasks were intended to assess the ability to react appropriately to unexpected driving hazards. Participants completed two simulated driving tasks, and a battery of cognitive and visual assessments including UFOV™, IT, ProPerVis Crowding, Working Memory, MMSE, visual acuity, and contrast sensitivity.

However, during the study it became evident that there were significant problems with the methodology. The main concern was the rate of dropout due to simulator sickness: 59% of participants were unable to complete the driving tasks, and this was known to be unacceptably high, given extant reports in the literature about this phenomenon. Furthermore, participants expressed concerns about the "realism" of the driving simulator, as a substitute for on-road driving, and about the relevance of the assessment tasks to on-road driving performance.

Based on this, and to inform future studies, the risk factors relating to simulator sickness in this sample were investigated. Survival Analysis methods were used to identify factors relating to the individual (e.g. gender, age) that were related to an increased risk of simulator sickness and dropout.

The study provided a discussion of these individual risk factors, and other factors inherent in the custom-made simulator that were likely to have contributed to the high simulator sickness rate. The results from Study 2 informed the methodology for the following studies.

Study 3

As a consequence of the difficulties encountered in Study 2, it was decided that a new approach was needed for the remaining studies. The main problems with the custom-built driving simulator were the high simulator sickness rate, which had not been anticipated, the physical setup of the simulator, and the relevance and realism of the scenarios. The experience with simulator sickness in Study 2 informed the practices for the studies that followed. For example, participants were more thoroughly screened prior to participating. The physical setup of the custom-designed simulator was noted as an issue during Study 2. The design included a small gaming wheel and gaming pedals, with participants seated on a standard desk chair. Participants indicated that the wheel felt too small and unrealistic, and they did not feel that the controls were representative of a real car. There were also problems noted with the programming of the vehicle handling; the steering was perceived to be much too sensitive, and braking and accelerating were too sudden. These issues led to many participants having difficulties adapting to the control of the vehicle, for example being unable to drive forward in a straight line at the appropriate speed, and being unable to turn corners. The issues with the controls and handling also contributed to sudden and jerky movement on the screens, which likely contributed to the high simulator sickness rates.

The custom-designed simulator ran two custom scenarios that were developed with the generous donation of time and resources by staff at Sydac Pty. Ltd. Sydac specialises in driver training simulation and the commercial application of simulation technologies[1]. However, the satisfactory development of the simulator and scenarios was a much more difficult process that initially anticipated. While staff at Sydac have had significant experience in train simulation and other areas of vehicle simulation, they have had only minor previous involvement with car simulators. The project was further constrained by limitations on the budget. The difficulties included implementing a realistic vehicle dynamics model (i.e. the handling of the vehicle); setting up automated scoring and recording of driver performance; and dealing with other unforeseen glitches and bugs in the program. Although these issues might have been solved with a higher budget and significantly more time invested in the project, it was ultimately decided that a ready-to-use driving simulator should be purchased for the remaining studies.

The aim was to purchase a simulator that appeared realistic and felt realistic to drive, could run scenarios that would be relevant for assessment of older drivers, and could be purchased within our budget. Research into the availability of such simulators indicated that many of the lower-cost simulators commonly used in driving research were outside of our budget; particularly those with customisable software and realistic controls (e.g. a full size steering wheel). We eventually decided on the SimWorx Driver Training Simulator - Light, running Carnetsoft software. This simulator was designed and built in Australia, thus reducing the shipping costs involved in purchasing from overseas. The SimWorx Simulator features a driver cab setup, with a realistic driver's seat, larger steering wheel, and realistic pedals and other controls. After trialling the simulator, it was evident that the control and handling of the vehicle was superior to that of the custom-designed simulator. Although no suitable simulators with customisable software were identified within the budget, the SimWorx simulator included a large range of automatically-scored scenarios that were deemed to be appropriate.

The first project with the new simulator investigated aspects of usability, test-retest reliability, validity (content, convergent, and face), and incidence of simulator sickness. Driving simulator validity is highly dependent on the specific equipment and scenarios used. Therefore, the aim of Study 3 was to investigate the rate of dropout due to simulator sickness and to determine the acceptability, reliability, and validity of the new

[1] see `http://www.sydac.com/en/` for more information

simulator using a variety of measures and methods. A small sample of older drivers was recruited for the study. The drivers completed four different scenarios on the new simulator and provided feedback about their experience. Task performance and rates of simulator sickness and dropout were also observed.

Results showed that dropout due to simulator sickness was 31%, which was substantially lower than the 59% experienced with the previous custom-built simulator, and was comparable to rates in the lower range reported in the literature for older drivers. Feedback from participants indicated that they were generally positive towards the simulator and perceived that the scenarios were an appropriate assessment of their driving ability. Overall, it was concluded that the simulator and tasks were suitable for use with older drivers.

Study 4

Given the promising results from Study 3, the validity and acceptability of the new simulator was further investigated in a larger sample of younger and older drivers. The data for the younger drivers was collected by Emmanuel Chalacas in 2015 and formed part of an Honours research project under my direction, supervised by Emeritus Professor Ted Nettelbeck. The younger and older drivers used the same driving simulator and completed the same driving assessments. All analyses reported in Study 4 are new and were conducted by me during my candidature. Aspects of the data for the older drivers have also been presented below as Study 5.

Study 4 aimed to build on the validation findings of Study 3 and included data to investigate differences between the younger and older drivers on simulated driving performance measures, simulator sickness rates, and on their feedback towards the simulator. A new sample of older adults were recruited and provided data for Study 4 and Study 5. The age group differences in performance in the two tasks were generally as expected and demonstrated discriminant validity for these tasks. It was found that older adults were more prone to simulator sickness and provided less positive feedback towards the simulator. However, feedback from both age groups was considered to be generally positive and, encouragingly, feedback was independent of task performance.

The results from Study 4 added to the results from Study 3 and suggested that the simulator and selected tasks were appropriate for assessment of older drivers. Results were discussed in relation to the validity and usability of the simulator for younger and older drivers. Studies 2 to 4 provided a discussion of methodological considerations around the use of driving simulators in older drivers, and informed the methodology for Study 5.

Study 5

The older adults from Study 4 provided data analysed for this study. The aim was to investigate the contribution of a range of functional measures (cognitive, visual, and physical activity) in relation to performance on the simulated driving tasks. Participants completed a battery of functional assessment measures (cognitive, visual, and physical activity) and two simulated driving tasks: a Brake Reaction Time task and a City Traffic Participation task, which involved following directions along a set route, navigating intersections and some hazards, and interacting with traffic.

Of particular interest were the cognitive measures IT and Crowding. In Study 1, these two tests contributed significantly to performance on UFOV™ Subtests 2 and 3. IT and Crowding have not previously been investigated in relation to driving performance.

The relationship between physical activity and driving performance was also of interest. Evidence (reviewed in the Introduction) suggested that physical activity and exercise may be related to cognitive function in older adults. An adequate level of physical function is also required to drive safely. Physical activity was measured using a self-report questionnaire, the Physical Activity Scale for the Elderly (PASE).

Performance on the cognitive, visual, and physical activity measures were used to predict performance on the two simulated driving tasks. The theoretical and practical implications of the results were discussed, including discussion of the cognitive processes contributing to driving performance, and how the results relate to current and proposed screening procedures for older drivers.

Chapter 4.

Study 1: Cognitive and visual predictors of UFOV performance in older adults

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2014). Cognitive and visual predictors of UFOV performance in older adults. *Accident Analysis and Prevention*, 70, 74-83. doi: 10.1016/j.aap.2014.03.011

Statement of Authorship

Title of paper	Cognitive and visual predictors of UFOV performance in older adults
Publication status	Published
Publication details	Matas, N. A., Nettelbeck, T., & Burns, N. R. (2014). Cognitive and visual predictors of UFOV performance in older adults. *Accident Analysis and Prevention*, 70, 74-83. doi: 10.1016/j.aap.2014.03.011

Principal Author

By signing the Statement of Authorship, I certify that this paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper

Name of principal author	Nicole Amy Matas		
Contribution to paper	Study concept and design, data collection, statistical analyses, writing and submitting manuscript, addressing referee comments		
Overall percentage	75%		
Signature		Date	13/9/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ted Nettelbeck		
Contribution to paper	Principal supervision, advice about development of concept and design of study, manuscript revision and approval		
Signature		Date	13. 9. 16

Name of Co-Author	Nick Burns		
Contribution to paper	Advice about development of concept and design of study, assistance with statistical analyses, manuscript revision and approval, ongoing		
Signature		Date	14/9/16

Abstract

Eighty two community dwelling older adults (52 females) aged $62-92$ years (mean = 75) completed a battery of cognitive and visual tests selected to assess functions relevant to driving performance. These were visual acuity, contrast sensitivity, general mental competence (Mini Mental State Examination, MMSE), processing speed (Inspection Time, IT), crowding across the visual field (Proficiency of Peripheral Visual Processing, ProPerVis) and change detection (DriverScan). These six tasks provided predictor variables for performance on the Useful Field of View Test™ (UFOV™), a well validated test of fitness to drive that includes subtests for (i) processing speed; (ii) divided attention; and (iii) selective attention. Relative importance regression analyses confirmed that UFOV™ is sensitive to attentional and speed processes but suggested that subtest (i) primarily reflects visual acuity and contrast sensitivity; subtest (ii) is better explained by change detection and processing speed; and subtest (iii) predominantly reflects crowding and contrast sensitivity. Unexpectedly, given no evidence of substantial cognitive decline, MMSE contributed significantly to performance on the more complex subtests (ii) and (iii).

Introduction

Older adults comprise a rapidly growing section of the population in Australia and throughout the world (Australian Bureau of Statistics [ABS], 2008; United Nations [UN], 2011). By 2035, one in four people is expected to be aged over 65, and by 2050 the number of people aged over 80 is expected to triple (OECD, 2001; UN, 2011). Driving will remain the preferred method of transport for most of these people (OECD, 2001). This is of potential concern because older drivers are a higher risk group for involvement in motor vehicle accidents. In Australia, people aged 70 years or older are the second most likely group to be involved in a fatal accident, second only to those aged $17 - 25$ (Bureau of Infrastructure, Transport and Regional Economics [BITRE], 2011). Additionally, the frailty of older adults means that they are more likely to be seriously injured or killed in the event of a crash (OECD, 2001; Li et al., 2003; Baldock & McLean, 2006).

Current licensing procedures for older drivers typically rely on medical and visual assessments, but the validity of existing procedures is questionable. For example, comparison of Australian jurisdictions found no reliable differences for crash data and serious accidents between jurisdictions with and without mandatory age-based fitness to drive testing (Langford, Fitzharris, Koppel, & Newstead, 2004; Langford, Fitzharris, Newstead, & Koppel, 2004). Moreover, the proportion of older drivers with self-reported cognitive and/or visual deficits appears to be comparable across jurisdictions, irrespective of whether mandatory aged-based testing applies or not (Ross, Browning, Luszcz, Mitchell, & Anstey, 2011). Thus, there is a need to develop procedures to better identify older drivers who may be at elevated risk for crash involvement. Investigation of the specific factors empirically related to poorer driving outcomes will assist development of targeted fitness to drive assessments that evaluate relevant functional abilities (OECD, 2001; Fildes, 2008; Anstey et al., 2012). Age itself does not reliably impact fitness to drive but declining medical, physical and cognitive functions typically associated with ageing have been found to increase crash risk (Janke, 1994; Charman, 1997; Daigneault et al., 2002; Anstey et al., 2005; Mathias & Lucas, 2009).

The most commonly investigated cognitive functions have included attention, processing speed, executive functioning, visuospatial skills, vision, and mental status; and several tests based on these abilities have been proposed as predictors of safe driving ability in older adults. One of the most successful fitness to drive tests[1] is the Useful Field of View Test™ (UFOV™; Ball & Owsley, 1993). UFOV™ performance has been shown to predict retrospective and prospective crash involvement, on-road driving performance, and driving simulator performance (Clay et al., 2005; Mathias & Lucas, 2009; Gentzler & Smither, 2012). UFOV™ Subtest 2 has been found to be particularly sensitive to driving outcomes. Subtest 2, designed to assess "processing speed for a divided attention task", is highly correlated with the Digit Symbol Substitution task which assesses visual processing speed (J. D. Edwards et al., 2006) (Edwards et al. 2006). Of the three subtests, Subtest 2 is most related to crashes (Owsley et al., 1998; Oxley et al., 2005; Ball et al., 2006), on-road driving (Bowers et al., 2013) and simulated driving performance (Molnar et al., 2007). Subtest 2 is therefore often considered on its own for purposes of brevity (Ball et al., 2006).

The test was originally developed to assess everyday visual difficulties encountered by older adults and has since been interpreted as assessing visual attention and processing speed (Ball et al., 1988, 1990, 1993; Owsley, 1994). The UFOV™ involves detection, localisation, and identification of stimuli located throughout the visual field and comprises three subtests of increasing difficulty that require identification of a central stimulus and localisation of a peripheral stimulus under different conditions (J. D. Edwards, Vance, et al., 2005). All subtests involve visual processing under limited time, justifying the conclusion that UFOV™ measures "an individual's speed of processing across increasingly complex visual displays" (J. D. Edwards, Vance, et al., 2005, p. 530).

Ball and colleagues have proposed that UFOV™ performance reflects the ability to rapidly scan the visual field and focus on salient features and that this in turn depends on speed of visual processing, divided attention, and selective attention (Ball et al., 1990; Ball, 1997). These three components are measured by the three subtests of UFOV™ which assess quality of performance when processing (i) brief stimuli located in central vision; (ii) similar stimuli presented concurrently in central and peripheral locations and (iii) detecting peripheral stimuli when distractors are also present in the visual field. Each

[1]After publication of this study, an anonymous reviewer has commented that the UFOV™ may be more accurately described here as "One of the more commonly used office-based tools used in research on assessing fitness to drive".

subtest has an independent and additive effect on overall UFOV™ performance (Ball et al., 1990).

UFOV™ has undergone only limited psychometric evaluation, and theories of visual attention have changed substantially since the test was first developed (Wolfe & Horowitz, 2004; Carrasco, 2011). Recently, Cosman, Lees, Lee, Rizzo, and Vecera (2012) have suggested that poor UFOV™ performance may be caused by deficits in attentional control processes, specifically reduced attentional disengagement (the ability to shift the focus of attention rapidly when required to do so), rather than a reduction in "attentional breadth"; i.e. the area in the visual field that can be searched within a single fixation. It is also possible that UFOV™ primarily reflects speed of information processing (Lunsman et al., 2008). Indeed, supporting evidence for this suggestion is correlation of UFOV™ performance with measures of processing speed, including WAIS-R digit symbol substitution (J. D. Edwards et al., 2006) and the Road Sign Test (J. D. Edwards, Vance, et al., 2005), with the digit symbol substitution test and UFOV™ showing a similar trajectory of decline accompanying older age (Lunsman et al., 2008).

Recent work has shown that other cognitive variables are also related to UFOV™ performance. Burns et al. (2005) found that selective attentional aspects of UFOV™ performance might more objectively be defined in terms of "crowding across the visual field", a widely studied phenomenon whereby visual interference occurs when a visually off-centre primary target is flanked by similar materials (Bouma, 1970; Pelli et al., 2004; Levi, 2008). Crowding has generally been considered a visual phenomenon but recent work suggests it may also involve limits to attentional resolution (He et al., 1996; Levi, 2008). Thus, a measure of crowding across the visual field has potential for more clearly identifying aspects of visual performance currently ascribed to visual attention, a term that has been widely applied to diverse functions (Plude, Enns, & Brodeur, 1994; Carrasco, 2011). Burns and White (2007) developed ProPerVis, a test of crowding across the visual field and visual processing speed. This test involves identifying briefly-presented flanked or unflanked stimuli in central, peripheral, and parafoveal vision. The crowding component of this test was found to be highly correlated with UFOV™ Subtest 3 (Burns et al., 2005). Results have also shown that performance on the crowding component declines with age, such that older adults find it much more difficult to extract information from a crowded visual field, especially when the stimuli of interest are presented in peripheral vision (Burns et al., 2005; Werner, 2008).

Hoffman et al. (2006) developed a visual change detection task (DriverScan) for older adults based on the "flicker" paradigm where a blank screen in displayed between an image and an altered version of that image, thereby masking luminance cues to the location of the change (Rensink et al., 1997). Rapid detection of change under these conditions requires focused attention to the area being changed (Rensink et al., 1997; Rensink, 2002). The DriverScan task involves detecting changes in "real world" driving scenes and on this basis it has been accepted as having ecological validity as a test for real-world driving skills, where drivers are required to be constantly vigilant for important changes in the environment like changes to traffic lights, position and velocity of other vehicles on the road, and the sudden presence of potential hazards. Along with sustained vigilance, successful detection of change also involves executive functioning, working memory, and processing speed (Pringle et al., 2001, 2004; Hoffman et al., 2006; Rizzo et al., 2009). Performance on change detection tasks is related to UFOV™ performance, leading to suggestions that change detection reflects the breadth of attentional focus (Pringle et al., 2001, 2004; Veiel, Storandt, & Abrams, 2006). Hoffman et al. (2005) found that performance on the change detection task was highly correlated with performance on UFOV™ divided attention (r = .50) and selective attention (r = .57) subtests. Furthermore, in a model incorporating measures of visual impairment, processing speed, and attention, both change detection and UFOV™ divided attention had independent significant direct effects on simulated driving performance. These results therefore suggest that, although DriverScan and UFOV™ Subtest 2 (divided attention) are related, they appear to rely on separate attentional processes.

Anstey et al. (2012) investigated the contribution of a battery of cognitive and visual tests to UFOV™ and two other tests of safe driving capacity: a hazard perception test and a change detection test. Factor analysis on the battery of tests revealed five factors: executive/speed, vision, spatial ability, visual closure, and working memory. Their results showed that UFOV™ performance was significantly related to the executive/speed, spatial ability, and working memory factors. Cognitive and visual factors together accounted for 40% of the variance in UFOV™ performance, 44% of variance in change detection performance, and 30% of variance in hazard perception performance.

Thus, UFOV™ performance has been shown to be related to a range of cognitive predictors including working memory, executive functioning, processing speed, visuospatial ability, and specific aspects of attention including attentional search and crowding across the visual field. Furthermore, UFOV™ performance is also related to age, education, vision, and eye health (Ball et al., 1993; J. D. Edwards et al., 2006).

Tests assessing variables found to be relevant to driving outcomes and UFOV™ performance in older adults were selected. These included tests of visual acuity and contrast sensitivity that are commonly used by licensing authorities to determine fitness to drive; a test of global cognitive functioning; and several recently developed cognitive tests tapping attention, processing speed, visuospatial skills and executive functioning. Specifically, the cognitive tests selected were DriverScan, ProPerVis, Inspection Time (IT), and the Snellgrove Maze Test. DriverScan and ProPerVis were selected due to their reported correlations with UFOV™. IT is a measure of processing speed and was included to verify the reported processing speed component of UFOV™. The IT task involves making a decision about which of two rapidly presented vertical lines is shorter (or longer). IT is notable for its high correlation to IQ, with a review of over 90 studies showing that the shared variance is around 25% (Grudnik & Kranzler, 2001). IT has been investigated as a biomarker for general cognitive decline (Gregory et al., 2008; Deary et al., 2010). IT predicts future results on tests of fluid reasoning, perceptual speed, and working memory, and changes in IT are predictive of future cognitive decline (Gregory et al., 2008). The Snellgrove Maze Test was included as quick, easy to administer assessment of executive function, visuoconstructional skills, and attention (Snellgrove, 2005). In a sample of cognitively-impaired older drivers, Snellgrove (2005) found that drivers who took longer than 60 seconds to complete the task were significantly more likely to fail an on-road driving test. The Snellgrove Maze Test and other versions of Maze Tests have been found to be related to driving ability, especially in older drivers with mild cognitive impairment or dementia (Ott et al., 2003; Whelihan, DiCarlo, & Paul, 2005; Ott et al., 2008; Carr, Barco, Wallendorf, Snellgrove, & Ott, 2011; Krishnasamy & Unsworth, 2011; Staplin, Gish, Lococo, Joyce, & Sifrit, 2013). The Mini Mental State Examination (MMSE) was included as a screening tool and as a measure of global cognitive functioning. Visual acuity and contrast sensitivity were included for use as covariates and because vision is often assessed as part of licensing procedures. The aim of the present study was to investigate the contribution of a set of cognitive, visual, and demographic variables to performance on UFOV™ in healthy older adults. On the basis of previous research it was predicted that the cognitive tests tapping change detection, response to visual field crowding, and visual processing speed would make the largest contribution to UFOV™ performance.

Materials and methods

Participants

Participants were 82 community dwelling older adults (52 females) living in Adelaide, South Australia, age 62 – 92 years, mean 75 years. Inclusion criteria were age over 60 years, living in the community, and free of severe mental impairment (MMSE > 24). Participants were recruited via advertisements placed in local newspapers and locations frequented by older adults, including gymnasiums and community centres. Participation was voluntary and no reimbursement was offered.

All participants were or had been drivers and 80 (95.1%) held a current South Australian license, six (7.3%) reported that they no longer drove a motor vehicle, 16 (19.5%) reported restricting their driving in some way, and 62 (73.2%) reported unrestricted driving.

Cognitive and visual measures

Visual acuity. Visual acuity was assessed using the Freiburg Visual Acuity Test (Bach, 1996). Participants indicated the orientation of a 'Tumbling E' optotype. Participants sat 1.65m from a computer screen and responded using the arrows on the computer keyboard. The size of the optotype presented varied on each trial depending on the current estimated threshold of the participant, calculated via the best Parameter Estimate by Sequential Testing procedure (Bach, 1996). There were 30 trials, with every sixth trial being an 'easy' trial where the optotype size was significantly larger than the current estimated threshold. Acuity was recorded as the logarithm of the Minimum Angle of Resolution (logMAR), with lower scores representing better visual acuity.

Contrast Sensitivity. Contrast sensitivity was measured using a Pelli-Robson contrast sensitivity chart (Pelli et al., 1988). The chart consists of eight rows of six letters, arranged into groups of three letters. The contrast of the letters reduces for each subsequent triplet. Participants viewed the chart at eye level from a distance of 1m while wearing normal vision correction. Participants were instructed to read the letters on the chart as far down as they could and were encouraged to guess if they were uncertain. The outcome was log contrast sensitivity, which was measured according to the faintest group of three letters for which the participant correctly identified two of the three letters. Possible scores range from 0 to 2.25, with higher scores indicating better contrast sensitivity.

Mini Mental State Examination (MMSE). The MMSE is a short, simple test of global cognitive function used to estimate the extent of any cognitive impairment (Folstein et al., 1975). It was administered according to standardised instructions and scoring proposed by Molloy and Standish (1997). Participants answer questions assessing short-term memory, orientation in space and time, constructional ability, executive functioning, and ability to follow instructions. Participants answer verbally, in writing, and by performing actions when requested. The MMSE is scored out of 30 and state of cognitive impairment can be interpreted as follows: $26 - 30$, 'could be normal'; $20 - 25$, 'mild'; $10 - 19$, 'moderate'; $0 - 9$, 'severe' (Vertesi et al., 2001).

Snellgrove Maze Test. The Snellgrove Maze test is a simple pencil and paper maze designed for use with people with dementia and mild cognitive impairment (Snellgrove, 2005). Performance on the task depends on attentional skills, executive functioning, and psychomotor speed. Older drivers with mild cognitive impairment and early dementia who took longer than 60 seconds to complete the maze were more likely to fail an on-road driving assessment (Snellgrove, 2005). The maze was presented in black on A4 paper with arrows indicating the entry and exit points. Participants used a pencil to complete the maze as quickly as they could while making as few errors as possible. Time to complete the maze was recorded using a stopwatch in seconds.

Inspection Time (IT). IT measures processing speed (Burns & Nettelbeck, 2003). Two high-contrast lines, one markedly shorter than the other, appear as a target on a computer screen. Participants indicate whether the shorter line was located left or right of a focal point. Time available for processing is limited by a backward masking procedure and reduced or extended using an adaptive staircase algorithm according to response accuracy. Targets are preceded by a warning cue ("+" in the centre of the screen, 370ms) and are immediately followed by a mask shaped like two lightning bolts. There were three sets of 10 practice trials with decreasing target presentation time for each set (835ms, 420ms, and 250ms). Participants were requires to answer 10/10 items correctly for the first and second set, and 9/10 items for the third set. IT was measured in ms as the duration between target onset and mask onset at which the viewer achieves 79% accuracy. Test-retest reliability for adults is .81 (Grudnik & Kranzler, 2001).

DriverScan. DriverScan is a change detection task designed to assess visual attention skills needed for safe driving (Hoffman et al., 2005). The test uses the change detection paradigm proposed by Rensink et al. (1997). Real world images of driving scenes (A) presented on a computer screen were alternated with a blank screen and an altered version of the image (A') to which a small change has been made, creating a blinking effect. Changes included deletion of objects (e.g. cars), colour/lettering changes (e.g. road signs), and signal changes (e.g. traffic lights, brake lights). The pattern of presentation is A, blank, A, blank, A', blank, A', blank, A, blank, A, blank... etc., with the images (A, A') being presented for 280 ms and the blank screen being presented for 80ms. There were four practice trials displaying different types of changes. If participants did not answer correctly or were not confident, the practice trials were repeated. This was followed by 25 trials where each trial was presented for 45 seconds or until a response was recorded. Participants viewed the images and attempted to detect the change between A and A', responding by clicking the mouse as soon as they detected a change and then verbally describing the change to the experimenter. The response time for each trial was recorded, and these times were used to generate an ability estimate for each participant using a constrained Graded Response Model (an application of Item Response Theory for polytomous data; see Samejima, 1970; Hoffman et al., 2006). Responses were classed as immediate response (< 8 s), delayed response ($8 - 45$s), no response, or incorrect response. Higher ability scores represent better performance on the test (fast, correct responses).Viewing distance was approximately 60cm.

Proficiency of Peripheral Visual Processing (ProPerVis). ProPerVis assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns & White, 2007). There are two subtests: Crowding and Inspection Time (see Burns et al., 2005). The stimuli are a four-square parent figure and six figures derived from it, resembling characters M, E, W, 3, 5, 2 (see Figure 4.1). In each of the 40 crowding trials, one of the six figures is flanked on either side by the parent figure and appears randomly in one of five lateral positions on the screen. Participants completed 10 practice trials before commencing. In each of the 60 IT trials, one of the six figures appears briefly on the screen in one of three lateral positions before being masked by the parent figure. Participants were required to respond correctly to five practice trials of 250ms in a row before commencing. In both subtests, participants respond by clicking on the figure that they saw on a response panel on the screen. Although both subtests were administered, only the crowding subtest was used for further analysis because of the inclusion of a separate test of IT. The outcome was the total number of errors made for the crowding subtest.

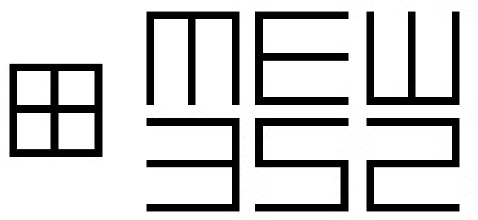

Figure 4.1: ProPerVis stimuli. The four-square parent figure is shown to the left of the six target stimuli.

Useful Field of View Test™

UFOV™ is a computer-based test of visual attention and processing speed involving detection and localisation of briefly presented targets throughout the visual field (Ball & Owsley, 1993). There are three subtests: processing speed, divided attention, and selective attention. Four practice trials were first completed for all subtests. Two stimuli are used; a silhouette of a car and a truck, sized 2 × 1.5 cm. In the processing speed subtest

(Subtest 1), one of these stimuli is briefly presented within a central fixation box of 3 × 3 cm at varying exposure durations across trials in central vision, and the observer indicates which stimulus they saw by clicking the appropriate symbol on the screen. For the divided attention subtest (Subtest 2), one of the two stimuli appears briefly in central vision as before, and the car appears simultaneously in the periphery at one of the eight cardinal or intercardinal points, 12.5 cm from the centre of the display. Participants indicate which object was presented in central vision, and the location of the car in the periphery. The task for the selective attention subtest (Subtest 3) is the same but with the addition of 47 distractors (triangles of the same size and luminance as the target stimuli) distributed evenly throughout the visual field. For each subtest, the score is exposure time (ms) for which 75% of responses are correct. UFOV™ total score was calculated by summing the score from each subtest. Test-retest reliability for total UFOV™ score is .88 (J. D. Edwards, Vance, et al., 2005).

Procedure

The study was conducted either in the laboratory or in the participant's home[2] and took 1−2 hours to complete. Participants were provided with information about the study and were given the opportunity to ask the researcher questions. They then completed a consent form and a safety induction form and provided demographic information (age, years of education, years of driving experience, driver's license, driving status, and crashes in the past five years). The tests were then conducted in the following order: MMSE, Snellgrove Maze, Pelli-Robson Contrast Sensitivity, Freiburg Visual Acuity, DriverScan, IT, UFOV™, ProPerVis. The order was the same for all participants. Participants were encouraged to complete the tests at their own pace and to take breaks when needed.

This research was approved by the University of Adelaide Human Research Ethics Committee. All participants provided written consent to participate.

[2]Testing procedure did not include an evaluation of light levels. It is therefore possible that this could have impacted contrast sensitivity.

Analysis

Two participants were excluded from the analysis because of missing data from equipment failure and incomplete tests. Linear regression models were used to assess amount of variance in UFOV™ performance that could be accounted for by the predictor variables. Each subtest and the total score were considered separately. Models were selected via backward elimination from an initial full model including age, MMSE, education, contrast sensitivity, visual acuity, DriverScan, IT, ProPerVis crowding, and the Maze test[3]. At each step, a single predictor was eliminated based on the largest reduction in the Akaike Information Criterion (AIC; Akaike, 1974). Elimination ceased when removal of additional predictors no longer reduced the AIC and therefore did not improve the goodness-of-fit of the model (R package `car`; Fox & Weisberg, 2011). Regression assumptions were checked using global validation of linear model assumptions (R package `gvlma`; Peña & Slate, 2006). Assumptions were acceptable for variables regressed on UFOV™ total score and UFOV™ Subtests 2 and 3. Assumptions were not satisfied for UFOV™ Subtest 1. This subtest displayed substantial floor effects; 61% of participants (n = 50) achieved the best possible score (16.7ms). This lack of variance meant that relationships between the predictors and performance were difficult to observe for this subtest. As is clear from Table 4.2, there was no evidence for multicollinearity (maximum Variance Inflation Factor (VIF) = 2.07; Menard, 1995).

The relative importance of each predictor in the final model after backward elimination was calculated. Relative importance regression is a method for calculating the share of the model R^2 that each regressor accounts for when the regressors are correlated. Relative importance is defined as "the proportionate contribution each predictor makes to the R^2, considering both its direct effect (i.e. its correlation with the criterion) and its effect when combined with other variables in the regression equation" (J. W. Johnson & LeBreton, 2004). There are several alternative methods for calculating relative importance (see Grömping, 2006). One method that is generally accepted to provide an

[3] After publication of this study, an anonymous reviewer has commented that it would have been useful to include driving status (drivers vs non-drivers) in the models. It is possible that driving status may have been a proxy for other underlying functional measures related to driving and therefore may have contributed to the overall variance in the models.

accurate decomposition of R^2 is the LMG method, first proposed by Lindeman, Merenda, and Gold (1980). LMG calculations of relative importance utilise semi-partial correlations between the predictors and the outcome. This method uses equivalent methods to dominance analysis (Budescu, 1993; Azen & Budescu, 2003; J. W. Johnson & LeBreton, 2004) and hierarchical partitioning (Chevan & Sutherland, 1991; Grömping, 2006). The LMG method calculates the contribution of each variable to the model across all possible orderings of regressors. When a regressor is entered into the model first, its share is its squared correlation with the outcome. When a regressor is entered last, its shared is the portion of the R^2 that has not already been accounted for by all other predictors in the model (i.e. its semi-partial correlation with the outcome). The percentage contribution of each regressor in each ordering is calculated and its relative importance value is the average percentage contribution across all orderings. Therefore the model incorporates the direct effect of each predictor (its share when entered first) and its effect when combined with all other predictors in the mode its share when entered last) as required by the definition stated by J. W. Johnson and LeBreton (2004). The LMG method was selected because it decomposes R^2 among the predictors in non-negative, non-zero shares, and is most suited for making causal interpretations (J. W. Johnson & LeBreton, 2004; Grömping, 2006, 2007; Chao, Zhao, Kupper, & Nylander-French, 2008).

All analyses were run using R version 2.15.3 (R Core Team, 2013). Relative importance metrics were produced using the R package `relaimpo` (Grömping, 2006). Results were considered statistically significant if associated probability $p < .05$.

Results

Descriptive statistics and correlations

Table 4.2 shows descriptive statistics for all predictor variables and the outcome measure (UFOV™ total score). Overall, the sample was cognitively healthy and highly educated. Most participants had good vision, but eight had poor visual acuity (with corrected vision) above logMAR 0.3, the minimum acuity required to hold a driver's license in Australia and most other countries (Bron et al., 2010), and seven had poor contrast sensitivity below 1.5, indicating impairment. Nine participants reported being involved in a minor crash within the past five years. There was no significant difference between crash-involved and non-crash-involved participants on any variable. The mean for UFOV™ total score (367.4 ± 209.5 SD ms) was significantly quicker than the norm of 481.9 ms reported by J. D. Edwards et al. (2006) for a large sample with similar characteristics, $t(81) = 4.95$, $p < .001$

Pearson correlations for all variables are shown in Table 4.2. All variables except age, education and contrast sensitivity were significantly skewed but reanalysis using Spearman's rho found essentially the same results. All correlations were in the expected directions. Age was significantly and moderately to highly correlated with all other variables. All cognitive variables were highly intercorrelated. All predictor variables were significantly related to total UFOV™ score. UFOV™ total score was most highly correlated with change detection (DriverScan; $r = -.64$), crowding (ProPerVis; $r = .59$), and cognitive integrity (MMSE; $r = -.57$).

Table 4.1: Descriptive statistics for all study variables

Variable	M	SD	Mdn	Range
Age (years)	75.12	7.79	73	62 – 92
Education (years)	15.13	4.28	15	5 – 25
MMSE	29.28	0.86	30	27 – 30
Contrast Sensitivity (log contrast)	1.76	0.18	1.65	1.2 – 1.95
Visual Acuity (logMAR)	0.11	0.15	0.07	−.15 – 0.69
Maze (s)	31.63	14.42	28.97	10.19 – 101.97
IT (ms)	66.35	26.03	61.88	33 – 199.38
Crowding Errors	13.1	6.19	12	1 – 31
DriverScan Ability Estimate	0.01	0.91	0.15	−3 – 1.81
UFOV™ Subtest 1	28.34	30.06	16.7	16.7 – 193.6
UFOV™ Subtest 2	115.5	106.69	85.1	16.7 – 500
UFOV™ Subtest 3	223.6	103.43	210.4	56.7 – 500
UFOV™ Total (ms)	367.39	209.52	337.05	90.10 – 1040.8

Note. MMSE = Mini Mental State Examination, IT = Inspection Time, UFOV™ = Useful Field of View Test™

Table 4.2: Pearson correlations between all study variables

	1	2	3	4	5	6	7	8	9	10	11	12
1. UFOV™ Subtest 1												
2. UFOV™ Subtest 2	.43											
3. UFOV™ Subtest 3	.49	.68										
4. UFOV™ Total	.61	.91	.91									
5. Age	.26	.48	.50	.53								
6. Education	-.16	-.29	-.28	-.31	-.41							
7. MMSE	-.38	-.45	-.57	-.57	-.33	.19						
8. Contrast sensitivity	-.47	-.36	-.59	-.55	-.51	.30	.47					
9. Visual acuity	.47	.37	.49	.50	.39	-.25	-.39	-.53				
10. Maze	.41	.39	.40	.45	.42	-.18	-.22	-.29	.28			
11. IT	.23	.40	.49	.48	.36	-.10	-.26	-.25	.28	.24		
12. Crowding	.41	.43	.63	.59	.41	-.34	-.39	-.46	.40	.38	.37	
13. DriverScan	-.53	-.57	-.55	-.64	-.48	.37	.39	.53	-.49	-.44	-.35	-.59

Note. MMSE = Mini Mental State Examination, CS = Contrast Sensitivity, VA = Visual Acuity, IT = Inspection Time, Crowding = ProPerVis Crowding, DriverScan = DriverScan change detection ability estimate. UFOV™ Total is the sum of UFOV™ Subtest 1, UFOV™ Subtest 2, and UFOV™ Subtest 3.

Regression

UFOV™ Subtests. For UFOV™ Subtest 1, the model selected via backward elimination included contrast sensitivity, change detection as measured by DriverScan, visual acuity, age, and the maze test as predictors, as shown in Table 4.3. DriverScan and the maze were significant predictors. This model explained 41% of the variance in UFOV™ Subtest 1 score, $F(5, 76) = 10.65$, $p < .01$. Relative importance regression showed that DriverScan accounted for 32% of the R^2, contrast sensitivity and visual acuity accounted for about 22% each, the maze accounted for 19% and age for 5%.

For UFOV™ Subtest 2, the model selected via backward elimination included MMSE, IT, change detection as measured by DriverScan, and age as predictors, as shown in Table 4.3. DriverScan and MMSE were significant predictors. This model explained 44% of the variance in UFOV™ Subtest 2 score, $F(4, 77) = 15.29$, $p < .01$. Relative importance regression showed that DriverScan accounted for 39% of the R^2, MMSE and age accounted for 23% each, and IT accounted for 15%.

For UFOV™ Subtest 3, the model selected via backward elimination included crowding errors as measures by ProPerVis, MMSE, contrast sensitivity, IT, and the maze test as predictors, as shown in Table 4.3. Crowding, MMSE, contrast sensitivity, and IT were significant predictors. This model explained 63% of the variance in UFOV™ Subtest 3 score, $F(5, 76) = 25.56$, $p < .01$. Relative importance regression showed that crowding accounted for 27% of the R^2, contrast sensitivity accounted for 24%, MMSE accounted for 23%, IT accounted for 17%, and the maze accounted for 9%. Contributions of the predictors to the R^2 for each subtest are shown in Figure 4.2.

UFOV™ Total. For UFOV™ Total score, the model selected via backward elimination included age, MMSE, IT, crowding errors as measured by ProPerVis, change detection as measured by DriverScan, and time to complete the maze as predictors, as shown in Table 4.3. IT, change detection, and MMSE were significant predictors. This model explained 64% of the variance in UFOV™ Total score, $F(9, 72) = 14.05$, $p < .01$. Relative importance regression showed that DriverScan accounted for 24% of the R^2, MMSE accounted for 22%, ProPerVis crowding accounted for 18%, and age and IT accounted for 13% each, and the maze accounted for 10%. Contributions of the predictors

Table 4.3: Regression of predictors on UFOV™ performance outcomes

Model	R^2	B	sig	SE B	95% CI B	lmg (R^2)	lmg (%)
UFOV™ Subtest 1	.41						
Intercept		131.85		54.03	24.24, 239.46		
DS		−9.77	*	3.86	−17.46, −2.09	.13	.32
CS		−39.70		19.97	−79.47, 0.07	.09	.22
VA		40.60		21.37	−1.95, 83.15	.09	.22
Maze		0.49	*	0.21	0.07, 0.91	.08	.19
Age		−0.71		0.43	−1.56, 0.15	.02	.05
UFOV™ Subtest 2	.44						
Intercept		681.05		372.96	−61.62, 1423.72		
DS		−39.80	**	12.11	−63.91, −15.68	.17	.39
MMSE		−27.32	*	11.64	−50.49, −4.14	.10	.23
Age		2.57		1.38	−.190, 5.32	.10	.23
IT		0.63		0.38	−.13, 1.40	.07	.15
UFOV™ Subtest 3	.63						
Intercept		1234.37	**	279.40	677.90, 1790.84		
Crowding		4.70	**	1.45	1.81, 7.59	.17	.27
CS		−151.16	**	50.13	−251.00, −51.31	.15	.24
MMSE		−30.41	**	9.79	−49.92, −10.91	.14	.23
IT		0.91	**	0.30	0.30, 1.52	.11	.17
Maze		0.77		0.55	−0.33, 1.87	.06	.09
UFOV™ Total	.64						
Intercept		1956.69	**	617.89	725.8, 3187.58		
DS		−58.95	**	22.20	−103.17, −14.72	.15	.24
MMSE		−70.45	**	19.24	−108.77, −32.13	.14	.22
Crowding		5.39		3.13	−108.77, −32.13	.11	.18
Age		3.33		2.33	−0.84, 11.63	.08	.13
IT		1.44	*	0.63	0.18, 2.71	.08	.13
Maze		1.80		1.19	−0.57, 4.18	.06	.10

Note. DS = DriverScan change detection ability estimate; CS = contrast sensitivity; VA = visual acuity; lmg (R^2), lmg (%) = Relative importance of each predictor calculated using the lmg method, expressed relative to the total R^2 (lmg values sum to R^2) and as a proportion of the total R^2 (lmg values sum to 1).

* $p < .05$, ** $p < .01$

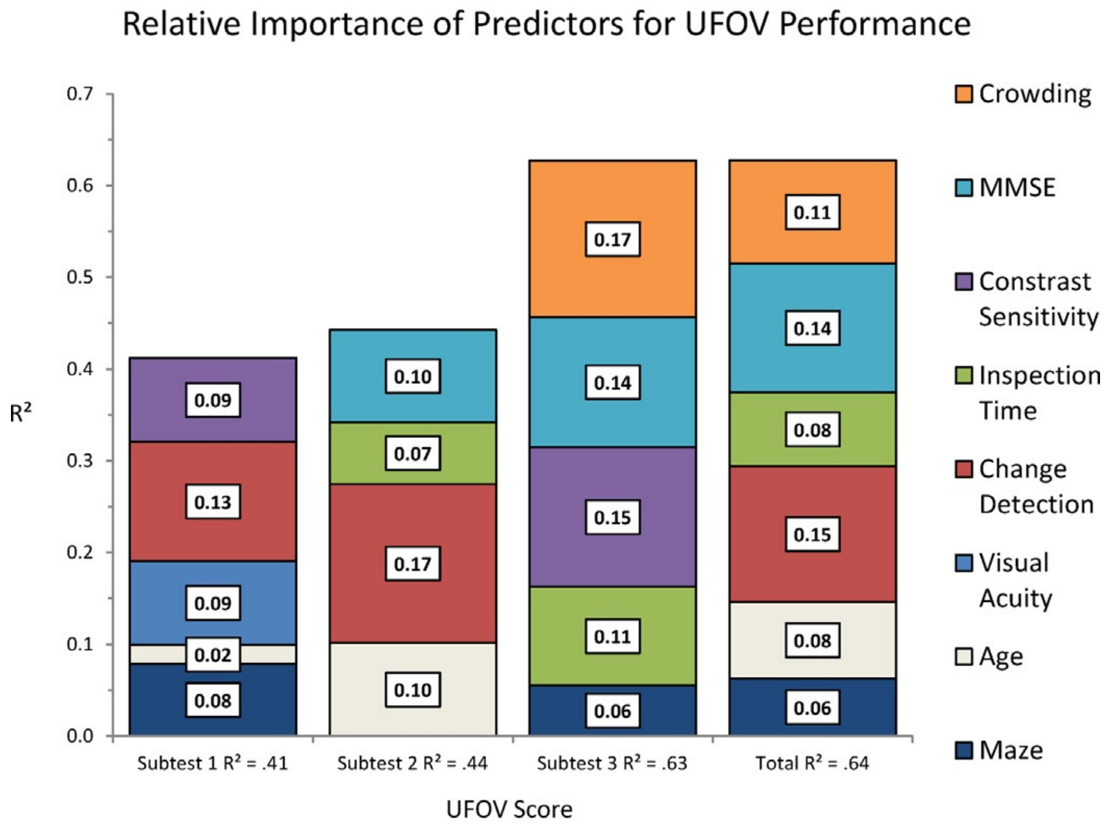

Figure 4.2: Relative importance of predictors for UFOV™ performance. UFOV™ Subtest 1 assesses "processing speed", Subtest 2 assesses "divided attention", and Subtest 3 assesses "selective attention" (Ball and Owsley, 1993). UFOV™ Total Score is the sum of scores from the three subtests. Predictors are crowding, MMSE, Contrast Sensitivity, IT, Change Detection, Visual Acuity, Age, and the Maze Test. Relative importance indices were calculated using the lmg method (Grömping, 2006). The relative importance of each variable represents the part of the total R^2 that can be attributed to that variable.

to the R^2 for each subtest are shown in Figure 4.2.

Discussion

The aim was to investigate the contribution of a battery of cognitive and visual tests to performance on the Useful Field of View Test™ (UFOV™). We predicted that tests reflecting attention and visual processing speed would be the most important contributors to UFOV™ performance. UFOV™ performance measures (individual subtests scores and total score) were regressed on the nine predictor variables: age, education, contrast sensitivity, visual acuity, MMSE, IT, crowding, and DriverScan. Results of relative importance analyses revealed a different profile of contributing measures for each subtest.

For UFOV™ Subtest 1 ("processing speed"; Ball & Owsley, 1993), the most important predictors were change detection, visual acuity, contrast sensitivity, and Maze Test performance. It should be noted that the regression model was a poor fit for the model because of skewed results on this subtest. Floor effects have been previously noted for this subtest (Burns et al., 2005; J. D. Edwards, Vance, et al., 2005; Rubin et al., 2007; Bentley, LeBlanc, Nicolela, & Chauhan, 2012). The norms reported by J. D. Edwards et al. (2006) for older adults on Subtest 1 (M = 30.66, SD = 40.51) suggest that most healthy older adults score at or near the minimum score of 16.7 ms, indicating that this subtest is too easy for this population. Nevertheless, the results tend to suggest that low-level visual processes are responsible for poorer performance; visual acuity and contrast sensitivity together accounted for 44% of R^2 (22% each). Speed of processing (IT) was not a predictor but change detection accounted for an additional 32% of the variance, suggesting that this subtest is more a test of focused visual attention than of processing speed, except to the extent that DriverScan relies on processing speed.

For UFOV™ Subtest 2 ("divided attention"), the most important predictors were change detection (39% of R^2), MMSE (23%), age (23%), and IT (15%). Subtest 2 requires detection of peripheral and central stimuli in an uncrowded visual field. The results therefore suggest that Subtest 2 reflects an age-related decline in change detection ability, efficiency of visual processing, and general cognitive ability. Given the very narrow range of MMSE scores (see Table 4.1), the latter result was not expected. It is discussed further, below. Change detection requires similar processes to visual search, including focussed attention (Rensink et al., 1997; Rensink, 2000). UFOV™ Subtest 2 therefore appears to reflect the efficiency of focussed divided attention over a larger spatial area. The ability to adjust the spatial focus of attention has been termed "attentional scaling", and this ability has been shown to decline with age (Greenwood & Parasuraman, 2004).

Of the three UFOV™ subtests, Subtest 2 has been found to be most related to crash occurrence for older drivers (Owsley et al., 1998; Ball et al., 2006). Of interest is what this subtest measures and why it in particular is most sensitive to driving outcomes. DriverScan and IT together accounted for 24% of the variance, confirming that UFOV™ Subtest 2 does involve aspect of visual attention and processing speed. Global cognitive functioning and age accounted for 20% of the variance; UFOV™ Subtest 2 is known to be related to these variables in normal samples (J. D. Edwards et al., 2006). As expected, Subtest 2 was not associated with low-level visual functioning. UFOV™ is not a test of visual sensory function, except to the extent that non-impaired vision is required for visual attention tasks (Ball et al., 1993). Visual acuity and contrast sensitivity, the vision tests included in the present study, have typically been reported to have small or no correlation with crash outcomes or driving performance (Wood, 2002; Anstey et al., 2005; Rubin et al., 2007; Owsley & McGwin Jr, 2010). Overall, the predictors in this model were able to account for 44% of the variance in UFOV™ Subtest 2 scores, meaning that an additional 56% in variance in scores was not explained by the variables included in the present study. It is likely that some of this variance is related to aspects of visual attention not picked up by DriverScan, and that these aspects, in addition to processing speed and attention skills required for change detection, and are what makes UFOV™ Subtest 2 particularly useful for predicting crashes.

For UFOV™ Subtest 3 (labelled "selective attention") the most important predictors were crowding, contrast sensitivity, MMSE, and IT. Subtest 3 involves detection of peripheral and central stimuli in a crowded visual field. The crowding measure accounted for 27% of R^2. This supports the results of Burns et al. (2005) who found that this subtest was principally a measure of sensitivity to crowding. Crowding is thought to reflect limits to visual processing, but may also reflect limits to attentional resolution (He et al., 1996; Levi, 2008; Whitney & Levi, 2011). Contrast sensitivity accounted for 24% of R^2 and IT accounted for 17% of R^2. As in Subtest 2, the MMSE made a significant contribution.

Overall, the results suggest that Subtest 3 reflects efficiency of low-level visual and attentional processing (processing speed, contrast sensitivity and susceptibility to crowding), and general cognitive ability.

For UFOV™ total score (sum of scores from the three subtests, reflecting overall speed of visual processing over varying visual conditions), the most important predictors were change detection (23%), MMSE (22%), and crowding (18%). Change detection contributed highly to performance on Subtests 1 and 2, whereas MMSE contributed highly to performance on Subtests 2 and 3. The influence of crowding comes entirely from Subtest 3. Age and IT accounted for additional variance (13% each) due to moderate influence across the three subtests. Therefore, as predicted, overall UFOV™ performance appears to reflect age-related declines in aspects of attention and efficient visual processing (change detection, crowding, and IT).

Change detection made a prominent contribution to UFOV™ performance overall, primarily through its large contribution to Subtest 2. Several previous studies have shown that UFOV™ and change detection are related, indicating that they both rely on similar attentional processes (Pringle et al., 2001, 2004; Hoffman et al., 2005; Veiel et al., 2006; Rizzo et al., 2009). Pringle et al. (2004) showed that individual differences in change detection ability could be predicted by visuospatial working memory, "attentional breadth" (UFOV™), and perceptual speed. Veiel et al. (2006) used eye-tracking analysis and showed that older adults used shorter saccades, returned to items that had previously been viewed, and fixated longer before responding. They argued that this could reflect reduced attentional breadth (thus requiring more eye movements) and could also reflect reduced Inhibition of Return (rechecking of previously fixated areas) and a more cautious strategy for responding (longer fixations). Rizzo et al. (2009) found that change detection ability and UFOV™ were correlated $r_s = -.50$, $p = .001$ in a combined sample of Alzheimer's disease (AD) patients and healthy controls (i.e. poorer change detection accompanied slower UFOV™ performance). This effect among AD patients ($r_s = -.75$, $p = .004$) was stronger than among healthy controls ($r_s = -.34$, $p = .089$). There was also a marginal correlation between UFOV™ and a test of short term memory in the combined sample, $r_s = -.32$, $p = .059$. Change detection requires focussed attention to be consciously directed by high-level processes to objects in the scene; without focussed attention to the change, appropriate encoding and comparison of objects is interrupted

(Rensink et al., 1997; Simons & Ambinder, 2005). These results suggest that the portion of UFOV™ performance predicted by change detection ability may reflect a shared reliance on attentional breadth and control of focussed attention. Further research is needed to examine the specific aspects of attention that explain UFOV™ performance.

Crowding contributed to UFOV™ performance through its strong influence on Subtest 3. Subtest 3 is the only subtest that incorporates visual noise. Performance on this subtest thus reflects the ability to discriminate stimuli in a crowded visual field. That crowding was such an important contributor suggests that this phenomenon may better explain performance on Subtest 3 than "selective attention", a somewhat vague label that is currently used to describe this subtest. This result supports Burns et al. (2005) who found that Subtest 3 could be explained by measure of crowding and processing speed. Like UFOV™ performance, susceptibility to crowding increases with age and eccentricity of stimuli in the visual field, and age and eccentricity interact such that older adults make more errors at greater eccentricities (Burns et al., 2005).

Visual processing speed (IT) made significant contributions to Subtests 2 and 3. This result concurs with several other studies linking UFOV™ and processing speed (Sekuler, Bennett, & Mamelak, 2000; J. D. Edwards, Vance, et al., 2005; J. D. Edwards et al., 2006; Lunsman et al., 2008). Processing speed is argued to contribute to performance across all subtests of UFOV™ (J. D. Edwards, Vance, et al., 2005). However, in the present study IT did not contribute significantly to Subtest 1. As mentioned above, clear floor effects were evident for Subtest 1, indicating that Subtest 1 is too easy and thus not a useful measure of "processing speed" in this type of sample (healthy older adults). Nonetheless, the contribution of IT to the two more complex subtests suggests that speed of processing contributes to overall UFOV™ performance such that UFOV™ reflects the efficiency of visual processing across a broad attentional area, with or without visual noise in the scene.

The Maze Test made small contributions to Subtest 1, Subtest 3, and total score. Maze navigation requires executive functioning skills including planning and foresight (Snellgrove, 2005), skills that are not normally mentioned in relation to UFOV™ performance. It is likely that there is not much overlap in the skills required for maze navigation and UFOV™ performance. It may be the case that maze test performance is related to driving performance but not to UFOV™ performance. Another explanation for the lack of relationship is that the maze test is usually used to assess clinical samples (Ott et al., 2003; Snellgrove, 2005; Whelihan et al., 2005; Ott et al., 2008; Carr et al., 2011; Krishnasamy & Unsworth, 2011; Ott et al., 2013). The sample in the present study was very high-functioning (mean MMSE 29.3). It is therefore possible that the Maze test is particularly useful for patients with mild cognitive impairment or dementia, but is not sensitive enough at higher ability levels. The majority of participants easily completed the maze before the 60 second cutpoint noted by Snellgrove (2005).

One surprising finding was the importance of the MMSE to performance on Subtests 2 and 3. This measure was included for screening purposes and was not expected to contribute significantly to UFOV™ performance, especially considering the restricted range in results; all participants scored 27 or above out of 30, within the "could be normal" range of the scale (Vertesi et al., 2001). Nevertheless, a strong linear relationship was evident between MMSE and UFOV™ total score, $r = -.57$. The relationship held when three participants with the lowest score of 27 were excluded. Findings of the relationship between MMSE and driving performance are mixed. Meta-analyses have shown correlations of between .4 to .6 for the MMSE and driving outcomes and the MMSE has been reported as having good discriminability for predicting driving performance (Adler et al., 2005; Mathias & Lucas, 2009). On the other hand, several studies have found no association between MMSE score and reported crashes or on-road assessment (see Molnar et al., 2006) and it has been argued that MMSE score alone is not an accurate predictor of driving ability (Crizzle et al., 2012). MMSE and UFOV™ score are correlated, with J. D. Edwards et al. (2006) reporting a correlation of $r = -.330$ for MMSE and UFOV™ 3 Subtest total in a sample with MMSE scores ranging from $24 - 30$. However, the fact that the MMSE still contributed significantly to UFOV™ performance when all other cognitive variables were entered into the regression model was unexpected. This suggests that global cognitive functioning and basic cognitive abilities, namely memory and ability to follow instructions, may be required for UFOV™ performance. It is possible that MMSE contributes to ability to perform well on the UFOV™ test, but is not reliably related to driving ability in non-impaired older drivers. Future research will further investigate the reliability of this result, and whether MMSE scores in the higher range can predict differences in on-road driving performance or future crash outcomes. If the relationship holds, it may suggest that the MMSE is particularly sensitive for identifying very mild cognitive decline which is relevant to driving ability in older drivers.

Conclusion

This study investigated the contribution of a number of cognitive and visual variables to performance on the Useful Field of View Test™ (UFOV™). Results of relative importance regression analyses showed a different profile of contributing variables for each subtest. Subtest 1 ("processing speed") was explained primarily by low-level visual factors, namely contrast sensitivity and visual acuity; however, this subtest was too easy for this sample with the majority of participants achieving the best possible score. Subtest 2 ("divided attention") was explained by change detection ability, processing speed, general cognitive ability, and age. Subtest 3 ("selective attention") was explained by crowding, contrast sensitivity, processing speed, and general cognitive ability. The strong influence of crowding on Subtest 3 suggests that reduced performance on this subtest may be better attributed to crowding effects than to "selective attention". The finding that general cognitive ability contributed to overall performance through the significant influence of the MMSE on the more complex Subtests 2 and 3 was novel and needs to be investigated further. Overall, this study has supported the hypothesis that UFOV™ measures attention and processing speed, however further research is needed to determine the specific aspects of attention that are involved.

Chapter 5.
Study 2: Simulator sickness and dropout during a driving simulator study: A survival analysis

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2015). Simulator sickness and dropout during a driving simulator study: A survival analysis. *Journal of Safety Research*, 55, 159-169. doi: 10.1016/j.jsr.2015.08.004

Statement of Authorship

Title of paper	Simulator sickness and dropout during a driving simulator study: A survival analysis
Publication status	Published
Publication details	Matas, N. A., Nettelbeck, T., & Burns, N. R. (2015). Simulator sickness and dropout during a driving simulator study: A survival analysis. *Journal of Safety Research*, 55, 159-169. doi: 10.1016/j.jsr.2015.08.004

Principal Author

By signing the Statement of Authorship, I certify that this paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper

Name of principal author	Nicole Amy Matas		
Contribution to paper	Study concept and design, data collection, statistical analyses, writing and submitting manuscript, addressing referee comments		
Overall percentage	80%		
Signature		Date	13/9/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ted Nettelbeck		
Contribution to paper	Principal supervision, advice about development of concept and design of study, manuscript revision and approval		
Signature		Date	13. 9. 16

Name of Co-Author	Nick Burns		
Contribution to paper	Advice about development of concept and design of study, assistance with statistical analyses, manuscript revision and approval, ongoing		
Signature		Date	14/9/16

Abstract

Introduction: Simulator Sickness (SS) is the occurrence of motion-sickness like symptoms that can occur during use of simulators and virtual reality technologies. This study investigated individual factors that contributed to SS and dropout while using a desktop driving simulator. **Method:** Eighty-eight older adult drivers (mean age 72.82 $\pm$ 5.42 years) attempted a practice drive and two test drives. Participants also completed a battery of cognitive and visual assessments, provided information on their health and driving habits, and reported their experience of SS symptoms throughout the study. **Results:** Fifty-two participants dropped out before completing the driving tasks. A time-dependent Cox Proportional Hazards model showed that female gender (HR = 2.02), prior motion sickness history (HR = 2.22) and Mini-Simulator Sickness Questionnaire (SSQ) score (HR = 1.55) were associated with dropout. There were no differences between dropouts and completers on any of the cognitive abilities tests. **Conclusions:** Older adults are a high-risk group for SS. Within this group, female gender and prior motion sickness history are related to simulator dropout. Higher reported experience of symptoms of SS increased rates of dropout. **Practical Applications:** The results highlight the importance of screening and monitoring of participants in driving simulation studies. Older adults, females, and those with a prior history of motion sickness may be especially at risk.

Introduction

Driving simulators are becoming more widely available and these instruments have many useful applications for research, training, assessment, rehabilitation, and entertainment (Allen et al., 2011; Classen & Brooks, 2014; Crisler et al., 2011; Dickerson et al., 2014; Pollatsek et al., 2011). The availability of lower-cost options means that driving simulators are now increasingly accessible to researchers and therapists. Simulators have been used successfully to investigate how to improve the training of novice drivers, (Allen et al., 2012; de Winter et al., 2009; Pollatsek et al., 2011), and does re-training older drivers and patients with acquired brain injury (Casutt et al., 2014; Pollatsek et al., 2012; Unsworth & Baker, 2014). They have also proved useful for investigating distractions common among on-road drivers including use of cell phones while driving (Caird, Willness, Steel, & Scialfa, 2008), text messaging (Caird, Johnston, Willness, Asbridge, & Steel, 2014), and use of in-vehicle entertainment systems (Engström et al., 2005; Horberry et al., 2006); and, more generally, for monitoring driver responses to challenging driving situations (Bélanger et al., 2010; de Waard, Dijksterhuis, & Brookhuis, 2009; Martin et al., 2010). They have also found wide application for studying the relationship between cognitive abilities and driving performance (Bélanger et al., 2010; Hoffman et al., 2005; Shanmugaratnam et al., 2010) and the effects of cognitive interventions on driving performance (Roenker et al., 2003). A survey of driver rehabilitation specialists found that 11% of specialists reported successfully using a simulator as part of assessment and training procedures (Dickerson, 2013), and a meta-analysis of occupational therapy interventions found that simulator interventions were the most commonly reported and were effective for use with older adults and brain injury patients (Unsworth & Baker, 2014). Driving simulators have been effectively used in different populations, including novice drivers (Allen et al., 2012; de Winter et al., 2009), older drivers (Hoffman & McDowd, 2010; Horberry et al., 2006; Lee, Cameron, & Lee, 2003; Martin et al., 2010; Stinchcombe & Gagnon, 2013), and clinical groups including patients with cognitive impairment (Devlin et al., 2012; Frittelli et al., 2009), HIV (Vance et al., 2014), diabetes (Cox et al., 2000), sleep disorders (Smolensky et al., 2011), and brain injury (Lew et al., 2005; Schultheis et al., 2006).

Driving simulators have several advantages compared to an on-road driving assessment. Most importantly, they are safer than on-road driving, allow dangerous and unusual situations to be assessed, and provide a consistent and repeatable test environment. They also avoid the cost, space, and personnel requirements of on-road testing (Allen et al., 2003; Classen et al., 2011; Classen & Brooks, 2014). Potential clinical patients, physicians, and users agree that driving simulators are an acceptable tool for assessment, research, and training (Crisler et al., 2011; Gibbons et al., 2014; Schultheis et al., 2007). A growing body of research indicates that driving simulators provide a valid representation of on-road driving behaviour, depending on the equipment used and the situation being evaluated (Mullen et al., 2011; Shechtman, 2010). For example, driving simulator performance predicted at-fault or partially at-fault crashes in the five years following assessment (Hoffman & McDowd, 2010), and, for learner drivers, performance on a driving simulator predicted performance on an on-road assessment 6 months later (de Winter et al., 2009). Discriminant validity has been demonstrated by significant differences in the performance of non-drivers, novice drivers, and experienced drivers both on a simulator and during on-road driving (Mayhew et al., 2011). Measures of overall performance, when compared between simulator and on-road assessment, display concurrent validity across all age groups from young adults to the elderly (Engström et al., 2005; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011). Specific aspects of driving are also related for simulated driving and on-road driving; for example, Shechtman et al. (2009) demonstrated relative validity for types of driving errors made, and Kaptein et al. (1996) showed absolute validity for route choice behaviour and relative validity for speed and lateral control. Furthermore, results have indicated that lower-fidelity simulators can produce results that are comparable to high cost, high fidelity simulators (Gibbons et al., 2014; Lemieux et al., 2014).

One of the potential disadvantages of using driving simulators is the occurrence of Simulator Sickness (SS), a well-documented side effect of using a wide range of simulators and virtual reality technology (Brooks et al., 2010; Classen et al., 2011; D. M. Johnson, 2005; Kennedy et al., 1993; McCauley, 1984; Stoner et al., 2011; Trick & Caird, 2011). Overall estimated prevalence of SS varies greatly: for example McCauley (1984) reported rates of $10-84\%$, and D. M. Johnson (2005) reported rates of $0-90\%$. Of 3,691 trials on a flight simulator, 50% of all users experienced some SS (Kennedy et al., 1993). Experience of SS is related to high rates of participant dropout in driving simulator studies; Trick and Caird (2011) reported estimated dropout rates of between 35% and 75% from various institutions conducting driving simulation research with older drivers, with an average of around 40% attrition. This high dropout rate is a concern for users of driving simulators, but also poses an ethical challenge when seeking to recruit research participants due to SS being considered as a potential risk (Brooks et al., 2010).

Simulator Sickness is usually measured through specialised self-report questionnaires, such as the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993). The SSQ has been called the “gold standard” for measuring SS (D. M. Johnson, 2005). Symptoms related to SS and measured by the SSQ include general discomfort, fatigue, headache, eyestrain, difficulty focussing, increased salivation, sweating, nausea, difficulty concentrating, feelings of fullness or pressure in the head, blurred vision, dizziness, vertigo, stomach awareness, and burping. Participants respond on a four-point scale the extent to which they are experiencing each of the 16 symptoms. The 16 symptoms form three factors: oculomotor symptoms (e.g. eyestrain), disorientation symptoms (e.g. dizziness), and nausea symptoms (e.g. nausea, stomach awareness; Kennedy et al., 1993). A short form of the SSQ, the mini-SSQ, has also been used (Mourant et al., 2007). This version was developed to avoid delays involved in repeated administration, and includes only six symptoms: general discomfort, headache, blurred vision, sweating, feeling faint, and stomach discomfort. The mini-SSQ was shown to be sensitive to changes in driving conditions (Mourant et al., 2007). G. D. Park, Allen, Fiorentino, Rosenthal, and Cook (2006) reported that higher increases in SSQ score were related to dropout, with participants who dropped out of the study displaying increased SS over time, compared to non-dropouts, whose SSQ scores remained stable over time.

Factors contributing to SS can be located within three categories: Factors related to the individual, factors related to the simulator, and factors related to the simulated task (Cassavaugh et al., 2011; Kolasinski, 1995). Of these, the simulator and task specifications can be controlled to an extent, for example by using a motion base simulator which replicates the pitch and roll movements of a real car (Stoner et al., 2011), using shorter scenarios (Cassavaugh et al., 2011), avoiding turns (Mourant et al., 2007; Stoner et al., 2011), and reducing the field of view (D. M. Johnson, 2005; Kolasinski, 1995). Factors related to the individual are harder to control because they are often related to inherent characteristics of the person, such as age, gender, and medical history (D. M. Johnson, 2005). It is nonetheless important to recognise these factors so that steps can be taken to identify risk-factors and take appropriate steps to ensure SS is kept to a minimum.

Age has been identified as an important individual factor contributing to SS. Early reviews stated that SS occurs most frequently for ages 2 − 12, declines rapidly for ages 12 − 21, and continues to decline as age increases so that it is almost non-existent beyond age 50 (D. M. Johnson, 2005; Kolasinski, 1995). However, many of these earlier reports were based on flight simulation and older adults were not specifically considered. Based on more recent driving simulation reviews, it appears that older drivers represent a particularly at-risk group (Cassavaugh et al., 2011; Classen et al., 2011; Trick & Caird, 2011). For example, in a review of recent driving simulation studies, Classen et al. (2011) reported that drivers over the age of 70 are particularly at risk for SS, and Cassavaugh et al. (2011) noted dropout rates from simulation studies of up to 50% among older adult drivers. Several recent studies have reported dropout rates of between 0% and 44% for older adults (e.g. Bélanger et al., 2010; Brooks et al., 2010; Caird, Chisholm, Edwards, & Creaser, 2007; Domeyer, Cassavaugh, & Backs, 2013; C. J. Edwards, Creaser, Caird, Lamsdale, & Chisholm, 2004; Lee, Lee, Cameron, & Li-Tsang, 2003; Shanmugaratnam et al., 2010; Sklar, Boissoneault, Fillmore, & Nixon, 2014) and between 0% and 17% for younger adults (e.g. Bélanger et al., 2010; Domeyer et al., 2013; Shechtman et al., 2007; Yang, Jaeger, & Mourant, 2006); see Table 5.1 for a summary. However, estimating a reliable average dropout rate is hampered because many driving simulation studies have not reported dropout information. Additionally, dropout rates vary depending on the configuration of the simulator and the demands of the simulated task. Nonetheless, in general, results show that older adults drop out more frequently than younger adults. However, due to the small sample sizes often participating in such studies, the differences have frequently not been statistically significant.

Gender is another individual factor that is related to SS. Generally, reviews have suggested that females are more at-risk than males, especially older females (Classen et al., 2011; D. M. Johnson, 2005; Trick & Caird, 2011). Females have been reported to be more susceptible to motion sickness, SS, and visually induced motion sickness (Keshavarz & Hecht, 2014; Allen et al., 2003; Klosterhalfen et al., 2005; Mourant & Thattacherry, 2000; G. D. Park et al., 2006). Females may be particularly sensitive to simulator scenarios involving high sensory conflict and increased vection (visual illusion of self-motion) and

Table 5.1: Reported dropout rates in driving simulation studies

Study	Older adult dropout	Younger adult dropout	Group difference (Fisher's Exact Test)	Notes
Bélanger et al. (2010)	37.5% (12/32)	0% (0/20)	Significant	Age 25 – 42 vs age 65 – 83
Brooks et al. (2010)	27.8% (15/54)	6.7% (4/60)	Significant	Age 18 – 50 vs age 65 – 81
Caird et al. (2007)	34.6% (9/26)	9.3% (7/75)	Significant	Over 65 vs Under 65
Domeyer et al. (2013	30% (12/40)	12.5% (5/40)	Non-significant	Young 18 – 28; old 60 – 90. (Middle age 30 – 58, 20% (8/40)). For remaining participants, there was no difference in SS scores between age groups after accounting for baseline scores.
C. J. Edwards et al. (2004)	40% (8/20)	14% (2/14)	Non-significant	Age 65 – 83 vs. age 19 – 22
Kaber et al. (2012)	16.7% (2/12)	0% (0/10)	Non-significant	Under 25 vs Over 65
Kawano et al. (2012)	17.6% (5/17)	0% (1/15)	Non-significant	Younger adults mean age 35.2 ± 5 SD vs Over 60 (mean 66.6 ± 4.7 SD)
Lee, Lee, et al. (2003)	0% (0/129)	-	-	9% experienced SS, but 0 dropped out. Age 60+.
Park et al. (2006)	37.3% (25/67)	13.7% (7/51)	Significant	Age 21 – 50 vs age 70 – 90
Roenker et al. (2003)	11.5% (3/26)	-	-	Age 55 – 86. Participants assigned to simulator training group
Schwebel et al. (2007)	10% (10/101)	-	-	Age 75+
Shanmugaratnam et al. (2010)	17.5% (5/18)	4.5% (2/44)	Significant	Age under 40 vs age over 40
Shechtman et al. (2007)	35% (10/30)	17% (4/23)	Non-significant	Age 25 – 25 vs age 65 – 85
Sklar et al. (2014)	37.9% (22/58)	4.5% (3/67)	Significant	Age 55 – 70 vs age 25 – 35
Trick et al. (2010)	44% (15/34)	-	-	Age range not reported. Mean age of non-dropouts 70.8 years ± 5.98. Participants were judged "at-risk" after completing screening and a practice drive and were removed from the study.
Yang et al. (2006)	-	0% (0/24)	-	Novice and experienced drivers aged 16 – 45

visual flow (Jäger, Gruber, Müri, Mosimann, & Nef, 2014). Thus, females have been found to report a more severe history of motion sickness than males (Flanagan, May, & Dobie, 2005) although Mourant et al. (2007) found no gender differences in driving SS among a sample of older adults (aged $50 - 65$). Graeber and Stanney (2002) have suggested that gender differences in SS and visually induced motion sickness may be accounted for by differences in susceptibility based on individuals' prior histories of experiencing motion sickness; when males and females were balanced for susceptibility, they found no difference in self-reported sickness between genders and no difference in study duration. Significantly higher levels of sickness were instead reported in the high-susceptibility group.

Health status is related to susceptibility to SS. Many researchers have suggested that individuals who are not in their usual state of fitness do not participate in simulator studies because they are at increased risk for SS (D. M. Johnson, 2005; Kennedy et al., 1993; Kolasinski, 1995; McCauley, 1984; Stoner et al., 2011). Specific health problems related to SS include head cold, influenza, upper respiratory illness, ear infection, ear blockage, and upset stomach (Kennedy et al., 1993). Fatigue, sleep loss, recent use of alcohol or drugs, and a history of motion sickness are also risk factors for SS (D. M. Johnson, 2005; Stoner et al., 2011). Experience with the real-world task may also be related to SS. Thus, Kolasinski (1995) reported that pilots who were more experienced and had accrued more real-world flying hours were more likely to suffer from SS. No evidence relating to motor vehicle driving is currently available about this possibility but, for driving simulation with older adults, such a trend would be of particular concern because many older drivers have been driving for most of their lives and may find it difficult to adapt to the simulator. Older adults are also more likely to be experiencing health concerns or to be using medication than younger adults (Eckert et al., 2013; Gu, Dillon, & Burt, 2010).

Knowledge about susceptibility will enable high-risk individuals to be more effectively informed and monitored and, if necessary, screened out. However, one concern is that it is possible that drivers who drop out due to SS may be different in some way from those who do not suffer from SS; for example, those who drop out of simulator studies may be more impaired in their everyday lives or may be more at-risk for adverse driving outcomes. There has been some investigation into systematic differences between SS sufferers and non-sufferers in terms of cognitive performance and driving ability. Where reported, analysis has shown that in general, there are no significant differences between dropout and non-dropouts on a range of cognitive tests and on-road driving measures (Kawano et al., 2012; Mullen, Weaver, Riendeau, Morrison, & Bédard, 2010).

The aim of this study was to investigate factors related to SS and dropout[1] in a sample of older drivers on a low-cost simulator. We aimed to develop a model to identify those older drivers who are most at risk of dropout. Based on previous work reported here, it was predicted that age, gender, medical status, and mental status would be related to SS and dropout. Additionally, it was predicted that SSQ score would also be related to dropout. Participants completed a battery of cognitive and visual tests, enabling investigation of whether there would be differences between dropouts and non-dropouts in cognitive abilities and visual status. Based on the foregoing review, we expected that there would be none.

Method

Participants

We recruited 117 volunteer older drivers from the community to participate in a study investigating cognitive predictors of simulated driving performance. Participants were required to hold a current Australian driver's license, be living independently in the community, and all reported being in good general physical and mental health.

The current analysis included 88 of these participants. These had attempted all aspects of the driving task (until stage of dropout) and had complete data available on all relevant covariates. Reasons for exclusion were voluntary withdrawal from the study prior to attempting the driving task ($n = 8$), participant or experimenter decision to complete only a subset of the driving task ($n = 7$), and missing data ($n = 14$, incomplete SSQ data).

[1]After publication of this study, an anonymous reviewer has commented that the terms "Simulator Sickness" and "dropout" are used interchangeably in parts of this manuscript. In this study, dropout was exclusively the consequence of SS.

Materials

Demographics and driving information. Demographic information collected included date of birth, gender, and information about the car that they drive most often (year of manufacture, make and model, transmission). Participants also completed items relating to driving avoidance, distractions while driving, and driving confidence.

Medical conditions. Participants answered YES or NO to the following questions: Are you currently experiencing (fatigue, sleep loss, hangover, upset stomach, headache, ear conditions or ear blockage, upper respiratory illness/cold/flu)? Have you recently used alcohol/drugs (today)? Do you have epilepsy or a seizure-related condition? Do you have prior experience of motion sickness? Due to low rates of responding to each item, results were considered as two binary items: prior motion sickness (YES/NO), and presence of other medical conditions (0 conditions/1 or more conditions).

Mini Mental State Examination (MMSE). The MMSE is a short questionnaire designed to identify possible dementia or mild cognitive impairment (Folstein et al., 1975). Scores of 26 or below out of 30 indicate possible impairment. The MMSE was administered using standardized instructions and scoring (Molloy, Alemayehu, & Roberts, 1991).

Mini-SSQ. The Mini-SSQ (Mourant et al., 2007) contains six questions and is a short form of the Kennedy SSQ (Kennedy et al., 1993). The Mini-SSQ is quick to administer, suitable for repeated administration, and sensitive to changes in driving conditions (Mourant et al., 2007). Six symptoms are assessed: general discomfort, headache, blurred vision, sweating, feeling faint, and stomach discomfort. Possible responses were on a four-point scale, corresponding to None (0), Slight (1), Moderate (2), Severe (3). Scores for each symptom were summed to give a total score. The mini-SSQ was administered via tablet computer and results were automatically recorded and collated. Participants completed the questionnaire at up to three time points throughout the study: after a practice drive (SSQ1), and after each of two test drives (see below). If dropout occurred, the participant completed the questionnaire as soon as they were able to following dropout. Due to concerns of SS symptoms being suggestible (S. D. Young, Adelstein, & Ellis, 2007), the

Mini-SSQ was not administered prior to the practice drive.

Computerised Cognitive Tasks.

Useful Field of View Test™ (UFOV™) Divided Attention subtest. The UFOV™ is a computerised test of visual attention and processing speed involving detection and localisation of briefly presented targets throughout the visual field (Ball & Owsley, 1993). There are three subtests: processing speed, divided attention, and selective attention. Only the divided attention subtest (UFOV™ Subtest 2) was administered. Of the three UFOV™ subtests, Subtest 2 is most correlated with UFOV™ Total score and best predicts driving outcomes (Ball et al., 2006; Bowers et al., 2013; J. D. Edwards et al., 2006; Owsley et al., 1998). In the divided attention subtest, participants are required to identify a briefly-presented central stimulus and locate a simultaneously presented peripheral stimulus. The score is exposure time (ms) for which 75% of trials were answered correctly.

Inspection Time (IT). IT is a measure of visual processing speed (Burns & Nettelbeck, 2003). Two high-contrast lines, one markedly shorter than the other, appear as a briefly-presented target before being masked. Participants indicate whether the shorter line was located left or right of a focal point. IT was measured in ms as the duration between target onset and mask onset at which the viewer achieves 79% accuracy.

Sentence Span. A sentence span task was used to assess working memory. Task specifications are described by Lewandowsky et al. (2010). Briefly, participants were presented with a series of sentences and to-be-remembered letters. Each trial consisted of between 4 and 8 sentence/letter pairs. Participants were required to answer TRUE or FALSE to each question (e.g. "All trees are plants"); after answering, a single letter was briefly presented on screen. At the end of each trial, participants were required to enter the remembered letters in order of presentation. The outcome was the overall proportion of correctly remembered letters.

ProPerVis (Crowding subtest). ProPerVis assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns & White, 2007). There are two subtests: crowding and inspection time (see Burns & White, 2007). Only the crowding subtest was administered. The stimuli are a four-square parent figure and six figures derived from it, resembling characters M, E, W, 3, 5, 2. In each of the 40 crowding trials, one of the six figures is flanked on either side by the parent figure and appears randomly in one of five lateral positions on the screen. Participants attempt to identify which of the six figures was presented. The outcome was the total number of errors made across the five positions.

Vision tests. Participants completed tests of visual acuity and contrast sensitivity. The Freiburg Visual Acuity Test (Bach, 1996) was computerised and was completed along with the computerised cognitive tasks. Acuity was recorded as the logarithm of the Minimum Angle of Resolution (logMAR), with lower scores representing better visual acuity. Contrast sensitivity was assessed using a Pelli-Robson contrast sensitivity wall chart (Pelli et al., 1988). The outcome was log contrast sensitivity (possible range: $0-2.25$, higher scores indicate better contrast sensitivity), which was measured according to the faintest group of three letters for which the participant correctly identified two of the three letters.

Driving Simulator and Tasks. The driving simulator was custom-designed and low-cost. The setup included three 42-inch high definition LCD monitors with a 100 Hz refresh rate. The screens provided approximately 140° horizontal field of view. Participants sat on a standard desk chair and controlled the car using a small force-feedback gaming wheel and pedals. Transmission was automatic. The setup is shown in Figure 5.1.

The driving task consisted of 3 stages: a practice drive of approximately 10 minutes duration (although participants could continue to practice until they felt comfortable with the controls); Drive 1, a test drive of approximately 15 minutes, which required participants to drive a set route around suburban and city areas and respond to hazardous events; and Drive 2, similar to Drive 1, but with the addition of a Peripheral Detection Task requiring participants to respond to flashing lights appearing in their side mirrors. Responses to the Peripheral Detection Task were made using a button on the steering wheel. Emergency events occurred eight times per drive and included cars pulling out suddenly from the side of the road or intersections, and pedestrians jaywalking across the path of the car. Participants were required to brake to avoid a collision. The speed limit was 50 km/h. The drives required participants to navigate intersections (signed and

Figure 5.1: Driving simulator setup

traffic light controlled), including stopping, and the route required multiple 90-degree turns. The city area was densely populated with roadside objects including buildings, pedestrians, parked cars, signage, and other typical objects (e.g. bins, benches, trees). Driving performance is not considered here.

Procedure

The study was approved by The University of Adelaide Human Research Ethics Committee and all participants provided written consent. Participants provided demographic information and completed questionnaires relating to driving behaviour, the MMSE, computerised cognitive assessments, vision screening, and the simulated driving task. The complete protocol took approximately 2 hours to complete. The order of administration was briefing information, consent and demographic information; questionnaires relating to driving behaviour and medical conditions; MMSE; practice drive; Drive 1; cognitive and visual assessments; Drive 2.

Participants completed the SSQ up to three times throughout the study depending on stage of dropout and were monitored for signs of SS. SSQ was completed after the practice drive (SSQ1), and after each of the two test drives (SSQ2 and SSQ3). Participants were told to alert the experimenter and stop driving if they wished to discontinue for any reason. They were also monitored by the experimenter and driving was stopped if they appeared visibly uncomfortable or distressed.

Frequent breaks were provided throughout the experiment. Mandatory $5 - 10$ minute breaks were provided after the Practice Drive, Drive 1, and Drive 2. Participants were told they could request additional breaks whenever required. The computerised cognitive tasks were completed between Drive 1 and Drive 2. If a participant dropped out of the driving component of the study, they could elect to cease participation or continue with the non-driving components after a mandatory break, and if they reported that they had recovered from SS symptoms.

Results

Analysis

Data analyses were completed in R (R Core Team, 2013) using the R packages `car` (Fox & Weisberg, 2011) and `survival` (Therneau, 2014). Logistic regression analysis was used to investigate predictors of dropout at any stage. Survival analysis (Kaplan-Meier) and Cox Proportional Hazards models were used to investigate stage of dropout.

Descriptive statistics

Of the 88 participants included, 52 (29.1%) dropped out of driving during the study as follows: 26 during the practice drive, 16 during Drive 1, 8 after Drive 1, and 2 during Drive 2. Median dropout stages are shown in Table 5.2, which shows longer survival for males compared with females, and marked advantage.

Table 5.2: Median survival stages

Population	Median Survival Stage
All	Completed Drive 1
Female	During Drive 1
Male	Completed Drive 2
Prior Motion Sickness	During Practice Drive
No Prior Motion Sickness	Completed Drive 2

Covariates of interest were age, MMSE, medical conditions, prior motion sickness, and SSQ scores. Descriptive statistics for these variables are shown in Table 5.3. Participants reported the following medical conditions: fatigue ($n = 4$), sleepiness ($n = 7$), hangover ($n = 1$), upset stomach ($n = 3$), cold or influenza symptoms ($n = 5$), ear conditions ($n = 7$), recent use of alcohol or medication ($n = 3$), and prior experience of motion sickness ($n = 23$). No participant reported epilepsy or a seizure related condition. Only 14 participants reported 2 or more conditions (11 reported 2 conditions, 2 reported 3 conditions, 1 reported 4 conditions). Prior motion sickness was considered separately from other medical conditions.

Table 5.3: Descriptive statistics for covariates of interest

Variable	Mean	SD	Range
Age	72.8	5.42	65 – 87
MMSE	29.3	0.90	26 – 30
Medical Conditions	0.26	0.46	0, 1
Prior Motion Sickness	0.29	0.44	0, 1
SSQ (all)	2.45	2.66	0 – 12
SSQ1	2.74	2.78	0 – 12
SSQ2	2.66	2.59	0 – 10
SSQ3	1.42	2.27	0 – 10

Note. SSQ (all) includes all SSQ results. SSQ1, SSQ2, and SSQ3 are SSQ scores from each of the three time points.

SSQ was analysed as both a time-independent and time-dependent covariate. For the time-independent analysis the reported score on the SSQ after completion of the practice drive (SSQ1) was available for all participants and was used as a predictor of dropout. SSQ1 was considered as informative for survival risk because it is measured early on and may represent an inherent susceptibility to SS. Subsequent measurements of SSQ were found to be correlated with each other (SSQ1 − SSQ2, $r = .69$, $n = 62$, $p < .01$; SSQ1 − SSQ3, $r = .24$, $n = 38$, $p = .15$; SSQ2 − SSQ3 $r = .78$, $n = 38$, $p < .01$).

For the time-dependent analysis, SSQ scores were available for participants at up to three time points in the study, depending on stage of dropout. These represent the change in reported SSQ over the duration of the study. Age was not significantly correlated with MMSE ($r = -.07$, $p = .54$) or SSQ score (SSQ1 $r = -.04$, $p = .72$; SSQ2 $r = -.19$, $p = .14$; SSQ3 $r = -.004$, $p = .98$), and was not associated with prior motion sickness or presence of other medical conditions. There were no gender differences in age, history of motion sickness, MMSE, or SSQ scores. Compared to males, females tended to be more likely to report having one or more other medical conditions (females 41%, males 22%; $t(61.3) = 1.84$, $p = .07$). In terms of driving, males reported higher confidence in their own overall driving ability than females (females 85.5%, males 92.2%; $t(56.5) = 3.4$, $p = .002$).

Differences between dropouts and non-dropouts

Cognitive and visual performance. Dropouts and non-dropouts were compared on the cognitive and visual variables: IT, UFOV™ Subtest 2, Crowding, Sentence Span, MMSE, visual acuity, and contrast sensitivity. For IT, UFOV™ Subtest 2, crowding, visual acuity, and contrast sensitivity, two dropouts had missing data (voluntary withdrawal). For Sentence Span, two completers and six dropouts had missing data (voluntary withdrawal/technical fault/insufficient computer skills). Means and standard deviations for each variable are shown in Table 5.4. According to Levene's Test, variances were equal for all variables. Independent samples t-tests found no significant differences between the groups on any of the variables, although there was a weak trend with a small-to-medium effect size ($p = .08$, Cohen's $D = 0.40$) for IT favouring completers. This effect was moderated by gender; for females, mean IT for completers and dropouts was 70ms, but for males, mean IT for completers was 63ms, and mean IT for dropouts was 82ms ($t(30.2) = 1.96$, $p = .059$).

Medical conditions and motion sickness. Dropouts were significantly more likely to report a history of motion sickness, $\chi^2(1) = 10.00$, $p = .002$. Of the 23 participants who reported prior motion sickness, 20 dropped out. There was no association between medical conditions (excluding motion sickness) and dropout, $\chi^2(1) = 1.57$, $p = .21$.

Table 5.4: Comparison between dropouts and completers on cognitive and visual measures and SSQ scores

	M (SD)				
Variable	Dropouts	Completers	t (df)	p	Cohen's D
MMSE	29.3 (1.02)	29.5 (0.70)	1.08 (86)	.28	0.22
IT	75.7 (35.8)	64.2 (18.7)	1.75 (84)	.08	0.40
UFOV™ Subtest 2	79.4 (72.3)	79.5 (63.9)	0.01 (84)	.99	0.00
Crowding	11.8 (4.91)	11.06 (4.71)	0.74 (84)	.46	0.16
Sentence Span	0.51 (0.17)	0.54 (0.15)	0.65 (78)	.52	0.19
Visual Acuity	0.11 (0.13)	0.11 (0.13)	0.14 (84)	.89	0.00
Contrast Sensitivity	1.80 (0.17)	1.83 (0.15)	0.81 (84)	.42	0.19
SSQ1	4.12 (2.78)	0.75 (1.03)	7.97 (69.1)	$< .001$	1.61
SSQ2	4.73 (2.33)	1.17 (1.54)	7.26 (60)	$< .001$	1.80
SSQ3	5.50 (0.71)	1.19 (2.11)	2.85 (36)	.01	2.74

Note. SSQ1 = SSQ Time 1, SSQ2 = SSQ Time 2, SSQ3 = SSQ Time 3

Gender. Females were significantly more likely to drop out than males, $\chi^2(1) = 6.92$, $p = .009$. Of the 56 male participants, 28 dropped out (50%). Of the 34 female participants, 26 dropped out (76%).

Age. There was no age difference between dropouts ($M = 72.58$ years) and completers ($M = 73.17$ years); $t(86) = 0.50$, $p = .62$.

SSQ. As expected, there were significant differences in SSQ scores for dropouts and completers (see Table 5.4) with dropouts reporting much higher SS symptoms than completers.

Logistic Regression: Dropout

Fifty-two participants dropped out and 36 participants completed all driving. Binary logistic regression analysis was used to determine if age, gender, prior motion sickness, medical conditions, MMSE, and SSQ were predictive of dropout (see Table 5.5). The dependent variable was dropout, considered as a binary variable (YES or NO for dropout at any stage of the study). Age, gender, motion sickness, medical conditions, and MMSE were entered into the model first. These are characteristics of the driver that are available prior to any simulated driving being attempted. This model was significant, $\chi^2(5) = 20.70$, $p < .001$, with pseudo-R^2 (Nagelkerke) of .283. Gender and prior motion sickness were significant predictors of dropout. SSQ1 score was added to the model in Step 2 and significantly improved prediction of dropout, $\chi^2(1) = 39.27$, $p < .001$ with a pseudo-R^2 (Nagelkerke) of .666. The model correctly classified 86.4% of cases. Motion sickness and SSQ1 score were significant predictors. Gender was no longer significant. For every 1 point increase in SSQ1 score, participants were 2.84 times more likely to drop out. Participants with prior history of motion sickness were 8 times more likely to dropout than those without.

Table 5.5: Logistic regression analysis for dropout

	B	SE *B*	OR	p
Step 1				
Age	-0.06	0.05	0.95	.25
Gender (Baseline: Male)	1.20	0.53	3.32	.02
MMSE	-0.42	0.30	0.65	.15
Prior Motion Sickness	2.03	0.72	7.61	.01
Other Conditions	0.48	0.56	1.62	.39
(Constant)	16.00	9.90		.11
Step 2				
Age	-0.05	0.06	0.95	.40
Gender (Baseline: Male)	0.66	0.73	1.94	.36
MMSE	-0.26	0.35	0.77	.46
Prior Motion Sickness	2.09	0.88	8.08	.02
Other Conditions	1.21	0.84	3.36	.15
SSQ1	1.04	0.26	2.84	$< .001$
(Constant)	8.44	11.78		.45

Note. Model 1 $\chi^2(5) = 20.70$, $p < .001$. Model 2 $\chi^2(6) = 59.97$. OR = Odds Ratio.

Survival analysis: Stage of dropout

There were 88 participants (54 male) with complete data available. Descriptive statistics for variables of interest are shown in Table 5.3. Survival analysis and Cox Proportional Hazards models were used to investigate dropout. Dropout stage was coded was coded as five stages: Stage 1 (Practice Drive), Stage 2 (during Drive 1), Stage 3 (after Drive 1), Stage 4 (during Drive 2), and Stage 5 (no dropout; completed Drive 2). The event of interest was dropout or withdrawal from the driving task. The stage that each participant dropped out or withdrew was recorded. Participants who completed all driving were considered censored at Stage 5 (that is, for the purposes of the survival analysis, the outcome of interest (dropout) did not occur during the observation period).

Participants tended to dropout early in the study rather than later; 29.5% of participants dropped out at stage 1, half of all dropouts. Fewer than half of all participants (40.9%) completed all driving. Median survival stages for all participants and subgroups based on gender and medical conditions are shown in Table 5.2. These median survival stages suggest that female gender and having one or more medical condition increases the risk of dropout.

Kaplan-Meier survival analysis. Figure 5.2 shows the Kaplan-Meier survival plot for all participants. There were 88 participants who had complete data available. Of these 88, 36 (41%) participants were censored at stage 5 (that is, they completed the study without drop out). The median survival stage was stage 3.

Figure 5.3 shows the Kaplan-Meier survival plot by gender. Males generally survived longer than females, with 49% of males surviving until stage 5, compared to 28.2% of females. The difference in survival plots was significant according to the log-rank test, $\chi^2(1) = 9.00$, $p = .002$. The median survival stage for females was 2, and the median survival stage for males was 5.

Figure 5.4 shows Kaplan-Meier survival plots by history of motion sickness. Participants with no history of motion sickness survived longer than those with a history of motion sickness. The plots were significantly different according to the log-rank test, $\chi^2(1) = 11.2$, $p < 001$. Median survival stage for participants with a history of motion sickness was 1, and the median survival stage for participants with no history of motion sickness was 5.

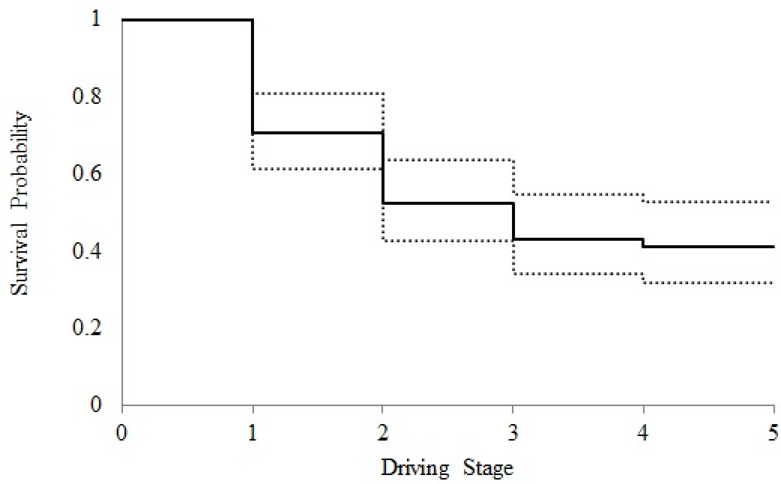

Figure 5.2: Kaplan-Meier survival plot for all participants. The dotted line shows the 95% confidence interval.

Cox Proportional Hazards model: Time independent analysis. Cox Proportional Hazards models were used to investigate the effects on individual predictors on dropout. Gender, motion sickness, other conditions, age, MMSE, and SSQ1 score were included. Analyses showed there were no significant interactions between these factors. Two models were calculated; the first included gender, motion sickness, other conditions, age, and MMSE, and the second added SSQ1 to these predictors. Both models are shown in Table 5.6. The log-likelihood test showed that Model 2 was a significant improvement on Model 1, $\chi^2(1) = 32.3$, $p < .001$. Gender was a significant predictor in Model 1, but was reduced to marginal significance in Model 2. Motion sickness was a significant predictor in both models, while other conditions, age and MMSE were not significant in either model. In Model 2, SSQ1 was a highly significant predictor, and the significance and HRs of gender and conditions decreased slightly. Prior motion sickness and SSQ1 were the most useful predictors of dropout. People with a history of motion sickness had 106% increased risk of dropout compared to people with no prior history of motion sickness. For each point increase in SSQ1 score, hazard rate increased by 35%.

Overall, the models met the assumption of proportional hazards, although age was

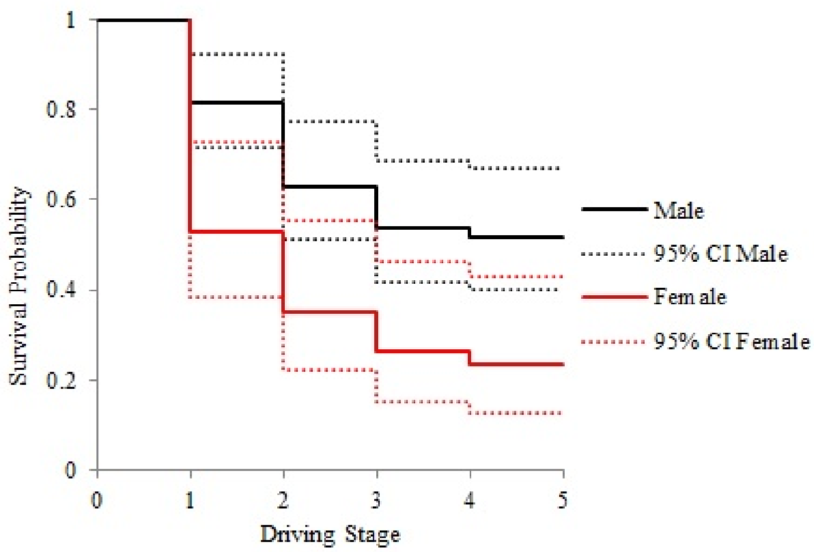

Figure 5.3: Kaplan-Meier survival plots by gender

shown to have non-proportional hazards.

Cox Proportional Hazards model: Time dependent analysis. It was also possible to consider SSQ score a time-dependent covariate. SSQ was completed up to three times by participants: after the practice drive, after drive 1, after drive 2, and/or after dropout if applicable. Therefore there were 5 possible stages for SSQ data to be collected: after practice drive (1), after dropout during drive 1 (2), after completion of drive 1 (3), after dropout during drive 2 (4), and after completion of drive 2 (5). SSQ scores were collected up to three times per participant. For example, a participant who dropped out during drive 1 would have two SSQ scores available, at stage 1 and stage 2. As another example, for participants who completed all driving, SSQ scores were collected at stages 1, 3, and 5. For these cases, missing data were filled by duplicating the next occurring rating (i.e. stage 2 was equal to stage 3; this is because ratings at stage 3 were assumed to reflect level of SS throughout the duration of the drive).

The Cox Proportional Hazards model in Table 5.7 included SSQ as a time dependent covariate. The log-likelihood test showed that the model was significant, $\chi^2(6) = 87.0$, $p < .001$. SSQ was a highly significant predictor of dropout. For every 1 point increase in SSQ score, participants had a 55% increased risk of dropout during that stage of the driving task. Gender remained a significant predictor of dropout, with females

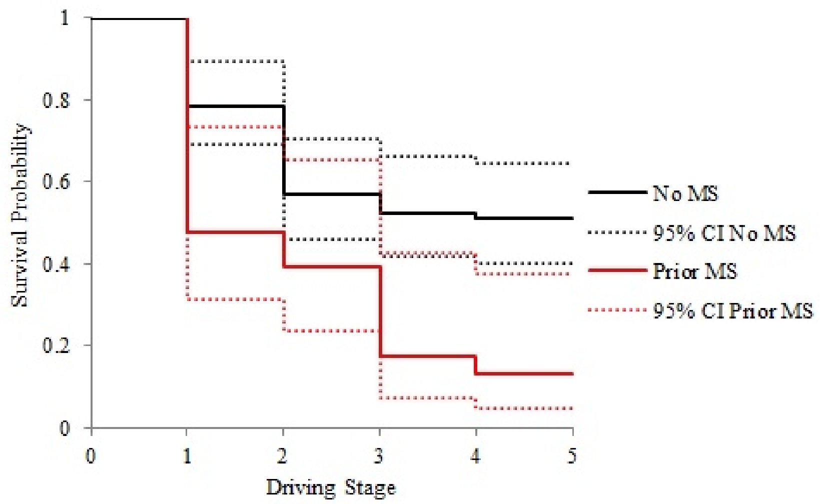

Figure 5.4: Kaplan-Meier survival plot by prior motion sickness

Table 5.6: Cox Proportional Hazards model for time to driving dropout

	HR	CI	p
Model 1			
Gender	1.94	1.07 – 3.51	.03
Motion Sickness	2.03	1.15 – 3.59	.02
Other Conditions	1.07	0.58 – 1.98	.82
Age	0.99	0.94 – 1.04	.66
MMSE	0.89	0.67 – 1.19	.43
Model 2			
Gender	1.71	0.95 – 3.06	.07
Motion Sickness	2.06	1.16 – 3.64	.01
Conditions	1.15	0.64 – 2.09	.64
Age	1.01	0.96 – 1.07	.64
MMSE	0.90	0.66 – 1.22	.50
SSQ1	1.35	1.22 – 1.50	$< .001$

Note. HR = Hazard Ratio, CI = 95% Confidence Interval

Table 5.7: Cox Proportional Hazards model for stage of study dropout, with SSQ as a time-dependent covariate

Variable	HR	CI	p
Gender	2.02	1.12 – 3.65	.02
Motion Sickness	2.22	1.24 – 3.97	.01
Other Conditions	0.71	0.38 – 1.34	.30
SSQ	1.55	1.40 – 1.72	$< .001$
Age	1.04	0.98 – 1.10	.20
MMSE	1.06	0.77 – 1.47	.72

Note. HR = Hazard Rate, CI = 95% Confidence Interval

being twice as likely to drop out compared to males. History of motion sickness was associated with a hazard ratio of 2.22. Other medical conditions, age, and MMSE did not significantly affect risk of dropout in the time-dependent model. The survival plot for this model is shown in Figure 5.5.

Discussion

The aim of this study was to investigate factors related to SS and dropout in a sample of older drivers on a low-cost simulator. We aimed to develop a model to identify those older drivers who are most at risk of dropout. Based on previous work reported here, it was predicted that age, gender, medical status, and mental status would be related to SS and dropout. Additionally, it was predicted that SSQ score would also be related to dropout. Participants completed a battery of cognitive and visual tests, enabling investigation of whether there would be differences between dropouts and non-dropouts in cognitive abilities and visual status. Based on the foregoing review, we expected that there would be none.

We investigated individual predictors of dropout and SS among older adults using a custom-designed, low-cost driving simulator. It was predicted that age, gender, and medical history would be related to SS and dropout, and that SS would be predictive

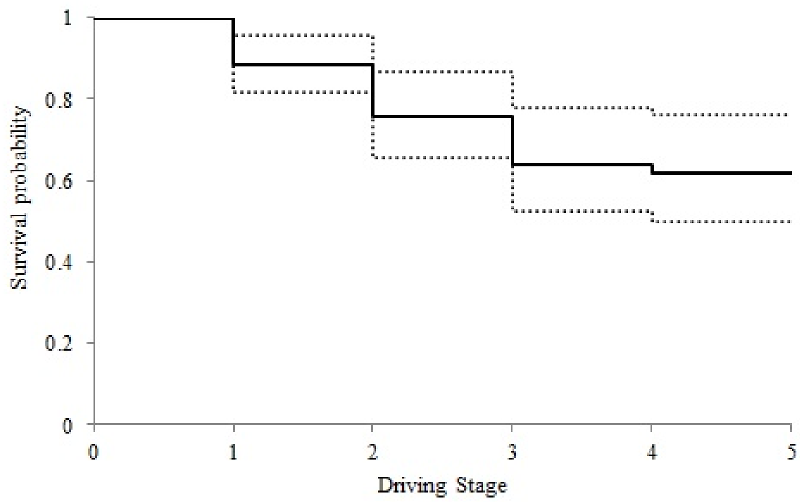

Figure 5.5: Cox survival plot displaying stage of dropout, controlling for gender, prior motion sickness, medical conditions, SSQ score, age, and MMSE. The dotted lines show the 95% confidence interval.

of stage of dropout. Using logistic regression and survival analysis models, the results showed that SSQ score was the best predictor of dropout. In the logistic regression analysis, adding SSQ1 score to the demographic predictors increased the pseudo-R^2 by 40%. This confirms that self-reported SS symptoms are related to dropout (G. D. Park et al., 2006). SSQ scores were considerably higher throughout the study for dropouts. On average, those completing all stages of driving reported a score of 1 on the SSQ throughout the study (equivalent to "slight" experience of one symptom), while dropouts reported scores above 4 at each time point (equivalent to slight, moderate or severe experience of two or more symptoms). Higher SSQ scores were associated with a higher hazard ratio in both the time-independent and time-dependent Cox Proportional Hazards models. In the time dependent analysis in particular, each point increase in SSQ score was associated with 55% increase in risk of dropout. This confirms the importance of monitoring SS symptoms throughout the study.

Of the covariates that would be available before study commencement, gender and history of motion sickness significantly predicted risk of dropout. Females dropped out more frequently than males, as reflected in the median survival stages, with females having a median survival stage corresponding to part way through Drive 1, while males on average completed all driving. The Cox Proportional Hazards models showed that females were at twice the risk of dropout as males. This corresponds with numerous reviews that have found that females are more susceptible to visually-induced motion sickness and SS (Allen et al., 2003; Keshavarz & Hecht, 2014; Mourant & Thattacherry, 2000; G. D. Park et al., 2006). Jäger et al. (2014) reported that females were more affected by scenes involving high sensory conflict and visual flow than males. This may have been related to the significantly higher dropout of females in the present study; the scene was very visually complex and involved numerous turns associated with high visual flow. It has been reported that females are also more likely than males to suffer from motion sickness in any form (Flanagan et al., 2005), but, in the present study, there was no reported gender difference on prior motion sickness (our study used a single item to assess motion sickness history; more comprehensive scales are more typical). There were also no gender differences in reported medical conditions, age, MMSE, or any of the cognitive or visual variables.

Prior experience of motion sickness significantly increased the risk of dropout. Thus, of 23 participants who reported a history of motion sickness, 20 dropped out during the study. As reflected by the median survival stages, participants with a history of motion sickness tended to drop out after attempting the practice drive, while participants with no history of motion sickness tended to complete the study. The Cox Proportional Hazards models showed that a history of motion sickness was associated with more than twice the increased risk of dropout than no prior history of motion sickness. As suggested by D. M. Johnson (2005), past behaviour is often the best predictor of future behaviour, and that was certainly the case in the present study.

Contrary to our expectation, presence of other medical conditions was not related to dropout. Drivers were required to be in good general health to participate in the study, and overall incidence of health conditions was low. Thus, self-selection and the screening procedure likely contributed to this outcome.

Although age was not related to dropout in this sample, the overall dropout rate in this study was very high (59%). A separate study was run with younger participants, undergraduate students aged $18 - 30$, using the same simulator and a very similar simulated task (Tan, 2014). Of 66 participants, only 2 dropped out (3%). Both were female and had prior experience of motion sickness but were otherwise in their usual state of good health. This dropout rate was significantly different from that in the present study, $\chi^2(1) = 52.1, p < .001$. This comparison adds strong evidence to the observations that older adults are a high-risk group for SS (Cassavaugh et al., 2011; Classen et al., 2011; Trick & Caird, 2011) and dropout rates between younger and older drivers (Brooks et al., 2010; Caird et al., 2007; G. D. Park et al., 2006; Sklar et al., 2014).

The 59% dropout observed in the current study was higher than has generally been previously reported[2]. For example Trick and Caird (2011) reported a dropout rate of 44%, and Cassavaugh et al. (2011) noted dropout rates of up to 50%, while most other studies have reported dropout rates between 10% and 40% (Bélanger et al., 2010; Brooks et al., 2010; Kaber, Zhang, Jin, Mosaly, & Garner, 2012; Kawano et al., 2012; Roenker et al., 2003). The high dropout rate here was likely a result of the simulator configuration and task demands of the current study. It is well known that simulator type and aspects of the simulated environment and task requirements are related to SS (Cassavaugh et al., 2011; Kolasinski, 1995). Our simulator contained a number of features that can increase SS, for example a wide field of view (Kolasinski, 1995; Stoner et al., 2011), turning several corners (Cassavaugh et al., 2011; C. J. Edwards et al., 2004; Mourant et al., 2007), a visually complex scene (G. D. Park et al., 2006), relatively long scenario duration (Cassavaugh et al., 2011; D. M. Johnson, 2005; Kolasinski, 1995), and elements of the scene prone to flicker (Kolasinski, 1995; Stoner et al., 2011). The control system was also prone to lag (delay between input and response) and was poorly calibrated (poor correspondence between degree of input and reaction of the vehicle, especially for steering; Classen et al., 2011; Kolasinski, 1995; Stoner et al., 2011). Moreover, many participants reported dissatisfaction with the gaming controls, expressing concerns that the wheel was too small and did not feel realistic, and unfavourable opinions have been related to increased SS (Schultheis et al., 2007). Modifications to the scenario, scene, and simulator setup can help to reduce SS (Cassavaugh et al., 2011), as can thorough screening procedures (Trick

[2]After publication of this study, an anonymous reviewer has commented that the high dropout rate might have been caused by the inclusion of drivers with a history of motion sickness. As mentioned later in the discussion, it is recommended that drivers with a history of motion sickness be excluded from participation. This recommendation was implemented for the remaining studies in this program of research, and SS rates dropped accordingly.

& Caird, 2011) and use of adaptation procedures (Domeyer et al., 2013).

It has been suggested that drivers with a history of motion sickness be excluded from simulator studies (Brooks et al., 2010; Stoner et al., 2011); and such screening procedures have been shown to reduce cases of SS (Trick, Toxopeus, & Wilson, 2010). It should be noted, however, that doing so would tend to introduce a male bias into the sampling of driving behaviour. Nonetheless, our results showed that there were no differences between dropouts and those who completed all stages on any of the cognitive or visual measures. These included tests of visual processing and attention, including the UFOV™, a well-established predictor of driving outcomes (Clay et al., 2005). This result suggests that older drivers who were unable to complete the simulated driving task were not more impaired than those who did. Results were remarkably uniform across the two groups. Similar findings have been reported by Kawano et al. (2012) and Bélanger et al. (2010). Mullen et al. (2010) compared dropouts and non-dropouts on an on-road driving test and found that those who dropped out actually committed fewer on-road driving errors than those who did not drop out. Evidence therefore suggests that people who drop out of simulator studies are not more impaired or at-risk than those who do not drop out. D. M. Johnson (2005) also reported that SS has little to no effect on cognitive and perceptual abilities; this was confirmed in the present study, where those who dropped out performed the cognitive testing after discontinuing driving, and performed equivalently to those who did not experience SS. Therefore, these results tend to suggest that it may be possible to screen out individuals who report prior motion sickness without biasing results from cognitive testing.

We used the Mini-SSQ, a much briefer version of the more frequently-used, "gold standard" SSQ (D. M. Johnson, 2005; Kennedy et al., 1993). The Mini-SSQ was developed to save time when repeated administration and monitoring is required (Mourant et al., 2007). Mourant et al. (2007) reported that the Mini-SSQ was sensitive to changes in driving environment, such as increased scene complexity and increased task demands. Our results show that the Mini-SSQ is sensitive to changes in SS over time and is able to accurately identify drivers who may be at increased risk for dropout. The Mini-SSQ was quick and easy to administer and was well accepted by the participants. It therefore appears to be appropriate for quickly and accurately monitoring SS symptoms over the course of a study, in situations where the more complete symptom breakdown of the full

SSQ is not required.

Overall, the results show that females and people with a history of motion sickness had a significantly increased risk of dropout from the study. Although age did not predict dropout within the sample, evidence from a study with younger participants using the same equipment has suggested that older adults in general are a high-risk group for SS. As expected, experience of SS symptoms was related to risk of dropout. The time-dependent analysis of SS confirmed the importance of monitoring symptoms throughout the study, because changes in SS affected risk of dropout. The study has highlighted the importance of thorough screening procedures and effective monitoring of participants. The results also showed that participants who dropped out were not cognitively impaired compared to those who did not dropout, and, given that cognitive testing was completed post-dropout, the results suggest that SS does not have a negative impact on cognitive performance.

Chapter 6.

Study 3: Assessment of driving simulator validity and acceptability for older adult drivers

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2016). Assessment of driving simulator validity and acceptability for older adult drivers. *Advances in Transportation Studies, 39*, 69-98.

Statement of Authorship

Title of paper	Assessment of driving simulator validity and acceptability for older adult drivers
Publication status	Published
Publication details	Matas, N. A., Nettelbeck, T., & Burns, N. R. (2016). Assessment of driving simulator validity and acceptability for older adult drivers. *Advances in Transportation Studies, 39*, 69-98.

Principal Author

By signing the Statement of Authorship, I certify that this paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper

Name of principal author	Nicole Amy Matas		
Contribution to paper	Study concept and design, data collection, statistical analyses, writing and submitting manuscript, addressing referee comments		
Overall percentage	80%		
Signature		Date	13/9/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ted Nettelbeck		
Contribution to paper	Principal supervision, advice about development of concept and design of study, manuscript revision and approval		
Signature		Date	13. 9. 16

Name of Co-Author	Nick Burns		
Contribution to paper	Advice about development of concept and design of study, assistance with statistical analyses, manuscript revision and approval, ongoing		
Signature		Date	14/9/16

Abstract

Driving simulators are now widely available and frequently used for research, training, and assessment. For driving simulators to be a useful tool they must be valid, reliable, and acceptable to users. Driving simulator validity depends on the simulator used, the particular scenarios selected, and the population under consideration. We investigated the validity, reliability, and user acceptability of a fixed-base three-screen driving simulator in a sample of older adult drivers (N = 26, mean age = 73 years, SD = 5.2 years). Participants completed four tasks in the driving simulator: 1. A steering practice scenario (Task 1: Braking Practice), 2. A brake reaction time test (Task 2: Brake Reaction Time (RT)), 3. A distracted driving assessment (Task 3: Distracted Driving), 4. Driving in city traffic (Task 4: City Traffic Participation). Participants provided feedback on each scenario and the driving simulator in general. Participants also completed questionnaires related to their driving habits and history, their health status, and their familiarity with computers and technology. Results showed that participants had generally positive opinions towards the simulator and scenarios; that the distracted driving task and the city driving task demonstrated evidence of content and convergent validity respectively; and that the Brake RT task had excellent test-retest reliability. Overall, the simulator and tasks appear to be appropriate for assessing driving performance in older adults.

Introduction

Driving simulators are now widely available and frequently used for research, training, and assessment (Allen et al., 2011; Classen & Brooks, 2014; Crisler et al., 2011; Dickerson et al., 2014). The availability of lower-cost options means that driving simulators are now increasingly accessible to researchers and results have indicated that, in certain situations, lower-fidelity simulators can produce results that are comparable to high cost, high fidelity simulators (Gibbons et al., 2014; Lemieux et al., 2014). For driving simulators to be useful, they must be a valid measure of driving behaviour; that is, performance in the simulator must accurately reflect on-road driving performance. The gold standard for validation of driving simulators is to compare measures of performance on a simulator with identical measures of performance on-road, where conditions in the simulator match those in the real environment (Shechtman, 2010). Studies using this method have been successful in establishing correspondence between simulator measures and on-road measures of driving performance (Engström et al., 2005; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011). However, the time, cost, and specialised equipment required for this method make it impractical in many cases.

Fortunately, there are other methods that can be used to establish the validity of a simulator. In general, driving simulators as measures of on-road driving ability have demonstrated acceptable validity across several modes of validation (Mullen et al., 2011; Shechtman, 2010). For example, driving simulator performance predicted at-fault or partially at-fault crashes in the five years following assessment (Hoffman & McDowd, 2010) and, for learner drivers, performance on a driving simulator predicted performance on an on-road assessment 6 months later (de Winter et al., 2009). Discriminant validity has been demonstrated by statistically significant differences in the performance of non-drivers, novice drivers, and experienced drivers both on a simulator and during on-road driving (Mayhew et al., 2011). Measures of overall performance display concurrent validity (between simulator and on-road measures) across age groups from young adults to the elderly (Engström et al., 2005; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011). Specific aspects of driving are related for simulated driving and on-road driving; for example Shechtman et al. (2009) demonstrated relative validity for types of driving errors made; and Kaptein et al. (1996) showed absolute validity for route choice behaviour and relative validity for speed and lateral control. Bédard et al. (2010) showed that simulated driving performance was related to cognitive ability, with measures of visual attention being related to simulator-recorded errors.

Simulator validity is highly dependent on the specific simulator, task, and population under consideration (Kaptein et al., 1996; Mullen et al., 2011; Shechtman, 2010). In the present study, we considered the reliability and validity of a three-screen fixed-base driving simulator. We aimed to assess the validity of the simulator by comparing drivers' self-reported fitness-to-drive, crash history, and recent infringements with their performance on the simulated driving tasks. We also compared driver age and cognitive ability with performance on the simulator, and compared performance on the simulated tasks with known real-world behaviours.

A concern with using driving simulators, particularly among older adult drivers, is Simulator Sickness (SS). SS is a set of motion sickness-like symptoms that can occur with use of simulators and virtual reality technology (Brooks et al., 2010; Classen et al., 2011; D. M. Johnson, 2005; Kennedy et al., 1993; Stoner et al., 2011). Due to SS symptoms, some users are unable to continue using the simulator and are thus unable to complete the intended assessment tasks. Estimated prevalence of SS varies greatly; for example, McCauley (1984) reported rates of $10-84\%$, and D. M. Johnson (2005) reported rates of $0-90\%$. In terms of dropout due to SS, Trick and Caird (2011) reported estimated dropout rates of between 35% and 75% from various institutions conducting driving simulation research with older drivers, with an average of around 40% attrition. There are a number of risk factors for SS, including female gender (Classen et al., 2011; D. M. Johnson, 2005; Trick & Caird, 2011), older age (Cassavaugh et al., 2011; Classen et al., 2011), and motion sickness history (D. M. Johnson, 2005; Lee, Cameron, & Lee, 2003). In the present study, we investigated rates of SS and risk factors for SS; in addition to demographic data, we also considered health (D. M. Johnson, 2005; Kennedy et al., 1993; Stoner et al., 2011), cognitive status (Freund & Green, 2006), simulator experience, driving experience, computer and video game use, and attitudes toward computers as potential risk factors for SS.

Finally, we investigated the usability and user acceptability of the simulator. This is a face validity issue; for the simulator to be a useful research tool, participants must take the assessment seriously and they must perceive the driving simulator and tasks to be acceptable and relevant. Usability is an important aspect of any new technology and relates to ease of use, user acceptance, user enjoyment, and accuracy of task performance (Schultheis et al., 2007). A usable technology should provide an accurate reflection of the behaviour of interest. Few studies have investigated whether older adults perceive driving simulators as being relevant for the assessment of their driving (Gibbons et al., 2014). In one such study, Schultheis et al. (2007) used a user feedback questionnaire to examine the usability of a Virtual Reality Driving Simulator with healthy adults (age range $21 - 64$) and acquired brain injury patients (age range $20 - 68$) and found that the simulator was generally acceptable to participants. They also found that user feedback ratings were related to onset of SS; people who rated the technology less favourably were more likely to experience SS. Additionally, older age was associated with less favourable feedback. Consistent with this, Gibbons et al. (2014) used a semi-structured interview to obtain feedback from older and middle-aged drivers; middle-aged drivers responded more positively than older drivers, but most participants agreed that the simulator could be useful for evaluation of their driving skills, training and teaching. In the present study we used a User Feedback Questionnaire derived from that of Schultheis et al. (2007) to determine if our driving simulator is acceptable to participants, and if usability ratings are related to SS. User feedback also indicated the face validity of the simulator in this population.

To summarise, the aim of the present study was to investigate the usability, validity, and acceptability of a fixed-base three-screen driving simulator for use with older adult drivers. Specifically, the aims of the study were to:

- Assess the test-retest reliability of the simulator, using a Brake RT assessment to provide evidence for the reliability of measurement of this test.

- Investigate the validity of the simulator by comparing individuals' self-reported driving habits, including accident and infringement history, with their driving performance on the simulator. If the simulator is a valid reflection of driving behaviour, we expect that simulated driving performance should be positively associated with better on-road driving habits and history. Additionally, we examined the relationship between simulated driving performance, age and cognitive ability. Again, relationships in the expected direction will provide evidence for the convergent validity of the simulator.

- Compare the effects of distraction of performance in our simulator with the effects of distraction from other simulator studies and on-road studies (content validity). If the effects in the simulator correspond to the literature, this provides evidence for the content validity of the distracted driving task.

- Investigate the usability and acceptability of the simulator and scenarios by obtaining user feedback. This indicates the face validity of the simulator and each of the selected scenarios.

- Report the extent of SS experienced by users of the simulator, and determine predictors of SS.

Method

Participants

Twenty-six participants (16 male, 10 female) were recruited from a database of volunteers who had expressed interest in participating in research related to older drives. Of these, 18 reported prior experience with simulators (10 male, 8 female), 15 (7 male, 8 female) of whom had previously participated in simulator research in our laboratory. All participants were aged over 65 years, held a current Australian driver's license and drove at least once per week, were in good general health, had normal or corrected-to-normal vision, and had not been diagnosed with dementia or cognitive impairment. The mean age of participants was 73 years (SD 5.2 years, age range $66 - 84$ years).

Materials

Pre-test questionnaire. Participants completed an online questionnaire prior to coming to the laboratory. The questionnaire contained items for demographic information, health and wellbeing, use of computers and simulators, and driving history and habits.

Demographic information. Gender, date of birth, level of education (qualifications ranging from "no formal education" to "doctorate degree"), and employment status.

Health and wellbeing. Self-rated health, current state of health, medical conditions, medications, and vision and eye conditions.

Simulator history. Prior experience with any type of simulator, and estimated hours of use of simulators.

Computer use. Estimated hours per day spent using a computer, hours per week spent playing video games, and genres of video games played.

Driving History. Year license first obtained, hours driven per day, kilometres driven per year, and special license conditions.

Motion Sickness Susceptibility Questionnaire (MSSQ). The MSSQ (Golding, 2006) predicts susceptibility to motion sickness based on reactions to various stimuli, including motion sickness experienced in cars, buses, small boats, etc. We used the structure and scoring developed by Golding (2006). Maximum score was 27, with higher scores indicating higher motion sickness susceptibility.

Attitudes Towards Computers Questionnaire (ATCQ). The ATCQ (Jay & Willis, 1992) contains items assessing attitudes toward computers. The questionnaire has been validated for use with older adults (Jay & Willis, 1992). We used the Comfort (feelings of comfort with computers) and Efficacy (feelings of competence with computers) subscales. Both subscales contained five items answered on a 5-point scale from "strongly disagree" to "strongly agree". Scores were calculated by summing responses to each item. Cronbach's α in the present sample was .92 for the Comfort Scale, .82 for the Efficacy Scale, and .88 for the Total Score.

Driving self-assessment questionnaire. The self-assessment questionnaire allows older drivers to reflect on their own driving performance across a number of areas. The items were adapted from a resource available from the Government of South Australia. Questions relate to driving safety behaviours, awareness of health concerns, awareness of road rules, and confidence and attitudes towards driving. Responses from 13 items were summed to give a total score. Maximum possible score was 52, with higher scores indicating a better self-assessment. Participants reported how many tickets, infringements, and warnings they had incurred over the past two years, and how many accidents they had been involved in over the past two years[1]

Mini-Simulator Sickness Questionnaire (SSQ). The Mini-SSQ (Mourant et al., 2007) contains six questions and is a short form of the Kennedy SSQ (Kennedy et al., 1993). The Mini-SSQ is quick to administer, suitable for repeated administration, and sensitive to changes in driving conditions and changes in SS symptoms over time (Matas, Nettelbeck, & Burns, 2015; Mourant et al., 2007). Six symptoms are included: general physical discomfort, headache, blurred vision, sweating, faintness/dizziness, and stomach discomfort/nausea. However, due to a transcription error, "sweating" was omitted from the list of symptoms, and therefore only five symptoms were assessed. Possible responses were on a four-point scale, corresponding to None (0), Slight (1), Moderate (2), Severe (3). Scores for each symptom were summed to give a total score. Participants completed the questionnaire after each of the four driving tasks (or as soon as possible after dropout). Five SSQ measures were calculated: Mini-SSQ score after each driving task (SSQ1, SSQ2, SSQ3, SSQ4) and the Maximum SSQ score recorded by each individual (SSQ Max).

Control Feedback. At the same time as the SSQ, participants completed a brief set of questions relating to how confident they were controlling the simulator at that point. Responses were on a 5-point scale from Strongly Disagree to Strongly Agree. The items assessed comfort with: following instructions provided by the simulator, controlling the brake, controlling the steering, controlling the accelerator, changing gears, and overall ability to control the vehicle.

Simulator and scenario feedback. Participants provided feedback relating to their overall experience with the simulator, and feedback relating to each task. Items in

[1]After publication of this study, an anonymous review has commented that the scoring of the self-assessment questionnaire requires clarification. Participants responded to 13 items. Possible responses were Never/almost never (0), Rarely (1), About half the time (2), Often (3), Always/Almost Always (4). Maximum score was 52, with a higher score representing better self-assessment of driving. Example items include "Do you signal in plenty of time and check for cars behind and beside you when you change lanes?" and "Do you feel you are reacting to dangerous driving situations later than you used to?" (reverse scored). Participants also self-reported number of infringements and crashed over the past two years; these two items were reported separately and were not included in the total self-assessment score.

the simulator feedback questionnaire were based on the questionnaire used by Schultheis et al. (2007). The questions related to various aspects of the simulator, including ability to use the controls, how the scenarios were presented, and the content of the tasks. Responses were made on a 5-point scale from Strongly Disagree (0) to Strongly Agree (4). Responses from the 17 items were summed to give a total Simulator Feedback Score (maximum 68). Participants were also asked how they felt about each task, including how challenging they perceived the task to be, the length of the task, confidence in controlling the vehicle during the task, whether the task made them uncomfortable or stressed, and whether the task was a relevant assessment of their driving skills.

Driving simulator and tasks. The driving simulator used was a SimWorx SX06DTS-L Driver Training (Figure 6.1). The simulator has a fixed base and is located in the laboratory. The simulator has a three screen display (three × 27in HD monitors, total resolution 5760 × 1080 pixels). The simulated horizontal field of view spans a total visual angle of 210°, but this can be accessed in its entirety only via head movement, or eye movement, or both. Thus, it is analogous to viewing through the front and side windows of a motor vehicle. The graphics were texture-rich. All scenarios required driving on the left hand side of the road, consistent with Australian road rules. The simulator has a driver cockpit including force feedback steering wheel, pedals, gearshift, handbrake, ignition switch, engine start button, and adjustable seat. The steering is controlled by a Logitech G27 unit modified with a steering wheel of 350mm diameter. The steering wheel and pedals were calibrated with the Logitech Profiler and the configuration menu on the simulator. The vehicle dynamics model is linear and computes the roll, pitch, yaw, longitudinal and lateral speed, and acceleration, based on vehicle and environmental forces. The handling is designed to resemble that of a 4-cylinder hatchback. Verbal instructions and feedback are provided to the driver. On-screen display elements include rear-view and side mirrors, a dashboard, and navigational assistance (arrows indicating the direction of the next turn, where required). Four tasks of increasing difficulty were included in this study. The sequence and content of all driving tasks were identical for all participants. The simulator is programmed to provide summary data for each driving task. Raw driving data is not provided.

- *Task 1: Braking practice.* This task introduced drivers to the control and handling of the vehicle, particularly braking accuracy. The car accelerates automatically and drivers are required to brake to bring the vehicle to a stop behind a stationary car positioned in their lane. Drivers completed five trials at a starting speed of 50km/h and five trials at 80km/h. The simulator provided feedback after each trial regarding braking time and stopping position, thus training drivers to brake correctly and stop the vehicle in an appropriate position. The road is mostly straight with a few curves on approach to the stationary vehicles. No performance measures were recorded for this task.

- *Task 2: Brake Reaction Time (RT).* This task assessed Brake RT to an on-screen stimulus. Drivers accelerate to a target speed (90km/h), after which a stop sign appears on the centre display. Drivers must brake as fast as possible and bring the vehicle to a complete stop. This is repeated 5 times per trial. Participants completed 3 trials (i.e. 3 sets of 5 measurements). The first trial was for practice and the data were excluded from analyses. The simulator recorded Brake RT, calculated as time elapsed between the appearance of the stop sign and any application of the brake pedal. Two scoring methods were used to calculate performance for each of the three trials: Scoring Method 1 was the mean of all 5 RT measurements, and Scoring Method 2 was the mean of the measurements excluding the two most extreme (highest and lowest) scores.

- *Task 3: Distracted Driving.* This task assessed the effects of visual/cognitive distraction on vehicle handling. During the initial baseline portion of the task, participants drive as normal. In the following distraction portion of the task, participants respond verbally to a series of 20 TRUE/FALSE questions presented in the bottom left corner of the right display. The questions related to the number of words in a statement, e.g. "Four words exactly? The sooner, the better" (TRUE)[2]. Drivers were instructed by the simulator to maintain a constant speed of 80 km/h. The road features several gentle curves and frequent oncoming traffic. Task duration is around 4 minutes. Performance measures (calculated for baseline and distracted driving)

[2]Responses to the distraction questions were not recorded. Future researchers might like to record responses as a measure of task engagement

were Standard Deviation of Lane Position (SDLP, measured in metres), lane exceedances (percentage of time outside lane boundaries), average speed (km/h), and number of accidents.

- *Task 4: City Traffic Participation.* Set in a small city including intersections, traffic, pedestrians, and some hazards. Directions are provided by the simulator. The speed limit is 50km/h. The course contains various intersections and turns, including traffic lights, give way signs, stop signs, and pedestrian crossings. The route, traffic, and content of the scene were programmed to be identical for all drivers. Performance was automatically scored and reported for 9 grade categories: General, Lane Position, Speed Control, Steering, Signs, Car Following, Priority, Signalling, and Lane Changing. The maximum score for each grade category was 10 points. The Total Score was the sum of scores from each grade category (maximum 90). An Overall Grade was also given for the task, which was the minimum score recorded from the nine grade categories (e.g. if a driver received a Grade of 5 for General and a Grade of 8 for all other categories, the Overall Grade would be 5). Higher scores indicate better performance. Items in each grade category are shown in Table 6.1; points are deducted for each instance of an error. For example, speeding results in a 1.5 point deduction; if a driver incurs three speeding infractions, 4.5 points are deducted. The weighting of points is based on the seriousness of the error. Number of accidents and route errors were recorded. Duration is approximately 10 minutes. An example scene from this task is shown in Figure 6.2.

Cognitive and visual assessments. Fifteen participants had previously completed a series of cognitive and visual assessments as part of a separate study, conducted $4-9$ months earlier. These measures had been selected based on the theoretical relationship between the constructs and driving performance (Anstey et al., 2005; Mathias & Lucas, 2009). They were:

Mini Mental State Examination (MMSE). The MMSE is a short questionnaire designed to identify possible dementia or mild cognitive impairment (Folstein et al., 1975). The standardized MMSE is a valid and reliable tool for assessing cognitive impairment in

Table 6.1: Scoring for Driving Task 4: City Traffic Participation

Grade Category	Items (Points deducted per infraction in square brackets)
General	Accident [5], Engine stalled [1.75], Route error [1], Driving off with handbrake on [1], Handbrake on while driving [1]
Lane Position	Driving on the wrong side of the road [3.5], Driving in the wrong lane [2.5], Stopping too far from the stop line [2.5], Failure to keep left when not overtaking [1.5], Driving too close to the line [1]
Speed Control	Cornering too fast [2.5], Speeding [1.5], Driving too fast while approaching intersection [1.5], Braking too suddenly [1]
Steering	Driving on pavement [1.5], Driving on hard shoulder [1.5], Driving too much to the right [1.5], Driving off the road on the left [1.5], Cutting corner on intersection [1.5], Swinging wide on intersection [1.5]
Signs	Ignored red traffic light [5], No-entry sign ignored [2.5], Stop sign ignored [2.5], Stopped on pedestrian crossing [2.5], Stopped in intersection [2], Ignored yellow traffic light [1], Waited too long for green traffic light [1]
Car Following	Too close to car in front [1.5]
Priority	Failure to give way [3.5]
Signalling	Did not indicate early enough [1.5], Indicator switched off too soon [0.5], Indicating without a valid reason [0.5],
Lane Changing	Cut in while changing lane [5], Did not indicate correctly while changing lane [1.5], Driving too close to car in front while changing lane [1.5]

Note. Each Grade Category has a maximum score of 10. Points are deducted in each Grade Category for each instance of an error in that category (points displayed in square brackets).

Figure 6.1: Driving simulator

Figure 6.2: Example scene from Task 4: City Traffic Participation

older adults (Molloy & Standish, 1997). Scoring is out of 30 with lower scores indicating greater impairment.

Inspection Time (IT). IT is a measure of visual processing speed (Burns & Nettelbeck, 2003). Test-retest reliability for adults is .81 (Grudnik & Kranzler, 2001). In the computerised IT task, two high-contrast lines, one markedly shorter than the other, appear as a briefly-presented target before being masked. Participants indicate whether the shorter line was located left or right of a focal point. IT was measured in ms as the duration between target onset and mask onset; therefore lower scores indicate faster processing performance.

Useful Field of View Test™ (UFOV™) Subtest 2. UFOV™ is a computer-based test of visual attention and processing speed involving detection and localisation of briefly presented targets throughout the visual field (Ball & Owsley, 1993; Wood & Owsley, 2014). Only the second subtest, divided attention , was administered; this subtest is most correlated with UFOV™ Total score and best predicts driving outcomes (Ball et al., 2006; J. D. Edwards et al., 2006; Owsley et al., 1998). Test-retest reliability for

UFOV™ Subtest 2 is .82 (J. D. Edwards, Vance, et al., 2005). Participants are required to identify a briefly-presented central stimulus and locate a simultaneously presented peripheral stimulus. Score is exposure time (ms) for which 75% of trials were answered correctly; lower scores indicate better performance.

ProPerVis Crowding. ProPerVis assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns & White, 2007). The stimuli are a four-square parent figure and six figures derived from it, resembling stylised characters M, E, W, 3, 5, 2. One of the six figures is flanked on either side by the parent figure and appears randomly in one of five lateral positions on the screen. Participants attempt to identify which of the six figures was presented. The outcome was total errors made across the five positions from 40 trials; lower scores indicate better performance.

Sentence Span. A sentence span task was used to assess working memory. Task specifications (along with development and validation) are described by Lewandowsky et al. (2010). Briefly, participants were presented with a series of sentences and to-be-remembered letters. Each trial consisted of between 4 and 8 sentence/letter pairs. Participants answer TRUE or FALSE to each question (e.g. "All trees are plants"); after answering, a single letter was briefly presented on screen, which participants attempted to remember in order. The outcome was the overall proportion of correctly remembered letters.

Freiburg Visual Acuity Test. The Freiburg Visual Acuity Test is an adaptive computerised assessment that has been validated as a measure of visual acuity (Bach, 1996). We used the "Tumbling E" optotype with a viewing distance of 1.7m. Acuity was recorded as the logarithm of the Minimum Angle of Resolution (logMAR), with lower scores representing better visual acuity.

Pelli-Robson contrast sensitivity wall chart. The Pelli-Robson wall chart is widely used and is a reliable and valid measure of contrast sensitivity (Pelli et al., 1988). The outcome was log contrast sensitivity (possible range: 0 to 2.25, higher scores indicate better contrast sensitivity).

Procedure

The study was approved by The University of Adelaide Human Research Ethics Committee (approval number 14/94) and all participants provided written consent. Test sessions were held in the laboratory, but participants were asked to complete the online pre-test questionnaire before coming in for testing. Participants completed the four driving tasks in the order specified. All participants used automatic transmission. After each task, participants completed the Mini-SSQ and control feedback questionnaire via tablet computer. Participants were monitored by the experimenter for signs of SS, and were told to alert the experimenter and stop driving if they wished to discontinue driving for any reason. Between each driving task, participants completed the questionnaires via tablet computer and had a break for as long as they needed (typically $1-2$ minutes). A compulsory longer break of at least 5 minutes was provided between Task 3 and Task 4. A fan provided ventilation and room temperature was approximately 23°C. Participants were provided with cold water. Trials 2 and 3 of Task 2: Break RT were separated by a short break during which participants' comfort with the procedures was checked, because these trials were used to establish test-retest reliability of Brake RT. After all driving tasks had been completed, participants completed the simulator and scenario feedback questionnaire. The complete protocol took approximately one hour to complete.

Results

Simulator Sickness and dropout

Of 26 participants, 18 (69%) were able to complete all four driving tasks ($n = 1$ completed all driving tasks but did not receive a score for Task 4: City Traffic Participation due to a technical glitch). Task 4: City Traffic Participation was associated with the highest Mini-SSQ scores and greater dropout ($n = 8$ participants did not complete Task 4; $n = 1$ participant did not complete Task 3: Distracted Driving; all participants completed Task 1: Braking practice and Task 2: Brake RT). There were no significant differences between dropouts and completers on SSQ1 and SSQ2, but dropouts reported higher simulator sickness scores for SSQ3 (mean score dropouts = 1.75, $SD = 1.49$; mean score completers = 0.28, $SD = 0.58$; $t(7.94) = 2.71$, $p = .03$) and for SSQ4 (mean score dropouts = 5.71, SD = 1.50; mean score completers = 1.11, SD = 1.23; $t(23) = 7.92$, $p < .001$).

Changes in SSQ scores over the study are shown in Figure 6.3, displayed separately for completers and dropouts. For completers, only SSQ4 was significantly different from the other SSQ measures (SSQ1 − SSQ4, $t(17) = 3.18$, $p = .01$; SSQ2 − SSQ4, $t(17) = 4.12$, $p < .01$; SSQ3 − SSQ4, $t(17) = 2.83$, $p = .01$). There were no significant differences between the first three SSQ measurements. For dropouts, SSQ4 was significantly higher than all other SSQ measures (SSQ1 − SSQ4, $t(6) = 7.67$, $p < .01$; SSQ2 − SSQ4, $t(6) = 9.50$, $p < .01$; SSQ3 − SSQ4, $t(5) = 9.00$, $p < .01$), and SSQ3 was significantly higher than SSQ2 ($t(6) = 2.50, p = .05$).

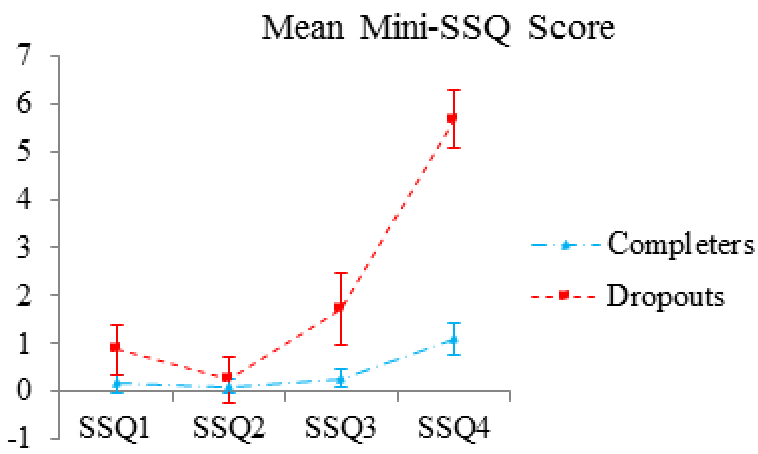

Figure 6.3: Mean Mini-SSQ scores. Error bars show within-subjects standard error.

Predictors of SS. None of the demographic variables were associated with SSQ Max ($r = -.15$ to .22, $p > .05$). Of the health variables, only motion sickness history score (MSSQ) was associated with SSQ Max ($r = .40$, $p = .04$; other variables $r = -.05$ to .27, $p > .05$). Participants with prior simulator experience experienced fewer SS symptoms (SSQ Max experienced $M = 1.44$, $SD = 1.65$; no experience $M = 5.00$, $SD = 1.93$; $t(24) = 4.82$, $p < .001$), and hours of simulator experience was also related to SSQ Max, with more hours of experience being related to fewer SS symptoms ($r = -.53$, $p < .01$). None of the technology use, driving history and habits, or cognitive and visual functioning variables were significantly associated with SSQ Max ($r = -.30$ to .30, $p > .05$). Overall, motion sickness history (MSSQ) and prior simulator experience were most related to SSQ Max. Together, these two variables accounted for 54% of the variance in SSQ Max, $F(2, 23) = 13.54$, $p < .001$. Similarly, dropout was significantly associated with motion sickness history and simulator experience. Dropouts scored significantly higher on the MSSQ than completers (mean for dropouts = 2.38, $SD = 3.34$; mean for completers = 0.64, $SD = 0.94$; $t(24) = 2.07$, $p = .05$). Prior experience with simulators was associated with less dropout, $\chi^2(1) = 10.61$, $p = .001$; of the 18 participants with prior simulator experience, only two dropped out; of the 8 participants with no simulator experience, 6 dropped out.

Simulator sickness/dropout and simulator performance. Dropouts and completers did not differ in regards to Brake RT, any of the driving measures Task 3: Distracted Driving, or Total Score or Overall Grade from Task 4: City Traffic Participation (Table 6.2). There was a significant difference in scores for the grade category "Steering", with dropouts performing better than completers (dropouts $M = 9.36$, $SD = 0.80$; completers $M = 6.94$, $SD = 2.80$; $t(22) = 2.22$, $p = .04$). However, this grade category was dependent on cumulative steering errors, and therefore dropouts had less time to incur infractions. There were no significant differences in the other grade categories from Task 4 ($t(22) = .11$ to 1.65, $p > .05$).

Reliability (Brake RT)

To assess test-retest reliability, the two "test" trials from Driving Task 2: Brake RT were compared. Test-retest reliability was high for both scoring methods; for the first scoring method, $r = .72$, $p < .001$, and for the second scoring method, $r = .82, p < .001$. For both scoring methods, there was no significant difference between the two trials ($t(25) = 1.14$ to 1.27, $p > .05$). Scoring method 2 produced significantly faster times for both trials ($t(25) = 2.84$ to 3.59, $p < .01$). Results from the Task 2: Brake RT task are shown in Table 6.3.

Table 6.2: Driving performance measures for completers and dropouts

	Completers		Dropouts		Difference	Effect Size
	Mean	SD	Mean	SD		(Cohen's D)
Driving Task 2: Brake RT[a]	0.59	0.09	0.62	0.08	$t(24) = 0.80$, $p = .43$	0.35
Driving Task 3: Distracted Driving						
Baseline Speed (km/h)	79.49	0.08	81.35	8.38	$t(22) = 0.90$, $p = .38$	0.36
Baseline SDLP (m)	0.39	0.10	0.38	0.10	$t(22) = 0.26$, $p = .80$	0.12
Baseline Lane Excursions (%)	5.18	4.51	5.09	4.19	$t(22) = 0.04$, $p = .97$	0.02
Distracted Speed (km/h)	77.08	3.64	74.89	7.23	$t(22) = 0.99$, $p = .34$	0.40
Distracted SDLP (m)	0.60	0.38	0.56	0.27	$t(22) = 0.23$, $p = .82$	0.12
Distracted Lane Excursions (%)	10.84	8.35	9.07	5.05	$t(22) = 0.49$, $p = .63$	0.26
Driving Task 4 (Total Score)	76.33	9.56	78.62	6.69	$t(22) = 0.58$, $p = .57$	0.28
Driving Task 4 (Overall Grade)	4.38	2.98	3.43	3.61	$t(22) = 0.67$, $p = .51$	0.29

[a] Mean Brake RT for Trial 3.

Table 6.3: Summary of results from Driving Task 2: Brake RT

	Score Method 1		Score Method 2	
	Mean (s)	SD (s)	Mean (s)	SD (s)
Brake RT Trial 1 (s)	0.62	0.10	0.61	0.10
Brake RT Trial 2 (s)	0.60	0.08	0.59	0.08

Validity

Age and simulated driving performance. A summary of results from Driving Task 4: City Traffic Participation is shown in Table 6.4. Our results showed that older age was associated with poorer overall performance on Driving Task 4: City Traffic Participation ($r = -.43$, $p = .04$). Older age was significantly associated with the grade categories Lane Position ($r = -.55$, $p = .01$) and tended towards association with Signalling ($r = -.40$, $p = .06$). The Lane Position category was graded based on appropriate positioning of the vehicle on the road, for example, being too far to the left or to the right, driving outside the lane, driving in the incorrect lane, and stopping position at stop signs. The Signalling category was based on appropriate use of turn signals, for example, applying the indicator when about to make a turn, signalling for an appropriate amount of time, and not using the indicator without a valid reason. Overall Grade and Other grade categories were not significantly correlated with age ($r = -.32$ to 0.29, $p > .05$)

Driving self-assessment and simulated driving performance. Descriptive statistics for the Driving self-assessment scale are shown in Table 6.5. There was no significant association between Driving Self-Assessment score and Driving Task 4: City Traffic Participation Total Score ($r = .16$, $p = .46$) or Overall Grade ($r = 0.24$, $p = .26$). Other grade categories for Driving Task 4 were not significantly associated with Driving Self-Assessment score ($r = -.23$ to .30, $p > .05$). Four drivers reported accident involvement in the past two years; accident involvement was significantly associated with poorer Total Score for Driving Task 4, $t(22) = 2.96$, $p = .01$, and suggested association with poorer Overall Grade for Driving Task 4, $t(22) = 1.93$, $p = .07$.

Table 6.4: Summary of results from Driving Task 4: City Traffic Participation ($n = 24$)

Score Category	Mean	SD	Range
Overall Grade	4.10	3.12	0 – 8.5
Total Score	77.00	8.74	52 – 88.5
General	7.46	3.81	0 – 10
Lane Position	7.96	3.08	0.5 – 10
Speed Control	8.13	2.51	1 – 10
Steering	7.65	2.62	0 – 10
Signs	7.73	2.71	0 – 10
Car Following	9.81	0.51	8.5 – 10
Priority	9.81	0.66	7.67 – 10
Signalling	9.04	1.35	4 – 10
Lane Change	9.42	1.15	6.67 – 10

Cognitive and visual measures and simulated driving performance. $n = 14$ participants who had a score available for Driving Task 4: City Traffic participation also had cognitive test scores available from a previous study. Descriptive statistics for these measures are also shown in Table 5. Total Score for Driving Task 4: City Traffic Participation was significantly correlated with IT ($r = -.69$, $p = .01$), UFOV™ Subtest 2 ($r = -.67$, $p = .01$), and suggested a moderate effect size with Crowding ($r = -.51$, $p = .06$). Overall Grade for Driving Task 4 was significantly correlated with IT ($r = -.74$, $p < .01$), Crowding ($r = -.62$, $p = .02$), and UFOV™ Subtest 2 ($r = -.66$, $p = .01$). All correlations were in the expected direction . There was no significant association between Overall Grade or Total Score for Driving Task 4 and MMSE, sentence span, visual acuity, or contrast sensitivity ($r = -.51$ to $.21$, $p > .05$).

Distracted driving assessment. Results from Driving Task 3: Distracted Driving showed mean speed was significantly lower during distraction, SDLP increased significantly during distraction, and lane excursions increased significantly during distraction. Table 6.6 shows descriptive statistics and effect sizes for each measure. There were sig-

Table 6.5: Descriptive statistics for self-assessment questionnaire, and cognitive and visual measures

Measure	Mean	SD	Range
Self-Assessment Total	42.79	4.93	30 – 51
MMSE	29.57	0.51	29 – 30
IT	67.13	19.12	45.38 – 101.75
Crowding	12.21	6.09	3 – 26
UFOV™ Subtest 2	55.29	43.99	17 – 147
Sentence Span	0.58	0.18	0.32 – 0.90
Contrast Sensitivity	1.80	0.16	1.65 – 1.95
Visual Acuity	0.09	0.07	−.03 – 0.24

Note. $n = 24$ for Self-Assessment Total; $n = 14$ for all other measures.

nificant differences in all three measures between baseline and distracted driving. These results will be compared to results from existing literature in the Discussion.

Table 6.6: Summary of results from Driving Task 3: Distracted Driving

	Baseline		Distracted		Difference	Effect Size
	M	SD	M	SD		(Cohen's D)
Speed (km/h)	79.96	4.38	76.53	4.70	$t(23) = 2.97$, $p = .007$	0.61
SDLP (m)	0.39	0.10	0.59	0.35	$t(23) = 3.43$, $p = .002$	1.18
Lane excursions (%)	5.16	4.34	10.40	7.60	$t(23) = 3.63$, $p = .001$	0.80

Face validity. Results from the Simulator Feedback questionnaire indicated that most participants agreed the scenarios accurately represented Australian roads (66%), and the scenarios represented situations relevant to everyday driving (93%). However, only 19% agreed using the simulator felt almost like driving a real car. Comments from participants during testing tended to suggest participants thought the physical setup of the simulator and the appearance and content of the scenarios were satisfactory, but they did not perceive the handling of the vehicle to be realistic. Particularly, participants reported that the steering was too sensitive and that the brake was too "hard".

Usability, acceptability and other feedback

General feedback. The Simulator Feedback Questionnaire showed good reliability (Cronbach's $\alpha = 0.76$). Overall, 77% of participants were generally positive towards the simulator. Participants were favourable towards the brake and accelerator (92% agreed they were comfortable controlling the brake and accelerator), but some participants reported feeling uncomfortable controlling the steering wheel (34% were neutral or disagreed that they were comfortable controlling the steering). Fifty-eight per cent of participants agreed that they felt like they were in control of the car, but only 38% agreed that the vehicle responded predictably. Mean Simulator Feedback Score was 43 ($SD = 6.12$, range $= 32 - 54$) out of a possible maximum of 68.

Vehicle control. Participants were generally confident controlling the brake, steering, and accelerator. Participants were comfortable following the directions provided by the simulator for each task (i.e. verbal and on-screen task instructions and route directions; over the 4 tasks, 92% agreed). After the final task, 88% agreed or strongly agreed that they could control the brake, 60% agreed or strongly agreed that they could control the steering, and 88% agreed or strongly agreed that they could control the accelerator, and 72% agreed or strongly agreed that they could control the vehicle overall. Participants were most confident controlling the vehicle in Task 2: Brake RT and during the no-distraction portion of Task 3: Distracted Driving, and were least confident controlling the vehicle while completing the secondary task portion of Task 3: Distracted Driving.

Predictors of user feedback. Higher feedback scores were associated with lower SSQ Max ($r = -.47$, $p = .04$); however, the feedback questionnaire was filled out after dropout (if applicable), therefore they may have provided less positive feedback because they experienced SS symptoms. Higher feedback scores tended to be associated with higher ATCQ Total score ($r = .35$, $p = .08$). Participants who had previous experience with simulators provided more positive feedback than those who had no previous simulator experience (means scores 44.78 and 39.13 respectively; $t(24) = 2.37$, $p = .03$). Although participants with simulator experience reported more favourable feedback and experienced fewer SS symptoms, there were no significant differences in simulated driving performance measures between participants with previous simulator experience and participants without previous simulator experience (Task 2: Average Brake RT (Trial 3): $t(24) = 0.60$, p = .56; Task 3: Distracted Driving measures, $t(22) = 0.08$ to 1.19, $p > .05$; Task 4: City Traffic Participation measures: $t(22) = 0.11$ to 1.90, $p > .05$). There was a significant relationship between user feedback and number of accidents in the past two years ($r = -.39$, $p = .05$), such that higher feedback was associated with fewer accidents. For the 15 participants with previous cognitive and visual assessment scores available, UFOV™ Subtest 2 and visual acuity were associated with user feedback. Better performance on UFOV™ Subtest 2 was associated with more positive feedback ($r = -.51$, $p = .05$). Better visual acuity was associated with more positive feedback ($r = -.53$, $p = .04$).

Discussion

The aim of the study was to assess the usability, user acceptability, validity, and reliability of a fixed-base three-screen driving simulator for older adult drivers. Participants completed four simulated driving tasks and provided feedback relating to the simulator and the scenario content. We investigated the validity of the simulator by a) comparing individual's demographic information, self-reported driving behaviour, and cognitive ability with their performance on the driving simulator; b) comparing measures from a distracted driving task with known effects of distraction from other simulator and on-road studies; and c) obtaining user ratings related to the face validity of the simulator. We also investigated the test-retest reliability of the Brake RT Task, the extent of SS symptoms exhibited by participants during the study, and the usability and acceptability of the simulator and the selected tasks from the point of view of the users.

Validity

In regards to validity, we found some evidence towards the convergent validity of the simulator in that age, cognitive test performance, and self-reported accident history were related to simulator performance as expected, although driving self-assessment scores were not related to simulated driving performance as expected. Evidence of content validity was found for Driving Task 3: Distracted Driving, where observed effects on the simulator corresponded to reported real-world effects. Evidence of face validity was found with participants indicating that the driving tasks were a relevant assessment of their driving skills and were relevant to everyday driving.

Our results showed significant correlations between simulated driving performance, as measured by the Total Score and Overall Grade from Driving Task 4: City Traffic Participation, and age and cognitive ability. Older age was associated with poorer Total Score, consistent with the finding that older age (more specifically, the cognitive decline generally associated with aging) tends to be associated with poorer driving performance. Poorer cognitive ability tends to be related to poorer driving outcomes (Anstey et al., 2005; Mathias & Lucas, 2009),and our results were consistent with this: UFOV™ Subtest 2, IT, and Crowding were all significantly correlated with Total Score and Overall Grade for Driving Task 4. UFOV™, particularly UFOV™ subtest 2, has been consistently linked to driving ability and various outcomes including retrospective and prospective crashes, on-road driving performance, and simulated driving performance (Ball et al., 2006; Clay et al., 2005; Mathias & Lucas, 2009; Wood & Owsley, 2014). IT, a measure of processing speed, and Crowding, a measure of efficiency of processing across the visual field, have both been found to be highly related to UFOV™ performance (Burns et al., 2005; Matas, Nettelbeck, & Burns, 2014). Total Score and Overall Grade for Driving Task 4: City Traffic Participation were not related to overall self-assessment score. This could indicate that the items on the self-assessment questionnaire were not sensitive enough to differentiate driving ability in a sample of high-functioning older drivers. Alternatively, this could be due to impaired drivers over-estimating their driving ability in the self-report questionnaire[3]. Drivers who reported accident involvement during the past two years performed significantly worse on Driving Task 4: City Traffic Participation. Although the accidents were self-reported, it has been suggested that in the Australian context self-report crash

[3]After publication of this study, an anonymous reviewer has commented that it would be useful to compare self-assessment scores with another measure such as informant ratings or on-road driving performance, in order to determine the accuracy of the self-assessment scores

data may actually be more accurate and preferable to official records because large numbers of older driver crashes are not reported to authorities, particularly minor crashes (Anstey, Wood, Caldwell, Kerr, & Lord, 2009). These results provide evidence towards the convergent validity of the simulator; that is, simulator performance is related to other constructs that are theoretically related to on-road driving ability (Mullen et al., 2010; Shechtman, 2010).

Results from Driving Task 3: Distracted Driving showed that drivers reduced their speed, increased lane position variability, and increased the occurrence of lane exceedances while completing the secondary task. To establish the content validity of this task, we aimed to compare the results from the present study with results from existing literature on driver distraction. Several previous simulator studies have established the effects of distractions including text messaging, mobile phone use, and other visual and cognitive distractions on speed and lateral position while driving. A meta-analysis of mainly simulator-based studies found a significant decrease in speed and lateral control while texting (Caird, Johnston, Willness, & Asbridge, 2014). Specifically, effect sizes of $D = 0.67$ (reading text messages) and $D = 0.79$ (typing and reading) were reported for measures of lane position (including SDLP and lane exceedances), with drivers showing decreased lateral control while text messaging, and an effect size of $D = 0.67$ was reported for speed, with drivers decreasing their speed while text messaging (Caird, Johnston, Willness, Asbridge, & Steel, 2014). These effects are generally similar to the effects of talking on a cell phone (Strayer, Watson, & Drews, 2011).

Few on-road text messaging studies have been conducted; Yager, Cooper, and Chrysler (2012) conducted an on-road, closed circuit study and showed that SDLP increased significantly while writing and reading text messages. Östlund et al. (2004) conducted a large-scale study investigating the effects of visual and cognitive distraction on simulated and on-road driving and reported that visual distraction had effects on lateral control and speed measures; however, effects on measures such as SDLP and lane exceedances were much more pronounced in the simulators than on the road. In a follow-up study utilising the same visual distractor task, results showed a significant reduction in speed while distracted, for both simulated and on-road driving, although there was no effect on SDLP or lane exceedances either in the simulator or on-road (Engström et al., 2005).

Our results were in agreement with these results from previous studies looking at the effects of distraction (particularly text messaging) on simulated and on-road driving (Caird, Johnston, Willness, & Asbridge, 2014; Caird, Johnston, Willness, Asbridge, & Steel, 2014; Strayer et al., 2011; Yager et al., 2012). Our results were closely aligned with previous simulator studies, with effect sizes of similar magnitude; we found effect sizes of $D = 0.61$ for speed, $D = 1.18$ for SDLP, and $D = 0.80$ for lane exceedances, compared to effect sizes of $D = 0.67$ for speed and $D = 0.67$ to 0.79 for lateral position measures as reported in meta-analyses (Caird, Johnston, Willness, & Asbridge, 2014; Caird, Johnston, Willness, Asbridge, & Steel, 2014). The results were also in general agreement with the few on-road studies that have been conducted on this topic. This comparison with existing literature provides evidence towards the content validity of the Distracted Driving task on our simulator, that is, whether real-world driving behaviour can be inferred from behaviour on the simulator (Shechtman, 2010).

In regards to face validity, most participants agreed that the scenarios accurately represented Australian roads, and that the scenarios represented situations relevant to everyday driving. Users were comfortable in the driver's seat, and comments from participants suggested that they perceived the controls and appearance of the simulator to be quite realistic. However, only a minority agreed that using the simulator felt almost like driving a real car. Comments from participants during testing tended to suggest that they did not perceive the handling of the vehicle to be realistic. Particularly, participants reported that the steering was too sensitive and that the brake was too "hard". These concerns about the controls were very similar to those reported by Gibbons et al. (2014),who noted that older drivers experienced difficulty with the steering and pedals, while younger drivers did not. Using terminology suggested by Shechtman (2010), results suggest that driver response validity (i.e. how driving behaviours in the simulator correspond to real-world driving) was satisfactory, while vehicle response validity (i.e. how well the dynamics and handling of the simulated vehicle correspond to real-world vehicle dynamics) was perceived to be problematic by the participants, particularly in regards to cornering and braking.

Reliability

Results from Driving Task 2: Brake RT indicated that it had excellent test-retest reliability ($r = .82$, $p < .01$). The results also indicated that a practice trial was a useful and necessary addition to the protocol; participants improved significantly from Trial 1 to Trial 2, but there was no significant difference in RTs from Trial 2 and Trial 3. The analysis of two separate scoring methods suggested that discarding the highest and lowest RT from each trial improved the reliability and stability of the measure. This method means that participants are not unfairly penalised for one unrepresentative slow reaction, and are not unduly advantaged if they record one significantly faster RT by attempting to guess when the stimuli would appear. The trial 2 and trial 3 mean RTs of $0.61s$ and $0.59s$ corresponded well to mean brake RTs reported in the literature for similar tasks (Hollis et al., 2013; Martin et al., 2010; Zhang et al., 2007).

Usability and user feedback

Results from the Simulator User Feedback Questionnaire and Scenario Feedback Questionnaire indicated that users were generally positive towards the simulator. The relatively high mean score on the Simulator User Feedback Questionnaire (43 out of a maximum of 68, with higher scores indicating more positive feedback) suggested that participants, in general, approved of the physical setup of the simulator, the appearance and content of the scenes and scenarios, and their ability to use the simulator. The items relating to the control of the car received less favourable responses. For example only 19% of participants agreed that the simulator felt almost like driving a real car, 76% agreed or were neutral towards the statement that the car would go off the road often, and 57% thought that the simulator did not give an accurate reflection of how they normally drive. The pattern of results and comments from participants suggests that they did not perceive certain elements of the vehicle dynamics, specifically the steering and the brake pedal, to be realistic.

We investigated whether user feedback was related to participant characteristics such as age, gender, computer use and attitudes, simulator history, and cognitive ability. The results showed that less positive feedback was associated with higher SSQ. The relationship between user feedback and SS has been noted previously; Schultheis et al.

(2007) showed that participants who experienced SS provided significantly less favourable feedback towards the simulator. Schultheis et al. (2007) suggested that user comfort with the simulator may be predictive of SS onset. In the present study, participants did not fill out the feedback questionnaire until they had completed driving, so it is possible that experience of SS symptoms during testing influenced how they responded to the questionnaire.

Simulator feedback score was marginally associated with ATCQ Total Score, suggesting that users who were less favourable towards computers were also less favourable towards the simulator. It has been suggested that older users may have difficulties using driving simulators due to unfamiliarity and anxiety relating to computer use (Classen & Brooks, 2014). However, as technology use becomes more prevalent among older adults and as the population ages, it is likely that attitudes towards computers among older adults will improve. The pattern of results relating to simulator feedback suggests that although negative feedback was associated with generally poorer outcomes on the predictor variables (i.e. higher SS, poorer attitudes towards computers, prior accident history, and poorer visual acuity and UFOV™ Subtest 2 score), having a negative opinion of the simulator did not contribute to poorer performance on the simulated driving tasks. Therefore, simulator performance appears to be independent of user attitude towards the simulator and towards technology in general.

Simulator sickness

A majority of participants (96%) were able to complete the first three driving tasks, and 69% of participants were able to complete all four driving tasks. Task 4: City Traffic Participation contained several elements known to contribute to SS, such as sharp corners, frequent stopping and starting, and a visually complex scene (Cassavaugh et al., 2011; C. J. Edwards et al., 2004; Mourant et al., 2007; G. D. Park et al., 2006). Nonetheless, a dropout rate of 31% is below the average simulator dropout rate for older adults of 40% reported by Trick and Caird (2011). However, our lower droupout rate may be attributable to the high percentage of participants with previous simulator experience, which Trick and Caird's figure does not account for.

Leaving the most demanding, potentially sickness-causing tasks to the end of the protocol means that data collection can be maximised from the simpler tasks, and also gives users more time to adapt to the simulator (Domeyer et al., 2013; Graeber, 2001). Overall, SS was not associated with performance on the simulated driving tasks (Mullen et al., 2010). As expected, we found that motion sickness history (MSSQ) was associated with higher SS, and experience with simulators was associated with lower SS. We found no effects for age or gender but this may be due to the relatively small sample size and restricted age range. Previous results have indicated that older age in general may be a risk factor for SS, that is, older adults as a group are more susceptible to SS than younger adults (Brooks et al., 2010; Cassavaugh et al., 2011; Classen et al., 2011; Matas et al., 2015; G. D. Park et al., 2006; Trick & Caird, 2011). We found no relationship between cognitive test measures and SS. However, our sample was restricted to cognitively intact older adults; Freund and Green (2006) reported that older drivers classified as impaired on the MMSE tended to be more likely to report SS. Our results also confirmed that the Mini-SSQ is an appropriate and sensitive tool for monitoring changes in SS over time (Matas et al., 2015; Mourant et al., 2007).

Limitations

This study has provided a promising starting point for confirming the validity, usability, and acceptability of the SimWorx SX06DTS-L driving simulator for assessment of older drivers.

The volunteer were healthy, active older drivers and we intentionally excluded drivers who were not living independently, not in good health and did not drive regularly. It is therefore likely that the sample consisted of high-functioning members of society and excluded those who may be experiencing problems with their driving. Our sample also included a high number of participants (69% of the sample) who had previous experience with simulators. These participants were self-selected volunteers, which may have affected the results, particularly relating to SS and dropout; participants experiencing SS previously may have been less inclined to participate. The participants with simulator experience were significantly more favourable in their feedback than those who had no prior simulator experience, which may also have been related to the previous experiences of the self-selected sample. However, in terms of driving performance, previous simulator experience was not associated with performance measures from the simulated driving tasks.

Simulator reliability and validity is dependent on the specific simulator and tasks used. Although we aimed to assess the reliability of the driving simulator, we only collected reliability measures relevant to Task 2: Brake RT. The other tasks selected did not have reliability data readily available, or would have required a significant increase in the length of the protocol to collect. Regarding validity, we aimed to include a range of different driving tasks and validation methods. Although the results described here are relevant for this simulator model and the selected tasks, the methodology described in the present study could be implemented to evaluate the validity and usability of other driving simulators and driving tasks[4].

Future work will expand on the present results, for example establishing the test-retest reliability of other simulated driving tasks, continuing to establish the convergent validity of the simulator by obtaining varied and accurate information about driver's on-road driving ability, and by assessing the validity and acceptability among other groups of users, for example younger drivers or clinical populations.

Conclusion

We investigated the acceptability and usability, and measures of reliability and validity, for a fixed-base three-screen driving simulator for use with older adult drivers. Users were generally positive towards the simulator but some reported difficulty adapting to controlling the vehicle. Participants tended to agree that the tasks were a relevant assessment of their driving skills and were relevant to everyday driving. Thirty-one per cent of participants experienced SS and had to stop driving, although this occurred primarily in the final driving task (Task 4: City Traffic Participation), which contained several features that can increase the risk of SS (for example, right-angle turns and frequent stopping and starting). All participants completed the first three tasks with minimal SS symptoms, with only one participant dropping out due to SS prior to Driving Task 4. Task 2: Brake RT demonstrated excellent test-retest reliability. Task 3: Distracted Driving was shown

[4]After publication of this study, an anonymous reviewer has noted additional study limitations of the small sample size and lack of on-road validation of the driving simulator. The results should be interpreted in consideration of these limitations. Ultimately, the validation of the simulator requires assessment in terms of on-road outcomes.

to be a valid measure of the effects of distraction on driving performance measures, and the driving assessment in Task 4: City Traffic Participation demonstrated validity in that performance was related to age, cognitive ability, and accident history. Overall, the simulator and tasks appear to be appropriate for assessing driving performance in older adults.

Chapter 7.

Study 4: Cognitive screening and driving simulation for younger and older drivers

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2016). Cognitive screening and driving simulation for younger and older drivers. (Manuscript submitted for publication)

Statement of Authorship

Title of paper	Cognitive screening and driving simulation for younger and older drivers
Publication status	Submitted for publication
Publication details	Submitted

Principal Author

By signing the Statement of Authorship, I certify that this paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper

Name of principal author	Nicole Amy Matas		
Contribution to paper	Study concept and design, data collection, statistical analyses, writing and submitting manuscript, addressing referee comments		
Overall percentage	80%		
Signature		Date	13/9/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ted Nettelbeck		
Contribution to paper	Principal supervision, advice about development of concept and design of study, manuscript revision and approval		
Signature		Date	13. 9. 16

Name of Co-Author	Nicholas Burns		
Contribution to paper	Advice about development of concept and design of study, assistance with statistical analyses, manuscript revision and approval, ongoing		
Signature		Date	14/9/16

Abstract

Introduction: Cognitive screening tools, discriminant validity, acceptability, and simulated driving performance were investigated for older and younger drivers using a three-screen fixed-base driving simulator. **Method:** Younger drivers ($N = 63$, mean age 20 years) and older drivers ($N = 82$, mean age 73 years) completed a Brake Reaction Time (RT) Task and a Traffic Participation Task on the simulator, Simulator Sickness Questionnaires, and a Simulator Feedback Questionnaire. Participants also completed a battery of cognitive assessments: Useful Field of View Test™ (UFOV™) Subtest 2, Inspection Time (IT), and ProPerVis Crowding. **Results:** The results showed age-group differences in performance on the simulated driving tasks. Younger participants recorded significantly faster Brake RT and better scores on the cognitive assessments. Older adults were significantly more affected by visual crowding (Cohen's $D = 2.44$), especially as the eccentricity of target presentation increased. On the Traffic Task, younger participants recorded more Speeding violations, and older participants recorded more errors in Steering, Signalling, and General (Accidents and Route Errors). These results reflected a difference in driving styles between the age groups. **Conclusions:** The results provided evidence for the discriminant validity and usability of the driving simulator, and for the utility of the cognitive screening tools for older drivers. Overall, simulator feedback was generally positive from both age groups. Older adults were significantly more prone to simulator sickness and dropout. **Practical applications:** Driving simulators are useful tools for assessing driving performance in older and younger drivers. For older adults, Simulator Sickness must be appropriately monitored. UFOV™ Subtest 2, IT and ProPerVis Crowding were sensitive to age-related cognitive declines that are relevant for driving performance in older adults and may be useful as screening tools for older drivers.

Introduction

For driving simulators to be useful, they must be a valid measure of driving behaviour; that is, performance in the simulator must accurately reflect on-road driving performance. Additionally, users must take the assessment seriously and they must perceive the driving simulator and tasks to be acceptable and relevant. Simulator validity and usability are highly dependent on the specific simulator, task, and population under consideration (Kaptein et al., 1996; Mullen et al., 2010; Shechtman, 2010). Therefore, the aim of the present study was to assess discriminant validity and acceptability of a three-screen, fixed-base driving simulator and two selected simulated driving tasks: a Brake RT Task and a Traffic Participation task. The previous pilot study investigated the validity (content, convergen, and face), test-retest reliability, and acceptability of our driving simulator for use with older drivers (Matas, Nettelbeck, & Burns, 2016). The study found support for the content and convergent validity of two simulated driving tasks (the Traffic Participation task and a Distracted Driving task), and the Brake RT task showed good test-retest reliability. A Simulator Feedback questionnaire was used to establish the usability and face validity of the simulator and tasks, and it was found that users were generally positive towards the simulator and considered that the scenarios used provided an appropriate assessment of their driving ability. To expand on these results, and to investigate the discriminant validity and acceptability of the driving simulator, the present study was conducted with a larger sample of healthy older drivers, and a comparison group of younger drivers. The participants completed the two simulated driving tasks, a brief battery of cognitive tests, a user feedback questionnaire, and a simulator sickness questionnaire. The inclusion of the younger group permitted age group comparisons, which were used to investigate the discriminant validity of the cognitive screening tasks and the simulated driving tasks.

In regards to driving performance, discriminant validity can be investigated by examining age-group differences in simulated driving performance and comparing these to known real-world effects. For example, one study used discriminant validity to demonstrate statistically significant differences in the performance of non-drivers, novice drivers, and experienced drivers both on a simulator and during on-road driving (Mayhew et al., 2011). In the present study, age-group comparisons were used to examine the discriminant validity of the selected driving tasks. For the Brake RT task, the younger drivers were expected to record faster Brake RT than the older drivers, based on the known slowing of speed of processing and RT with age (Deary et al., 2009; Der & Deary, 2006) and age differences in Brake RT where younger drivers tend to react faster in both simulated and on-road assessments (Dickerson, Reistetter, Burhans, & Apple, 2016; Makishita & Matsunaga, 2008; Warshawsky-Livne & Shinar, 2002). For the Traffic Task, it was expected that older and younger drivers would perform differently on different aspects of the driving tasks. Younger drivers have generally been reported to commit more deliberate violations than older drivers (e.g. speeding, close following; de Winter & Dodou, 2010; Reason, Manstead, Stradling, Baxter, & Campbell, 1990) and their crashes have often been attributed to risk-taking behavior (McGwin Jr & Brown, 1999). On the other hand, while they commit fewer deliberate violations, older drivers have been reported to be more susceptible to errors relating to making timely and accurate decisions in complex situations such as navigating intersections (Clarke, Ward, Bartle, & Truman, 2010; Langford & Koppel, 2006; McGwin Jr & Brown, 1999) and have been found to be more prone to errors and violations involving signs (Reason et al., 1990). Therefore it was expected that the younger drivers would perform worse on the Traffic Task grade categories relating to intentional violations, while the older drivers would perform worse on Traffic Task grade categories relating to errors in judgement and decision-making.

Usability and acceptability of the simulator need to be considered because it is important that users perceive the driving simulator and tasks to be a relevant and accurate assessment of their driving ability. A simulator with high usability will contribute to the user's accuracy of task performance, enjoyment of the experience, and acceptance of the technology (Schultheis et al., 2007). Few studies have investigated issues of usability and acceptability relating to driving simulation. In one such study, Schultheis et al. (2007) used a feedback questionnaire to investigate the usability of a Virtual Reality driving rehabilitation simulator in a sample of participants with acquired brain injury and healthy controls (age range 20 – 68). They found that those with acquired brain injuries rated the simulator less favourably, and that older age was also associated with less favourable ratings. They also found that user feedback ratings were related to simulator sickness, with those reporting more positive feedback being less likely to experience simulator sickness. In another study, Gibbons et al. (2014) used a semi-structured interview to obtain feedback from older (age 65+) and middle-aged (age 40 – 55) drivers in relation to a one-screen and three-screen simulator for assessment of fitness-to-drive; middle-aged drivers responded more positively than older drivers, and 85% of middle-aged and 60% of older drivers preferred the three-screen simulator. The middle-aged drivers were more likely to respond that they perceived the simulator to be realistic, that the simulated tasks provided a reasonable reflection of their driving skills, and that a simulator assessment could enhance current procedures for examining fitness-to-drive. In the present study, we used a Simulator Feedback Questionnaire derived from that of Schultheis et al. (2007) to determine if our driving simulator was acceptable to participants, and to investigate age group differences regarding participants' perceptions of usability.

An issue related to simulator usability and acceptability is Simulator Sickness. Simulator Sickness refers to a set of motion sickness-like symptoms that can occur with use of simulators and virtual reality technology. Simulator Sickness and dropout rates vary considerably depending on the configuration of the simulator and the demands of the simulated task. The onset of Simulator Sickness can affect a user's ability to complete the intended assessment tasks. This is of particular concern for older drivers, who tend to be more susceptible to Simulator Sickness (Cassavaugh et al., 2011; Classen et al., 2011; Trick & Caird, 2011). Several recent driving simulator studies have reported dropout rates of between 0% and 44% for older adults (e.g. Bélanger et al., 2010; Caird et al., 2007; Domeyer et al., 2013; Lee, Lee, Cameron, & Li-Tsang, 2003; Sklar et al., 2014; Trick et al., 2010) compared to dropout rates of between 0% and 17% for younger adults (e.g. Bélanger et al., 2010; Domeyer et al., 2013; Shechtman et al., 2007; Yang et al., 2006). In the present study, we investigated rates of Simulator Sickness and Dropout on our simulator among older and younger drivers.

We also examined three cognitive screening tools that have the potential to be used for driver assessment: UFOV™ Subtest 2, IT, and ProPerVis Crowding. Safe driving ability depends on cognitive, sensory, and physical factors, and cognitive factors typically show the strongest association with various outcomes (Anstey et al., 2005). Cognitive screening tools may therefore be useful for identifying at-risk older drivers, particularly those already showing signs of functional declines or risky driving behaviours. For these assessments to be acceptable as screening tools, they should be sensitive to age-related cognitive declines that are relevant for driving performance in older adults.

The aim of the present study was to assess the discriminant validity and usability of the driving simulator by comparing younger and older drivers, and to investigate age-group differences in simulated driving performance, simulator sickness, and user perceptions of usability. To measure driving performance, two simulated driving tasks were selected: a Brake RT Task, and a Traffic Participation Task. Participants also completed a brief battery of cognitive assessments assessing processing speed and visual attention. It was hypothesised that:

- The younger drivers will record quicker Brake RTs
- There are age group differences in performance on the Traffic Task
- Younger drivers will provide more positive feedback about the simulator
- Older drivers are more susceptible to Simulator Sickness
- Younger drivers perform better on the cognitive measures reflecting processing speed and visual attention (Inspection Time (IT), Useful Field of View Test™ (UFOV™) Subtest 2, and ProPerVis Crowding)

Method

Participants

Older participants were aged 65 years or over, living independently in the community, held a valid Australian driver's license and drove at least once per week, had normal or corrected-to-normal vision, and reported good physical and mental health. Participants were excluded if they reported a history of epilepsy or seizure-related conditions, a diagnosis of cognitive impairment, or a history of motion sickness or simulator sickness. These participants were recruited from a database of older adults who had previously participated in, or expressed interest in participating in driving research. Participants were also recruited through flyers and announcements at various community groups, and through snowball sampling. Interested participants were directed to a website containing further information about the study and a preliminary questionnaire assessing their eligibility for the study.

The online questionnaire was accessed by 92 older individuals. Of these, 82 (54 males, 28 females; mean age 73 years) were able to attend the on-campus test session (n = 3 were screened out due to poor health; n = 3 were unavailable to attend; n = 2 cancelled due to illness; n = 1 cancelled due to personal commitments; n = 1 was uncontactable). Eighty participants attempted the simulated driving tasks and 54 completed all driving tasks (n = 26 dropped out due to simulator sickness). There were no significant differences between those who attended the study and those who did not in age, education level, self-rated health, number of medical conditions, motion sickness history, or self-rated driving skill level (response to single item on 5-point scale from "poor" to "excellent").

Younger participants (N = 63, 41 female, 22 male; mean age 20 years) were recruited from first year students studying Psychology at the University of Adelaide. Participants were aged 17 – 24, held at least a P1 provisional driver's license, and drove at least once per week.

Participants were excluded if they had an extensive history of motion sickness, suffered from epilepsy or another seizure-related condition, or did not regularly drive at least once per week, or were currently not in their usual state of health and fitness. The study was advertised via posters throughout the University and on the School of Psychology's online research participation system. Participation was voluntary and participants received course credit.

Materials

Driving simulator. The simulator used was a SimWorx SX06DTS-L Driver Training Cockpit (Figure 7.1). The simulator has a fixed base and is located in the laboratory. The simulator has a three screen display (3 × 27in HD monitors, total resolution 5760 × 1080 pixels). The simulated horizontal field of view spans a total visual angle of 210°, but this can be accessed in its entirety only via head movement, or eye movement, or both. Thus, it is analogous to viewing through the front and side windows of a motor vehicle. The graphics were texture-rich. All scenarios required driving on the left hand side of the road, consistent with Australian road rules. The simulator has a driver cockpit including force feedback steering wheel, pedals, gearshift, handbrake, ignition switch, engine start button, and adjustable seat. The steering is controlled by a Logitech G27 unit modified with a steering wheel of 350mm diameter. The steering wheel and pedals were calibrated with the Logitech Profiler and the configuration menu on the simulator. The vehicle dynamics model is linear and computes the roll, pitch, yaw, longitudinal and lateral speed, and acceleration, based on vehicle and environmental forces. The handling is designed to resemble that of a 4-cylinder hatchback. Verbal instructions and feedback are provided to the driver. On-screen display elements include rear-view and side mirrors, a dashboard, and navigational assistance (arrows indicating the direction of the next turn, where required).

The simulator was programmed to provide summary data for each driving task. Raw driving data was not provided. The sequence and content of all driving tasks were identical for all participants. There were three simulated driving tasks:

- *Practice Drive.* The aim of the practice drive was for drivers to familiarise themselves with the simulator. Performance on the practice drive was not assessed. The speed of the vehicle was automatically controlled by the simulator, and the driver had control of the steering wheel, indicators, and some control of the brake (the simulator was programmed to automatically slow down to an appropriate speed on approach to corners, and the driver could brake at other times, though this was not necessary to complete the task successfully). The vehicle initially accelerated to 60km/h, and gradually increased the speed to 100km/h during the task. The driver was required to steer the vehicle on bends and corners. Task duration was approximately 8 minutes.

- *Brake Reaction Time (RT).* This task assessed Brake RT to an on-screen stimulus. Drivers accelerated to a target speed (90km/h) along a straight road. After the target speed had been reached, a stop sign randomly appeared on the centre display. Drivers were instructed to brake as quickly as possible and bring the vehicle to a complete stop. This was repeated five times per trial. The simulator recorded Brake RT, calculated as time elapsed between the appearance of the stop sign and any application of the brake pedal. The outcome was Average Brake RT, calculated as the mean Brake RT from the five test trials. Drivers completed one practice run and one test run, with a brief break in between.

- *Traffic Task.* This task was set in a small city including intersections, traffic, pedestrians, and some hazards. Directions are provided by the simulator. The speed limit is 50km/h. The course contains various intersections and turns, including traffic lights, give way signs, stop signs, and pedestrian crossings. The route, traffic, and content of the scene were programmed to be identical for all drivers. Performance was automatically scored and reported for eight grade categories: General, Lane Position, Speed Control, Steering, Signs, Car Following, Priority, and Signalling. The maximum score for each grade category was 10 points. The Total Score was the sum of scores from each grade category (maximum 80). An Overall Grade was also given for the task, which was the minimum score recorded from the eight grade categories (e.g. if a driver received a Grade of 5 for General and a Grade of 8 for all other categories, the Overall Grade would be 5). Higher scores indicate better performance. Items in each grade category are shown in Table 7.1; points were deducted for each instance of an error. For example, speeding results in a 1.5 point deduction; if a driver incurs three speeding infractions, 4.5 points are deducted. The weighting of points is based on the seriousness of the error, and this is calculated automatically by the simulator based on pre-programmed item weights. Number of accidents and route errors were recorded. Duration is approximately 10 minutes. An example scene from this task is shown in Figure 7.2.

Figure 7.1: Driving simulator

Mini-Simulator Sickness Questionnaire (SSQ). The Mini-SSQ (Mourant et al., 2007) contains six questions and is a short form of the Simulator Sickness Question-

Table 7.1: Scoring for Traffic Participation task

Grade Category	Items (Points deducted per infraction in square brackets)
General	Accident [5], Engine stalled [1.75], Route error [1], Driving off with handbrake on [1], Handbrake on while driving [1]
Lane Position	Driving on the wrong side of the road [3.5], Driving in the wrong lane [2.5], Stopping too far from the stop line [2.5], Failure to keep left when not overtaking [1.5], Driving too close to the line [1]
Speed Control	Cornering too fast [2.5], Speeding [1.5], Driving too fast while approaching intersection [1.5], Braking too suddenly [1]
Steering	Driving on pavement [1.5], Driving on hard shoulder [1.5], Driving too much to the right [1.5], Driving off the road on the left [1.5], Cutting corner on intersection [1.5], Swinging wide on intersection [1.5]
Signs	Ignored red traffic light [5], No-entry sign ignored [2.5], Stop sign ignored [2.5], Stopped on pedestrian crossing [2.5], Stopped in intersection [2], Ignored yellow traffic light [1], Waited too long for green traffic light [1]
Car Following	Too close to car in front [1.5]
Priority	Failure to give way [3.5]
Signalling	Did not indicate early enough [1.5], Indicator switched off too soon [0.5], Indicating without a valid reason [0.5],
Lane Changing	Cut in while changing lane [5], Did not indicate correctly while changing lane [1.5], Driving too close to car in front while changing lane [1.5]

Note. Each Grade Category has a maximum score of 10. Points are deducted in each Grade Category for each instance of an error in that category (points displayed in square brackets).

Figure 7.2: Example scene from the Traffic Participation task

naire (Kennedy et al., 1993). The Mini-SSQ is quick to administer, suitable for repeated administration, and sensitive to changes in driving conditions and changes in SS symptoms over time (Matas et al., 2015; Mourant et al., 2007). Six symptoms were included: general physical discomfort, headache, blurred vision, sweating, faintness/dizziness, and stomach discomfort/nausea. Possible responses were on a 4-point scale, corresponding to None (0), Slight (1), Moderate (2), Severe (3). Scores for each symptom were summed to give a total score. The Mini-SSQ was administered after the Practice Drive (SSQ1) and after the Traffic Participation Task (SSQ3). An additional measure, SSQ Max, was recorded as the highest SSQ reported by the participant.

Simulator feedback questionnaire. Participants provided feedback relating to their overall experience with the simulator. Items in the simulator feedback questionnaire were based on the questionnaire used by Schultheis et al. (2007). The questions related to various aspects of the simulator, including ability to use the controls, how the scenarios were presented, and the content of the tasks. Responses were made on a 5-point scale from Strongly Disagree (1) to Strongly Agree (5). Responses from the 11 items were summed to give a total Simulator Feedback Score (maximum 55).

Inspection Time (IT). IT measures processing speed (Burns & Nettelbeck, 2003). Two high-contrast lines, one markedly shorter than the other, appear as a target on a computer screen. Participants indicate whether the shorter line was located left or right of a focal point. Time available for processing is limited by a backward masking procedure and reduced or extended using an adaptive staircase algorithm according to response accuracy. Targets are preceded by a warning cue ("+" in the centre of the screen, 370ms) and are immediately followed by a mask figure shaped like two lightning bolts. First, there were three sets of 10 practice trials with decreasing target presentation time for each set (835ms, 420ms, and 250ms) that required answering 10/10 items correctly for the first and second set, and 9/10 items for the third set. IT was measured in ms as the duration between target onset and mask onset at which the viewer achieved 79% accuracy. Test-retest reliability for adults has been reported to be .81 (Grudnik & Kranzler, 2001).

ProPerVis Crowding. The ProPerVis Crowding subtest assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns et al., 2005). The stimuli are a four-square parent figure and six figures derived from it, resembling stylised characters M, E, W, 3, 5, 2. On each trial, one of the six figures is presented, flanked on either side by the parent figure. The target and flankers appear randomly in one of five lateral positions on the screen. Participants attempt to identify which of the six figures was presented. The outcome was total errors made across the five positions from 40 trials; lower scores indicate better performance.

Useful Field of View Test™ (UFOV™) Subtest 2. UFOV™ is a computer-based test of visual attention and processing speed involving detection and localisation of targets briefly presented throughout the visual field (Ball & Owsley, 1993; Wood & Owsley, 2014). UFOV™ Subtest 2 is most correlated with UFOV™ Total score and best predicts driving outcomes (Ball et al., 2006; J. D. Edwards et al., 2006; Owsley et al., 1998). In UFOV™ Subtest 2, one of two stimuli (either a car or a truck) appears briefly in central vision, and the car appears simultaneously in the periphery at one of the eight cardinal or intercardinal points, 12.5 cm from the centre of the display. Participants indicate which object was presented in central vision, and the location of the car in the periphery. Test-retest reliability for UFOV™ Subtest 2 is reportedly .82 (J. D. Edwards, Vance, et al., 2005). Score is exposure time (ms) for which 75% of trials were answered correctly; lower scores indicate better performance.

Procedure

The study was approved by The University of Adelaide Human Research Ethics Committee. All participants provided written consent and were provided with digital and hard copies of the study information sheet.

Test sessions were held in the laboratory. The order and content of tasks differed slightly for older and younger participants.

Older participants first completed the cognitive assessments amongst a battery of cognitive and visual tests. Within the battery, the tests were completed in the following order: IT, Crowding, and UFOV™ Subtest 2. The older participants then completed the three driving tasks in the following order: Practice Drive, Brake RT, and Traffic Task. After each task, participants completed the Mini-SSQ via tablet computer. After all driving tasks had been completed, older participants remained in the laboratory for approximately 10 minutes while they were monitored for symptoms of SS and completed feedback questionnaire via tablet computer.

Younger participants completed an online questionnaire containing questions about their driving habits, followed by the cognitive assessments in the following order: UFOV™ (all subtests), Crowding, and IT. They then completed the driving tasks in the same order as the older participants, but they also completed a 10-minute distracted driving task between Brake RT and the Traffic Task. Because of the known lower occurrence of simulator sickness in young adults, the Mini-SSQ was only administered after the Practice Drive and after the Traffic Task. The simulator feedback questionnaire was completed after the driving tasks.

The protocol for all participants took approximately 1.5 hours to complete. All participants were monitored by the experimenter for signs of SS, and were told to alert the experimenter and stop driving if they wished to discontinue driving for any reason. Between each driving task, participants took a break for as long as they needed. A fan provided ventilation and room temperature was approximately 23°C. Participants were provided with cold water. All participants used automatic transmission.

Results

Analysis

Data were analysed using R version 3.2.4 (R Core Team, 2016).

Of the 82 older participants who attended testing, one participant did not bring vision correction and was unable to complete cognitive testing or attempt the simulated driving tasks. One participant was unable to complete UFOV™ Subtest 2 due to difficulty understanding the task requirements. One participant elected not to attempt the simulated driving task due to health concerns. For the simulated driving tasks, 73 older participants (89%) completed the Brake RT task, and 53 participants (64%) completed the Traffic Task. In regards to driving performance measures, only participants who completed the driving tasks were included in analyses (i.e. $n = 73$ for the Brake RT task and $n = 53$ for the Traffic Task).

Of the 63 younger participants who attended testing, one participant did not complete the Crowding task due to difficulty distinguishing the task stimuli. One participant did not complete the Traffic Task due to simulator sickness.

Descriptive statistics

Table 7.2 displays descriptive statistics for cognitive and driving performance measures for younger and older drivers.

For older drivers, Traffic Task Total Score suggested good overall performance, with a median score of 71.5 out of 80. However, the median score for Traffic Grade was 5 out of 10. Fifteen participants (28%) had a Traffic Grade of less than 5. This indicates that the average participant achieved a poor score in at least one Grade Category. As seen in Table 7.2, the worst Grade Category was Lane Position with a mean score of 7.63 (SD = 2.89) out of 10 (median 7.5), and 7 participants scoring less than 5 out of 10. The best categories were Car Following, Priority, and Signalling, with no participants scoring less than 5 out of 10 for these categories.

Younger drivers also performed well overall on the Traffic Task, with a median score of 72.6 out of 80. As for older drivers, the median score for Traffic Grade was 5 out of 10. Twenty four participants (39%) had a Traffic Grade of less than 5. As seen in Table 7.1, the worst Grade Category for younger drivers was Speed Control, with a mean score of 6.23 (SD = 3.34) out of 10 (median 7), with 20 participants scoring less than 5 out of 10. The best categories were Lane Position, Car Following, Priority, and Signalling, with no participants scoring less than 5 out of 10 for these categories.

Cognitive test performance

Welch independent samples *t*-tests indicated that the younger participants performed significantly better than the older participants on the three cognitive performance

Table 7.2: Descriptive statistics for cognitive and driving performance measures for older and younger drivers

Variable	Older					Younger					Difference	
	N	Mean	SD	Range	Fail	N	Mean	SD	Range	Fail	$t(df)/W$	p
Age	82	72.93	4.39	65 to 84		63	20.20	1.51	17 to 24			
IT	82	61.5	18.09	30.25 to 147.13		63	41.39	2.57	23.38 to 70.13		8.50 (131.46)	< .001
Crowding Total	81	12.15	4.86	2 to 24		62	2.32	2.57	0 to 12		15.58 (126.84)	< .001
UFOV™ Subtest 2	81	64.96	55.3	17 to 250		63	23.38	19.55	16.7 to 133.3		6.28 (104.3)	< .001
Brake RT	73	0.61	0.12	0.34 to 1.06		63	0.58	0.09	0.36 to 0.82		1.783 (129.04)	.039
Traffic Task Total	53	70.09	6.54	52 to 80		62	71.55	5.92	56 to 80		1850.5	.981
Traffic Task Grade	53	5.25	2.67	0 to 10	15	62	4.99	3.14	0 to 10	24	1637.5	.999
General	53	8.34	2.56	0 to 10	4	62	9.19	2.56	0 to 10	4	2065	.009
Lane Position	53	7.63	2.89	0 to 10	7	62	9.55	1.05	6.5 to 10	0	2338	< .001
Speed Control	53	8.51	2.28	0 to 10	3	62	6.23	3.34	0 to 10	20	972	< .001
Steering	53	8.42	1.73	4 to 10	1	62	9.07	1.40	2.5 to 10	1	1994	.195
Signs	53	8.24	1.99	0.5 to 10	2	62	8.21	2.14	0 to 10	3	1994	.195
Car Following	53	9.97	0.21	8.50 to 10	0	62	9.95	0.27	8.5 to 10	0	1621	.999
Priority	53	9.74	0.75	7.67 to 10	0	62	9.57	0.99	5.33 to 10	0	1537.5	.999
Signalling	53	9.25	1.02	6 to 10	0	62	9.78	0.59	7.5 to 10	0	2201	< .001
Simulator Feedback	74	35.38	6.39	24 to 55		63	43.49	5.53	31 to 55		7.96 (134.96)	< .001

Note. Fail = Number of participants with grade less than 5. a Performance on the Brake RT test was compared between younger and older drivers with a one-sided Welch independent samples t-test. Performance on the Traffic Task was compared between younger and older drivers with the Wilcoxon Rank Sum Test. p was corrected for multiple comparisons using Holm correction.

measures (Table 7.2). The effect size for the difference (Cohen's D) showed large to very large age group differences for all three measures (UFOV™ Subtest 2, $D = 0.96$; IT, $D = 1.33$; Crowding, $D = 2.44$). The distribution of scores for the older and younger drivers for IT, UFOV™ Subtest 2, and Crowding is shown in Figure 7.3. Floor effects were evident for the younger drivers on UFOV™ Subtest 2 and Crowding, and for the older drivers on UFOV™ Subtest 2.

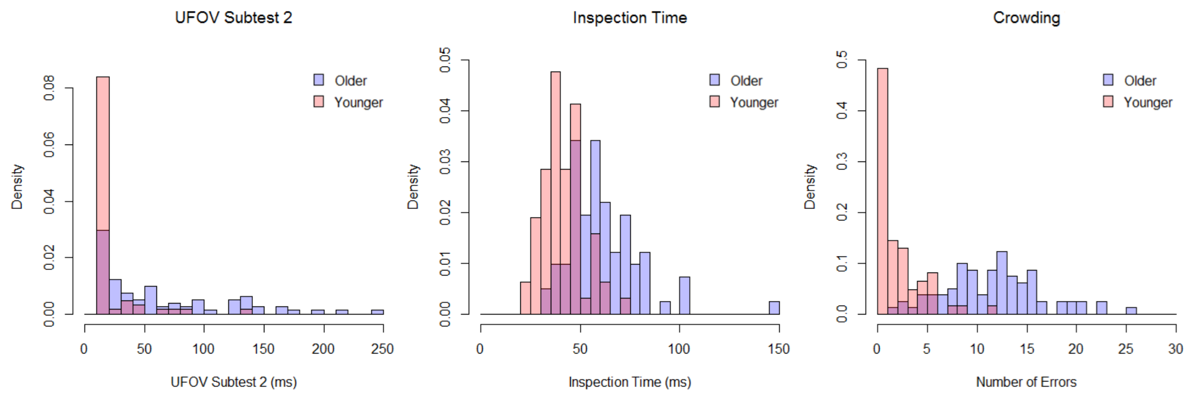

Figure 7.3: Distribution of cognitive test performance by age group

Further investigation of the Crowding measure showed significant age group differences at each of the five stimulus positions ($p < .001$ for all positions). A multilevel model was used to investigate the interaction between age and stimulus position (central, parafoveal and peripheral)[1]. The interaction is shown in Figure 7.4. The model indicated significant main effects for stimulus position $\chi^2(2) = 242.00$, $p < .001$ and age $\chi^2(1) = 122.10$, $p < .001$. The interaction between age and stimulus position was also significant, $\chi^2(2) = 182.44$, $p < .001$. Contrasts showed a significant difference between

[1]Peripheral left and peripheral right were combined to provide mean errors for the peripheral position; parafoveal left and parafoveal were combined to proved total errors for the parafovel position. In the older age group, there was no significant difference in scores for the two parafoveal positions, and for the two peripheral positions. In the younger age group, there was no significant difference in scores at the parafoveal position, but there was a significant difference in scores for the peripheral position, with participants recording more errors at the right peripheral position. Additional analysis showed that if the higher mean score for position 5 was used for the younger adults instead of the combined score, the parameters of the multilevel model remained approximately equal.

older and younger adults when comparing the central and parafoveal positions ($b = 0.90$, $t(282) = 5.00$, $p < .001$) and for the parafoveal and peripheral positions ($b = 1.90$, $t(282) = 10.54$, $p < .001$).

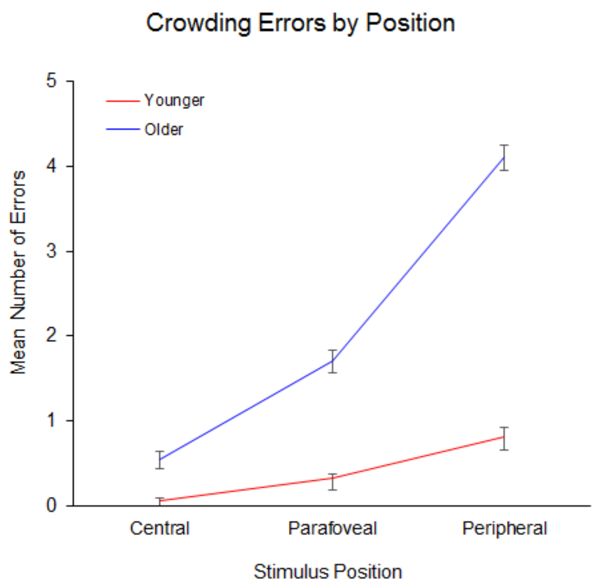

Figure 7.4: Interaction between age group and target position for crowding errors. Error bars display within-subjects standard error.

Simulator sickness and dropout

Simulator sickness was measured using the Mini-SSQ and was assessed after the Practice Drive (SSQ1) and after the Traffic Task (SSQ3). Of the older participants, 80 participants attempted the simulated driving tasks. Of these, $n = 3$ participants (4%) dropped out during or immediately after the Practice Drive, assessment, $n = 8$ participants dropped out during or immediately after the Brake RT task (10%), and $n = 15$ dropped out during the Traffic Task (19%; total dropouts: $n = 26$, 33%). Of the younger participants, $n = 1$ (1.5%) participant dropped out during the Traffic Task.

For the older participants paired samples t-tests indicated SSQ3 was significantly higher than SSQ1 ($t(66) = 4.57$, $p < .001$). For the younger participants, a paired samples t-test indicated that SSQ3 was significantly higher than SSQ1 ($t(62) = 2.76$, $p = .008$).

Welch independent samples t-tests indicated that younger participants reported significantly lower SS scores than older participants. This was the case when all participants were included (SSQ1: $t(100.48) = 6.485$, $p < .001$; SSQ3: $t(115.29) = 4.107$, $p < .001$), and when only completers were included (SSQ1: $t(68.19) = 4.192$, $p < .001$; SSQ3: $t(78.65) = 3.474$, $p < .001$).

Table 7.3: Descriptive statistics for SSQ scores

SSQ Measure	Older					Younger				
	N	Mean	Mdn	SD	Range	N	Mean	Mdn	SD	Range
SSQ1	80	1.27	1	1.44	0 to 6	63	0.16	0	0.48	0 to 3
Complete	53	0.91	1	1.20	0 to 5	62	0.16	0	0.49	0 to 3
Dropout	27	2.00	2	1.62	0 to 6	1	0	0	-	-
SSQ3	67	2.4	2	2.75	0 to 10	63	1.82	0	0.75	0 to 10
Complete	52[a]	1.93	1	2.51	0 to 10	62	0.58	0	1.4	0 to 9
Dropout	15[b]	3.93	3	3.08	0 to 9	1	10	10	-	-

[a,b] SSQ3 Scores for one older dropout and one older completer were not recorded

Simulated driving performance measures for older and younger drivers

Brake RT. Brake RT scores for younger and older drivers are shown in Table 7.2. A one-sided Welch independent samples t-test indicated that younger drivers recorded significantly faster Brake RT than older drivers, $t(129.04) = 1.783$, $p = .039$. Cohen's D showed a small effect for the group difference ($D = 0.299$).

Traffic Task. Traffic Task scores for younger and older drivers are shown in Table 7.2. Wilcoxon Rank Sum tests with Holm correction for multiple comparisons indicated that there was no significant difference in Traffic Task Total Score or overall Grade. However, younger drivers scored significantly higher than older drivers in the grade categories General, Lane Position, and Signalling. Older drivers scored significantly higher than younger drivers in the grade category Speed Control. The difference was approximately two points for Speed Control and Lane position; approximately one point for General; and approximately half a point for Signalling.

Data were also recorded for number of Accidents and Route Errors. Of the older drivers, 15 of the 53 (28%) participants who completed the Traffic Task recorded an accident. Of the younger drivers, 3 of the 62 participants (5%) recorded an accident. This difference in accident rate was significant, $\chi^2(1) = 10.20$, $p < .001$. Of the older drivers, 8 of the 53 (16%) participants recorded a Route Error. Of the younger drivers, none of the 62 participants recorded a route error. This difference in Route Errors was significant, $\chi^2(1) = 7.86$, $p < .001$.

Simulator feedback

Descriptive statistics for the Simulator Feedback Questionnaire Total Score are displayed in Table 7.2. Cronbach's α indicated that the internal consistency reliability of the scale was 0.89 for younger drivers and 0.86 for older drivers. Overall, feedback was positive in both age groups. A Welch independent samples t-test indicated that younger drivers reported significantly higher Feedback Total scores than older drivers, $t(134.96) = 7.965$, $p < .001$. This represented a very large difference between the age groups (Cohen's $D = 1.35$). Analysis of individual scale items showed medium to very large difference for all items (Cohen's $D = 0.56$ to 1.27), with younger participants providing more favourable responses for all items.

The relationship between Simulator Sickness and Feedback Total was also assessed. It was found that there was a significant correlation between SSQ Max and Feedback Total, $r = -.45$, $p < .001$.

A linear regression model was assessed with Feedback Total as the outcome and SSQ Max and Age Group as predictors. Both predictors were significant and the model explained 38% of the variance in Feedback Total, $F(2, 134) = 42.05$, $p < .001$. An additional model investigated the interaction between SSQ Max and Age Group; the interaction was not statistically significant. The relationship between the variables is shown in Figure 7.5. No significant correlations were found between Feedback scores and simulated driving performance.

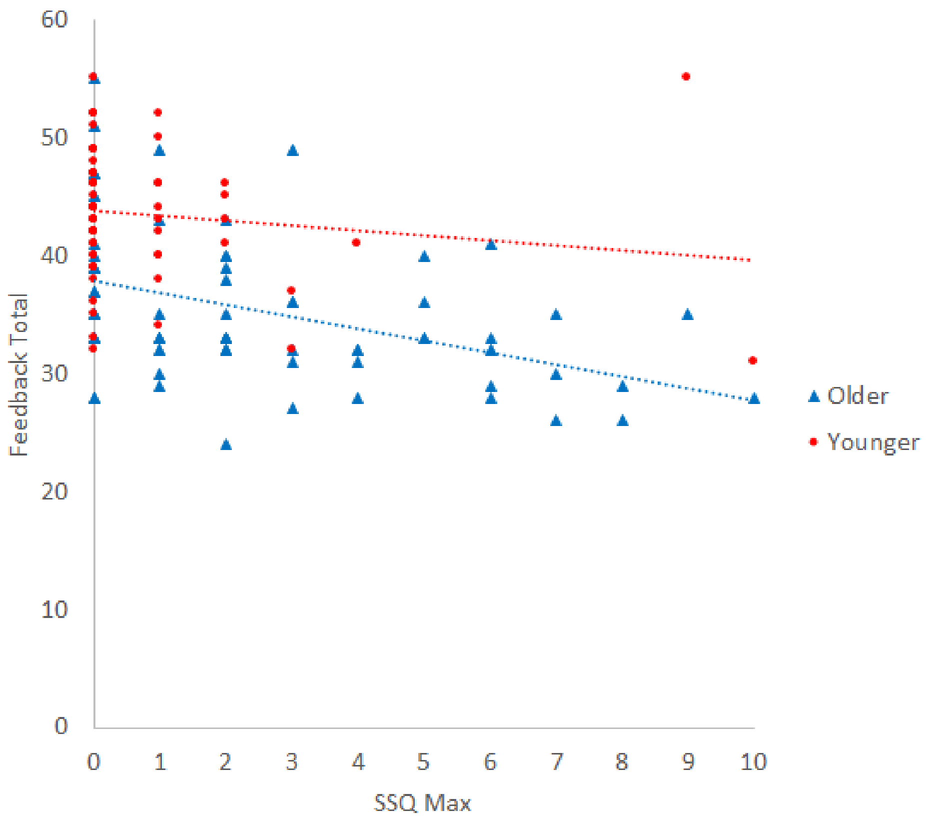

Figure 7.5: Relationship between Feedback Total, SSQ Max, and age group

Discussion

This study aimed to provide evidence for the discriminant validity, usability, and acceptability of the driving simulator used. A convenience sample of healthy younger and older drivers completed a series of simulated driving tasks to assess their driving performance. Simulator sickness symptoms were monitored and participants provided feedback regarding their experience with the simulator. Participants also completed a brief battery of cognitive assessments. To provide evidence for the discriminant validity of the simulator, we investigated age group differences in driving performance on two tasks: a Brake RT task and a Traffic Participation task. To assess the usability and acceptability of the simulator, we analysed responses and age group differences in simulator sickness, dropout, and feedback.

The results indicated that the younger drivers recorded significantly faster Brake RT scores, as hypothesised. This provided evidence for the discriminant validity of the task; Brake RT has generally been reported to be slower in older adults (Dickerson et al., 2016; Makishita & Matsunaga, 2008; Warshawsky-Livne & Shinar, 2002). The effect is often small, especially for Simple RT tasks, which is consistent with the result. The slowed Brake RT for older adults has been attributed to the perceptual response time to the simulator, rather than the physical movement time from accelerator to brake (Warshawsky-Livne & Shinar, 2002), although this was not measured in the present study. However, consistent with this, the present results indicated significant age group differences in performance on the cognitive assessments, which reflected aspects of processing speed and visual attention. The present result is also consistent with the general slowing of cognitive speed and RT with age (Der & Deary, 2006).

Although there were no age differences in overall scores for the Traffic Task, there were significant age group differences in some of the individual grade categories. It was hypothesised that younger drivers perform worse in the grade categories reflecting intentional violations, and that older drivers would perform worse in the grade categories reflecting errors in information processing and decision making. Some parts of the results partially supported the hypothesis. Results showed that the younger drivers scored significantly lower than the older drivers on the Grade Category Speed Control. This category penalised speeding, which is more prevalent in younger drivers (de Winter & Dodou, 2010;

Reason et al., 1990). Although it was noted that the younger drivers incurred more speeding penalties than older drivers, it was not observed whether the older drivers kept to the posted speed limit. Future studies should include this point in analysis, because older drivers might be compensating for slower reaction times by lowering their travel speed. Conversely, younger drivers scored significantly lower than older drivers in the grade categories General, Lane Position, and Signalling. The Grade Category General penalised accidents and route errors, which occurred significantly more frequently in the older age group. The Traffic Task involved navigating intersections and interacting with other traffic, factors that are known to be associated with older driver crashes (Clarke et al., 2010; Langford & Koppel, 2006; McGwin Jr & Brown, 1999). Data relating to the types of accidents encountered by the older and younger drivers was not recorded in the present study. Future researchers might like to investigate this further in regards to the types of crash that are more likely to occur in older drivers; this would be particularly relevant in a driving simulator which had the ability to modify the programming in order to link crashes to their causes, for example brake response or swerving in response to hazards. The Grade Category Lane Position penalised vehicle handling errors including driving in the wrong lane, stopping too far from the stop line, and driving too close to the edge of the lane. It is possible that the difference in this grade category is attributable to the difficulties experienced by the older drivers in adapting to the handling of the simulator (Matas et al., 2016). Similarly, the difference in the Grade Category Signalling could reflect unintentional errors by the older drivers in adapting to the sensitivity of the indicator and accurately monitoring the state of the turn signal. Older drivers experience difficulty when visual processing demands are high and experience greater attentional demands (Leversen, Hopkins, & Sigmundsson, 2013; Stinchcombe, Gagnon, Zhang, Montembeault, & Bedard, 2011). In the context of the Grade Category Signalling, the older drivers may have been more focussed on the roadway environment and had difficulty monitoring the state of the indicator and the turn signal on the dashboard.

Simulator sickness affected the older drivers much more than the younger drivers, and the older drivers were more likely to dropout before completing the driving tasks. Of the older drivers, 33% dropped out, whereas only one younger participant (1.5%) dropped out. Older drivers also reported significantly more simulator sickness symptoms than younger drivers. This is consistent with findings that older age is a risk factor for simulator sickness (Cassavaugh et al., 2011; Classen et al., 2011; Matas et al., 2015), and other studies that have reported a higher dropout rate for older drivers compared to younger drivers (Brooks et al., 2010; Caird et al., 2007; G. D. Park et al., 2006; Sklar et al., 2014). This result emphasises the importance of appropriate screening and monitoring of older adults using driving simulators (Matas et al., 2015; Trick & Caird, 2011; Trick et al., 2010).

Feedback from both age groups was generally positive overall. As hypothesised, the younger drivers were significantly more positive in their responses to the Simulator Feedback Questionnaire than the older drivers. There was also a significant correlation between Simulator Sickness scores and Feedback; higher SSQ scores were associated with lower Feedback. There was no significant interaction between Age Group and SSQ score. This result is important because acceptability and usability affect how accurately users can perform tasks on the simulator, and how relevant they perceive the simulator and driving tasks to be as accurate assessments of their driving ability. A possible reason for the age group difference could be that the older users may have difficulties using driving simulators due to unfamiliarity and anxiety relating to computer use (Classen & Brooks, 2014; Matas et al., 2016). As technology use becomes more prevalent among older adults and as the population ages, it is likely that attitudes towards computers among older adults will improve. It is also notable that having a negative opinion of the simulator was not related to poorer performance on the driving tasks. Therefore, simulator performance appears to be independent of user attitude towards the simulator and towards technology in general.

Participants also completed cognitive assessment. It was found that the younger drivers performed significantly better on these tests, as hypothesised. The differences in performance was particuarly evident for the Crowding measure, which showed a very large effect size for the mean difference (Cohen's D = 2.44). The significant large to very large age group differences in cognitive test performance occurred even though the older participants were high-functioning and highly educated compared to the general population of older drivers. For the younger drivers, there was no significant association between the cognitive abilities and driving performance (see Chalacas, 2015). For the older drivers, all three cognitive measures showed small but significant correlations with the Brake RT task, and Crowding showed a small but significant correlation with the

Traffic task. This result demonstrated that the three cognitive measures (IT, UFOV™ Subtest 2, and particularly Crowding) were sensitive to age-related cognitive declines that are relevant for driving performance in older adults. The usefulness of these cognitive tests as potential screening assessments for older drivers will be considered in future research. Crowding may be particularly relevant for driving performance in complex scenarios, such as city driving with high traffic density. This may also be related to the difficulties that older drivers experience with intersections and gap selection where they must efficiently and accurately process the location of traffic in peripheral vision.

The results from the present study have provided initial evidence for the discriminant validity of the driving simulator and tasks. The age group differences in performance in the two tasks were generally as expected and demonstrate discriminant validity for these tasks. The results relating to simulator sickness and user feedback have implications for the usability of the simulator in younger and older adults. Older adults were more prone to simulator sickness and provided less positive feedback towards the simulator. However, feedback from both age groups was considered to be generally positive, and encouragingly, feedback was independent of task performance. The current study adds to the results reported in Matas et al. (2016) and suggests that the simulator and selected tasks may be an appropriate option for assessment of older drivers. It should be noted that ultimate validation of the driving simulator will require assessment in terms of on-road driving.

Practical applications

The results suggest that driving simulators may be useful tools for assessing driving performance in older drivers and avoid the risks and costs associated with on-road driving tests. Users should be aware that older adults are prone to simulator sickness and dropout. The cognitive assessment tasks (UFOV™ Subtest 2, IT and ProPerVis Crowding) showed very large age-group differences indicating that they were sensitive to age-related cognitive declines that are relevant for driving performance in older adults. These tasks may be useful as screening tools for identifying at-risk older drivers.

Chapter 8.
Study 5: Predictors of simulated driving performance in healthy older adults: Cognitive ability, visual function, and physical activity

Matas, N. A. (2016). Predictors of simulated driving performance in healthy older adults: Cognitive ability, visual function, and physical activity. (Unpublished Manuscript)

Statement of Authorship

Title of paper	Predictors of simulated driving performance in healthy older adults: Cognitive ability, visual function, and physical activity
Publication status	Publication style
Publication details	Unpublished manuscript

Principal Author

By signing the Statement of Authorship, I certify that this paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper

Name of principal author	Nicole Amy Matas		
Contribution to paper	Study concept and design, data collection, statistical analyses, writing and submitting manuscript, addressing referee comments		
Overall percentage	80%		
Signature		Date	13/9/16

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: i. the candidate's stated contribution to the publication is accurate (as detailed above); ii. permission is granted for the candidate in include the publication in the thesis; and iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ted Nettelbeck		
Contribution to paper	Principal supervision, advice about development of concept and design of study, manuscript revision and approval		
Signature		Date	13. 9. 16

Name of Co-Author	Nick Burns		
Contribution to paper	Advice about development of concept and design of study, assistance with statistical analyses, manuscript revision and approval, ongoing		
Signature		Date	14/9/16

Abstract

Driving is a complex task requiring intact physical, visual, and cognitive function, factors that are known to decline with normal aging. Screening tests based on various cognitive, sensory, and physical factors may therefore be useful for identifying at-risk older drivers. This study investigated several of these factors in relation to simulated driving performance. Eighty two healthy older drivers (54 males) aged 65 – 84 years (mean 73 years) completed a battery of cognitive tests assessing general cognitive functioning (Mini Mental State Examination, MMSE), processing speed (Inspection Time, IT), crowding across the visual field (Proficiency of Peripheral Visual Processing (ProPerVis) Crowding), and divided visual attention (Useful Field of View Test™ (UFOV™) Subtest 2). Participants also completed two measures of visual function (Visual Acuity and Contrast Sensitivity) and a self-report measure of physical activity (Physical Activity Scale for the Elderly (PASE)). Performance on the cognitive, visual, and physical activity measures were used to predict performance on two simulated driving tasks: a Brake Reaction Time (RT) Task, and a Traffic Task requiring interaction with other traffic, pedestrians, and some hazards in a small city environment. The results showed small but significant correlations between cognitive performance (IT, Crowding, and UFOV™ Subtest 2) and Brake RT. For the Traffic Task, only Crowding was significantly correlated with Traffic Task Grade. MMSE, PASE, and the visual measures were not associated with simulated driving performance. Results and implications are discussed in relation to older driver screening models.

Introduction

Older drivers represent an increasing proportion of active drivers and, as the population ages, more older drivers are predicted to hold and maintain a driver's license than has been the case for any previous generation (Koppel & Berecki-Gisolf, 2015; Lyman et al., 2002; OECD, 2001). Older drivers have often been reported to be over-represented in serious crashes, especially when crash rate is considered per distance driven (Li et al., 2003; OECD, 2001), and are more likely to be considered at-fault compared to younger drivers (L. E. Hakamies-Blomqvist, 1993; McGwin Jr & Brown, 1999). If they do become involved in a crash, older drivers are at increased risk of serious injury or death because of their greater fragility (Li et al., 2003). However, it has been argued that older drivers are not over-represented in crashes when the "low-mileage bias" (drivers travelling fewer kilometres have higher crash-risk, and older drivers are more likely to driver fewer kilometres than other age groups) is taken into consideration (Langford et al., 2006).

Nonetheless, because of the apparent risk faced by older drivers, much research has focussed on identifying those older drivers who may be most at-risk. It has been recognised that mandatory age-based testing for older drivers is not an effective strategy for reducing crash risk (Fildes et al., 2008). For example, comparison of Australian jurisdictions found no reliable differences for crash data and serious accidents between jurisdictions with and without mandatory age-based fitness to drive testing (Langford, Fitzharris, Koppel, & Newstead, 2004; Langford, Fitzharris, Newstead, & Koppel, 2004). It has also been noted that in Australia, the proportion of drivers with cognitive or visual impairments was similar regardless of the requirement for age-based testing, and that mandatory age-based testing has been associated with lower rates of driving (Ross et al., 2011).

As an alternative to mandatory age-based testing, it has been suggested that driver screening be applied only to those drivers who are most at-risk, for example those drivers who are showing functional impairments that may affect their driving (Fildes, 2008; Langford, 2008; OECD, 2001). Although age itself is not a good predictor of driving safety, chronological age is associated with functional declines that can affect driving ability (Anstey & Wood, 2011). Reviews suggest that testing various cognitive, sensory, and physical factors may prove useful in identifying those older adults who may be at greater risk for driving difficulty (Anstey et al., 2005; Janke, 1994; Mathias & Lucas, 2009). Based on this, Fildes et al. (2008) described a targeted License Assessment procedure which relies on at-risk drivers being referred for assessment by various community members, for example medical and health practitioners, family members, police, and the licensing authority. Once referred, drivers may be required to undergo validated off-road screening tests, which would be used to make a decision of "fit to drive", "unfit to drive", or "unclear" (further assessment required). Complementary to this model, Anstey et al. (2005) have described a Multifactorial Model of Driving Safety, which specifies that "Capacity to Drive Safely" depends on cognition, vision, and physical function, which are themselves inter-related. Actual driving behaviour is determined by both the Capacity to Drive Safely and Self-Monitoring Beliefs (accurate insight into driving ability). A key factor in both of these models is that measures of functional ability (cognitive, sensory, and physical) should reliably and validly predict driving outcomes. There has been much research on this issue, but there remains a lack of consensus in regards to which screening tests should be used.

In regards to cognitive abilities, it is well established that cognitive performance declines with age, and that several of the cognitive abilities prone to decline are associated with driving ability (Anstey et al., 2005; Mathias & Lucas, 2009). The cognitive domains most commonly associated with safe driving ability in older adults include visual attention, processing speed, working memory, executive functioning, and general cognitive function. These relationships have been observed for many driving outcomes, including performance on on-road driving assessments, real-world driving outcomes such as accidents and infringements, and performance on simulated driving tasks (Mathias & Lucas, 2009).

In the present study, we aimed to investigate the relationship between three brief cognitive assessments that are potentially useful as driver screening tests, and performance on simulated driving tasks. The three tasks were Useful Field of View Test™ (UFOV™) Subtest 2, Inspection Time (IT), and ProPerVis Crowding, and they were selected to represent aspects of attention and visual processing which are relevant to the driving task. The three tasks were brief, taking approximately 5 minutes or less each to complete, easy to administer via personal computer, and automatically scored.

We also measured general cognitive function using the Mini-Mental State Examination (MMSE; Folstein et al., 1975). The MMSE tests general cognitive function used for identifying possible dementia or cognitive impairment. It assesses areas such as orientation to time and place, short-term memory, and attention. In a review focussing on older drivers with dementia, it was found that MMSE was moderately correlated with on-road and simulated driving performance (Adler et al., 2005). On the other hand, some studies have found no association between MMSE score and driving outcomes in both impaired (Molnar et al., 2006) and non-impaired samples (Crizzle et al., 2012; Joseph et al., 2014). It has been argued that MMSE is unlikely to be sensitive enough as a predictor of driving outcomes in non-impaired older drivers (Crizzle et al., 2012). In contrast, Matas et al. (2014) found that MMSE independently predicted UFOV™ performance in older adults and argued that it may be sensitive to very mild cognitive decline relevant to driving ability. The MMSE was included in the present study to test this argument, and as a screen for study inclusion (participants were required to be free from cognitive impairment). We expected that MMSE scores would be high in the present sample of volunteers.

The Useful Field of View Test™ (UFOV™) is a driver screening test that has been consistently reported to predict driving outcomes, including retrospective and prospective crashes, on-road driving performance, and simulated driving performance (Clay et al., 2005; Gentzler & Smither, 2012; Mathias & Lucas, 2009). The UFOV™ is a computerised test consisting of three (or less frequently, four) subtests assessing aspects of visual attention and processing speed (Ball & Owsley, 1993; Wood & Owsley, 2014). Of the three subtests, UFOV™ Subtest 2, which reflects the efficiency of focussed attention over a large spatial area (Matas et al., 2014) is most related to crashes (Ball et al., 2006; Owsley et al., 1998; Oxley et al., 2005), on-road driving (Bowers et al., 2013) and simulated driving performance (Molnar et al., 2007). Subtest 2 is therefore often considered on its own for purposes of brevity (Ball et al., 2006).

Inspection Time (IT) is a measure of processing speed that is related to a general "speediness" factor of intelligence (Nettelbeck, 2001). IT is notable for its high correlation with IQ (around 25% shared variance), and is related to other cognitive abilities including fluid reasoning and short-term memory (Grudnik & Kranzler, 2001). Furthermore, there is evidence that IT may be useful as a biomarker for general cognitive decline; IT generally lengthens with normal aging, and is associated with poorer future performance on a range of outcomes, including future cognitive ability and performance of everyday tasks (Deary et al., 2010; Gregory et al., 2008). These results suggest that IT may be useful for detecting older adults who are currently experiencing slowed processing speed and who may be at risk for future cognitive decline and increased driving problems. Processing speed is an important contributor to driving safety, and various tests of processing speed have consistently been associated with driving outcomes (Anstey et al., 2005; Lundberg, Hakamies-Blomqvist, Almkvist, & Johansson, 1998). IT has been shown to be significantly correlated with scores on the UFOV™ processing speed subtest (Burns et al., 2005), and in a small sample of healthy older drivers ($N = 14$), faster IT was significantly correlated with better performance on a simulated driving task (Matas et al., 2016).

The Proficiency of Peripheral Visual Processing (ProPerVis) Crowding task assesses crowding across the visual field and requires identifying briefly-presented flanked stimuli in central, peripheral, and parafoveal vision. Visual crowding is a phenomenon whereby visual interference occurs when a visually off-centre primary target is flanked by similar materials, resulting in difficulty identifying a crowded stimuli (Bouma, 1970; Levi, 2008; Pelli et al., 2004). Crowding has generally been considered a visual phenomenon but recent work suggests it may also involve limits to attentional resolution (He et al., 1996; Levi, 2008). Crowding influences the successful execution of everyday tasks, including reading, interacting with the environment, and driving because the visual scene is often cluttered (Whitney & Levi, 2011). In a cluttered scene, crowding does not affect the detection of an object; rather, it affects the identification of an object, making it appear indistinct or jumbled with surrounding objects (Levi, 2008). The effect of crowding increases with age, such that older adults find it much more difficult to extract information from a crowded visual field, especially when the stimuli of interest are presented in peripheral vision (Burns & Nettelbeck, 2003; Werner, 2008). ProPerVis Crowding has been reported to be significantly correlated with UFOV™ Subtests 2 and 3 (Burns et al., 2005; Matas et al., 2014), and in a recent study with a small sample of older drivers, ProPerVis Crowding tended to be associated with performance on a simulated driving task (Matas et al., 2016).

Along with cognitive abilities, the role of vision in predicting driving performance and crash risk has been extensively studied. Most countries, including Australia, require drivers to meet nominal visual acuity standards (typically logMAR 0.3) to hold a driver's license (Austroads, 2012; Bron et al., 2010). The Austroads guidelines for assessing fitness to drive also specify minimum standards for visual fields, and recommend use of contrast sensitivity and other specialised vision tests where appropriate. However, the evidence linking visual function with crash risk is equivocal. In particular, reviews indicate that visual acuity demonstrates only a very weak correlation with crash risk (Owsley & McGwin Jr, 2010; Wood, 2002). Similarly, contrast sensitivity is only weakly associated with driving performance (Anstey et al., 2005). Despite this, low-level visual function has a high level of face validity for the driving task, and adequate visual function is required in order to process the driving environment. For this reason, vision (visual acuity and contrast sensitivity) is included as a predictor in Anstey et al.'s (2005) Multifactorial Model of Driving Safety. In the present study, we assessed the correlation of visual acuity and contrast sensitivity with performance on a simulated driving task.

Finally, the driving task requires adequate physical function. It is intuitively recognisable that a satisfactory level of physical fitness is required to operate a vehicle safely; this is also reflected in Anstey et al.'s (2005) proposed Multi-Factorial Model of Driving Safety, where cognitive function, sensory function, and physical function all contribute to the Capacity to Drive Safely. The Australian guidelines for assessing fitness to drive (Austroads, 2012) emphasise that the driving tasks relies on both sensory input (e.g. vision and hearing) and musculoskeletal actions requiring muscle power and coordination, among other factors.

Physical activity and fitness may also contribute to driving safety through an association with cognitive functions relevant for driving. A study by Marmeleira et al. (2012) showed that self-reported physical activity, as measured by the International Physical Activity Questionnaire, was associated with better performance on tests of processing speed and divided attention. Similarly, Roth et al. (2003) reported that two measures of self-reported physical activity, the Exercise Participation Questionnaire and the Physical Activity Scale for the Elderly (PASE), were both significantly correlated with UFOV™ performance. Divided attention and processing speed, especially UFOV™, are known to be related to driving safety and crash risk (Anstey et al., 2005; Clay et al., 2005; Mathias & Lucas, 2009). Meta-analyses and reviews have generally indicated that physical activity is positively associated with brain function and cognitive ability in older adults (Bherer et al., 2013; Erickson et al., 2015; Marmeleira, 2012; Stine-Morrow & Basak, 2011), and that physical activity interventions for older adults (particularly aerobic or combined aerobic strength interventions) can have a positive effect on cognitive performance across several domains (Carvalho et al., 2014; Colcombe & Kramer, 2003). However, in contrast, two recent reviews and meta-analyses of Randomised Control Trials have failed to find consistent evidence of cognitive benefit from physical exercise (Kelly et al., 2014; J. Young et al., 2015).

Few studies have investigated the direct relationship between physical activity and driving. Marottoli et al. (2007) conducted a randomised control trial to investigate if participation in a physical conditioning program could improve driving performance in older drivers with physical impairments such as neck rotation, trunk rotation, or gait speed. The intervention targeted flexibility, coordination, and speed of movement. The results showed that, compared to the control group, the intervention group were able to maintain their driving performance on an on-road driving assessment, while the driving performance of the control group declined (Marmeleira, De Melo, Tlemcani, & Godinho, 2011). In a pilot study Caragata et al. (2009) found that older drivers who participated in a fitness program showed gains in physical function measures, and also reported that they felt their driving skills had improved and their driving confidence was higher. These results indicate that physical activity may be an important factor in predicting safe driving. To investigate further the link between physical activity and driving, we correlated self-reported physical fitness, as measured by the PASE (a validated physical activity questionnaire specifically developed for older adults) with performance on simulated driving tasks.

Therefore, the aim of the present study was to investigate the contribution of cognitive, visual, and physical activity factors to driving task performance in older adults. We also aimed to determine the best combination of these variables for predicting driving performance. To measure driving performance, two simulated driving tasks were selected: a Brake RT Task, and a Traffic Participation Task. Driving simulators are now frequently used for research and assessment, and they provide a safe, cost effective alternative to on-road driving tests. In general, driving simulators have demonstrated validity across a number of criteria, including comparisons with on-road driving (Engström et al., 2005; Lee, Cameron, & Lee, 2003; Mayhew et al., 2011), predictive validity for crash outcomes (Hoffman & McDowd, 2010), and discriminant validity across driver experience levels (Mayhew et al., 2011). Importantly, driving simulator validity is dependent on the specific simulator and tasks, and preliminary research with the simulator and tasks used in the present study suggested that the simulator and tasks demonstrated evidence of content and convergent validity, and appeared to be appropriate for assessing driving performance in older adults (Matas et al., 2016).

Based on previous results, it was hypothesised that:

- Better performance on cognitive ability measures reflecting aspects of attention and speed of visual processing (IT, UFOV™ Subtest 2, and Crowding) is associated with better performance on the simulated driving tasks
- Higher physical activity engagement, as measured by the PASE self-report measure, is associated with better performance on the simulated driving tasks
- The cognitive ability measures are more strongly associated with performance on the simulated driving tasks than the low-level vision measures (visual acuity and contrast sensitivity)

Method

Participants

Participants were aged 65 years or over, living independently in the community, held a valid Australian driver's license, drove at least once per week, had normal or corrected-to-normal vision, and reported good physical and mental health. Participants were excluded if they reported a history of epilepsy or seizure-related conditions, a diagnosis of cognitive impairment, or a history of motion sickness or Simulator Sickness (SS). Participants were recruited from a database of older adults who had previously participated in, or expressed interest in participating in driving research. Participants were also recruited through flyers and announcements at various community groups, and through snowball sampling. Interested potential participants were directed to a website containing further information about the study and a preliminary questionnaire assessing their eligibility for the study.

The online questionnaire was accessed by 92 individuals. Of these, 82 were able to attend the on-campus test session. Those screened out were: $n = 3$ poor health; $n = 3$

subsequently unavailable to attend; $n = 2$ cancelled due to illness; $n = 1$ cancelled due to personal commitments; $n = 1$ subsequently uncontactable.

These 82 participants (54 males, 28 females; age range 65 to 84 years, mean age 73 years (SD 4 years)) who attended testing were considered for analysis in the present study. Participants drove an average of approximately 10,000 km per year (SD 6663 km, range 200 to 33000 km) and approximately 7 hours per week (SD 6 hours, range 0.5 to 35 hours). There were no significant differences between those who attended the study and those who did not in age, education level, self-rated health, number of medical conditions, motion sickness history, or self-rated driving skill level.

Of the 82 participants who attended, $n = 80$ attempted the driving tasks ($n = 1$ dropped out prior to attempting driving due to inadequate vision correction; $n = 1$ dropped out prior to attempting driving due to health concerns). Of the 80 participants who attempted the simulated driving tasks, $n = 73$ completed the Brake RT Task, and $n = 54$ completed the Brake RT Task and the Traffic Task ($n = 26$ [33%] dropped out due to SS). However, due to a technical error, driving performance data for $n = 1$ participant on the Traffic Task were not recorded, leaving $n = 53$ participants with complete data for the Traffic Task. One participant did not complete UFOV™ Subtest 2 due to difficulty understanding task requirements; this participant subsequently dropped out during the Brake RT Task. One participant did not bring adequate vision correction and was unable to complete the Crowding Task or the vision measures, and did not attempt driving.

Materials

Pre-test questionnaire. The online questionnaire included items to check participant eligibility for the study (age, driving status, health status, motion sickness history, and simulator sickness history); and the physical activity scale.

Physical Activity Scale for the Elderly (PASE). The PASE is a brief physical activity questionnaire designed specifically for the elderly (Washburn et al., 1993). The questionnaire takes approximately five minutes to complete and asks older adults to recall their physical activity over the past seven days, in categories of leisure time/recreational activity, household activity, and work-related activity. Respondents report if they participated in the listed activities during the past seven days. For the leisure time and recreational activities, participants indicated how many days out of seven they had engaged in that activity, and how many hours per day on average they engaged in the activity. For household activities, participants indicated either “yes” or “no” to having engaged in each listed activity at any time during the past seven days. For work related activities, participants indicated either “yes” or “no” to having engaged in work related activities (for pay or as a volunteer) at any time during the past seven days. If yes, they indicated the number of hours worked during the past seven days, and the type of work they did (mainly sitting, sitting or standing with some walking, walking with some material handling, walking and heavy manual work). Scores are calculated based on the level of engagement in each of the listed activities; each item is weighted according to its contribution to overall physical activity (Washburn et al., 1993). Higher scores indicate higher physical activity.

Mini Mental State Examination (MMSE). The MMSE is a short test of global cognitive function used to estimate the extent of any cognitive impairment (Folstein et al., 1975). It was administered according to standardised instructions and scoring; the standardized MMSE is a valid and reliable tool for assessing cognitive impairment in older adults (Molloy & Standish, 1997). Participants answer questions assessing short-term memory, orientation in space and time, constructional ability, executive functioning, and ability to follow instructions. Scoring is out of 30 with lower scores indicating greater impairment. Participants were required to score 24 or above on the MMSE to proceed with participation. All participants met this requirement.

Useful Field of View Test™ (UFOV™) Subtest 2. UFOV™ is a computer-based test of visual attention and processing speed involving detection and localisation of targets briefly presented throughout the visual field (Ball & Owsley, 1993; Wood & Owsley, 2014). Only the second subtest, divided attention, was administered; this subtest has been found to be most correlated with UFOV™ Total score and to best predict driving outcomes (Ball et al., 2006; J. D. Edwards et al., 2006; Owsley et al., 1998). For the divided attention subtest, one of two stimuli (either a car or a truck) appears briefly in central vision, and the car appears simultaneously in the periphery at one of the eight cardinal or intercardinal points, 12.5 cm from the centre of the display. Participants indicate which object was presented in central vision, and the location of the car in the periphery. Test-retest reliability for UFOV™ Subtest 2 is reportedly .82 (J. D. Edwards, Vance, et al., 2005). Score is exposure time (ms) for which 75% of trials were answered correctly; lower scores indicate better performance.

Inspection Time (IT). IT measures processing speed (Burns & Nettelbeck, 2003). Two high-contrast lines, one markedly shorter than the other, appear as a target on a computer screen. Participants indicate whether the shorter line was located left or right of a focal point. Time available for processing is limited by a backward masking procedure and reduced or extended using an adaptive staircase algorithm according to response accuracy. Targets are preceded by a warning cue ("+" in the centre of the screen, 370ms) and are immediately followed by a mask figure shaped like two lightning bolts. First, there were three sets of 10 practice trials with decreasing target presentation time for each set (835ms, 420ms, and 250ms) that required answering 10/10 items correctly for the first and second set, and 9/10 items for the third set. IT was measured in ms as the duration between target onset and mask onset at which the viewer achieved 79% accuracy. Test-retest reliability for adults is .81 (Grudnik & Kranzler, 2001).

ProPerVis Crowding. ProPerVis assesses visual processing of briefly presented stimuli across the visual field on a computer screen (Burns et al., 2005). The stimuli are a four-square parent figure and six figures derived from it, resembling stylised characters M, E, W, 3, 5, 2. On each trial, one of the six figures is presented, flanked on either side by the parent figure. The target and flankers appear randomly in one of five lateral positions on the screen. Participants attempt to identify which of the six figures was presented. The outcome was total errors made across the five positions from 40 trials; lower scores indicate better performance.

Visual acuity. Visual acuity was assessed using the Freiburg Visual Acuity Test (Bach, 1996). Participants indicated the orientation of a "Tumbling E" figure. Participants sat 1.65m from a computer screen and responded using the arrows on the computer keyboard. The size of the "Tumbling E" figure varied on each trial depending on the current estimated acuity threshold of the participant, calculated via the best Parameter

Estimate by Sequential Testing procedure (Bach, 1996). There were 30 trials, with every sixth trial being an "easy" trial where the "Tumbling E" figure size was significantly larger than the current estimated threshold. Acuity was recorded as the logarithm of the Minimum Angle of Resolution (logMAR), with lower scores representing better visual acuity.

Contrast sensitivity. A Pelli-Robson Contrast Sensitivity wall chart (Pelli et al., 1988) was used to measure contrast sensitivity. The chart consists of eight rows of six letters, arranged into groups of three letters. The contrast of the letters reduces for each subsequent triplet. Participants viewed the chart at eye level from a distance of 1m while wearing normal vision correction. Participants were instructed to read the letters on the chart as far down as they could and were encouraged to guess if they were uncertain. The outcome was log contrast sensitivity, which was measured according to the faintest group of three letters for which the participant correctly identified two of the three letters. Possible scores range from 0 to 2.25, with higher scores indicating better contrast sensitivity.

Mini Simulator Sickness Questionnaire (SSQ). The Mini-SSQ (Mourant et al., 2007) contains six questions and is a short form of the Simulator Sickness Questionnaire (Kennedy et al., 1993). The Mini-SSQ is quick to administer, suitable for repeated administration, and sensitive to changes in driving conditions and changes in Simulator Sickness (SS) symptoms over time (Matas et al., 2015; Mourant et al., 2007). Six symptoms are included: general physical discomfort, headache, blurred vision, sweating, faintness/dizziness, and stomach discomfort/nausea. Possible responses were on a four-point scale, corresponding to None (0), Slight (1), Moderate (2), Severe (3). Scores for each symptom were summed to give a total score. Mini-SSQ scores were recorded after the Practice Drive (SSQ1), after the Brake RT Task (SSQ2), and after the Traffic Task (SSQ3).

Driving simulator. The simulator used was a SimWorx SX06DTS-L Driver Training Cockpit (Figure 8.1). The simulator has a fixed base and is located in the laboratory. The simulator has a three screen display (3×27in HD monitors, total resolution 5760×1080 pixels). The simulated horizontal field of view spans a total visual angle of 210°, but this can be accessed in its entirety only via head movement, or eye movement, or both. Thus, it is analogous to viewing through the front and side windows of a motor vehicle. The graphics were texture-rich. All scenarios required driving on the left hand side of the road, consistent with Australian road rules. The simulator has a driver cockpit including force feedback steering wheel, pedals, gearshift, handbrake, ignition switch, engine start button, and adjustable seat. The steering is controlled by a Logitech G27 unit modified with a steering wheel of 350mm diameter. The steering wheel and pedals were calibrated with the Logitech Profiler and the configuration menu on the simulator. The vehicle dynamics model is linear and computes the roll, pitch, yaw, longitudinal and lateral speed, and acceleration, based on vehicle and environmental forces. The handling is designed to resemble that of a 4-cylinder hatchback. Verbal instructions and feedback are provided to the driver. On-screen display elements include rear-view and side mirrors, a dashboard, and navigational assistance (arrows indicating the direction of the next turn, where required).

The simulator is programmed to provide summary data for each driving task. Raw driving data are not provided. The sequence and content of all driving tasks were identical for all participants. There were three simulated driving tasks:

- *Practice Drive.* The aim of the practice drive was for drivers to familiarise themselves with the simulator. Performance on the practice drive was not assessed. The speed of the vehicle was automatically controlled by the simulator, and the driver had control of the steering wheel, indicators, and some control of the brake (the simulator was programmed to automatically slow down to an appropriate speed on approach to corners, and the driver could brake at other times, though this was not necessary to complete the task successfully). The vehicle initially accelerated to 60km/h, and gradually increased the speed to 100km/h during the task. The driver was required to steer the vehicle on bends and corners. Task duration was approximately 8 minutes.

- *Brake Reaction Time (RT)* This task assessed Brake RT to an on-screen stimulus. Drivers accelerate to a target speed (90km/h) along a straight road. After the target speed has been reached, a stop sign randomly appears on the centre display. Drivers must brake as quickly as possible and bring the vehicle to a complete stop. This is repeated five times per trial. The simulator recorded Brake RT, calculated as time elapsed between the appearance of the stop sign and any application of the brake pedal. The outcome was Average Brake RT, calculated as the mean Brake RT from the five test trials. Drivers completed one practice run and one test run, with a brief break in between.

- *Traffic Task.* This task is set in a small city including intersections, traffic, pedestrians, and some hazards. Directions are provided by the simulator. The speed limit is 50km/h. The course contains various intersections and turns, including traffic lights, give way signs, stop signs, and pedestrian crossings. The route, traffic, and content of the scene were programmed to be identical for all drivers. The Traffic Task has been associated with SS in previous studies (Study 3 and Study 4); however, it was retained in the present study because it was considered to be most relevant task to the study aims of the available tasks pre-programmed on the simulator. Due to the risk of SS, participants were carefully screened prior to participation and were monitored throughout the study. Performance was automatically scored and reported for eight grade categories: General, Lane Position, Speed Control, Steering, Signs, Car Following, Priority, and Signalling. The maximum score for each grade category was 10 points. The Total Score was the sum of scores from each grade category (maximum 80). An Overall Grade was also given for the task, which was the minimum score recorded from the eight grade categories (e.g. if a driver received a Grade of 5 for General and a Grade of 8 for all other categories, the Overall Grade would be 5). Higher scores indicate better performance. Items in each grade category are shown in Table 8.1; points were deducted for each instance of an error. For example, speeding results in a 1.5 point deduction; if a driver incurs three speeding infractions, 4.5 points are deducted. The weighting of points is based on the seriousness of the error. Number of accidents and route errors were recorded. Duration is approximately 10 minutes. An example scene from this task is shown in Figure 8.2.

Table 8.1: Scoring for Traffic Participation task

Grade Category	Items (Points deducted per infraction in square brackets)
General	Accident [5], Engine stalled [1.75], Route error [1], Driving off with handbrake on [1], Handbrake on while driving [1]
Lane Position	Driving on the wrong side of the road [3.5], Driving in the wrong lane [2.5], Stopping too far from the stop line [2.5], Failure to keep left when not overtaking [1.5], Driving too close to the line [1]
Speed Control	Cornering too fast [2.5], Speeding [1.5], Driving too fast while approaching intersection [1.5], Braking too suddenly [1]
Steering	Driving on pavement [1.5], Driving on hard shoulder [1.5], Driving too much to the right [1.5], Driving off the road on the left [1.5], Cutting corner on intersection [1.5], Swinging wide on intersection [1.5]
Signs	Ignored red traffic light [5], No-entry sign ignored [2.5], Stop sign ignored [2.5], Stopped on pedestrian crossing [2.5], Stopped in intersection [2], Ignored yellow traffic light [1], Waited too long for green traffic light [1]
Car Following	Too close to car in front [1.5]
Priority	Failure to give way [3.5]
Signalling	Did not indicate early enough [1.5], Indicator switched off too soon [0.5], Indicating without a valid reason [0.5],
Lane Changing	Cut in while changing lane [5], Did not indicate correctly while changing lane [1.5], Driving too close to car in front while changing lane [1.5]

Note. Each Grade Category has a maximum score of 10. Points are deducted in each Grade Category for each instance of an error in that category (points displayed in square brackets).

Figure 8.1: Driving simulator

Figure 8.2: Example scene from the Traffic Participation task

Procedure

The study was approved by The University of Adelaide Human Research Ethics Committee. All participants provided written consent and were provided with digital and hard copies of the study information sheet.

Test sessions were held in the laboratory, but participants were asked to complete the online pre-test questionnaire before coming in for testing. The order and content of tasks was identical for all participants.

Participants first completed the cognitive and visual assessments in the following order: MMSE, IT, ProPerVis Crowding, UFOV™ Subtest 2, Freiburg Visual Acuity Test, and Pelli-Robson Contrast Sensitivity Chart. Cognitive and visual testing took approximately 30 minutes. Participants then had a compulsory break of approximately 5 minutes, which was extended if needed.

Participants then completed the three driving tasks in the following order: Practice Drive, Brake RT, and Traffic Task. All participants used automatic transmission.

After each task, participants completed the Mini-SSQ via tablet computer. Participants were monitored by the experimenter throughout for signs of SS, and were told to alert the experimenter and stop driving if they wished to discontinue driving for any reason. Between each driving task, participants completed the SSQ and had a break for as long as they needed (typically $1-2$ minutes between Practice Drive and Brake RT, and 5 minutes between Brake RT and Traffic Task, during which participants were required to move off of the simulator). A fan provided ventilation and room temperature was approximately 23°C. Participants were provided with cold water.

After all driving tasks had been completed, participants remained in the laboratory for approximately 10 minutes while they were monitored for symptoms of SS and completed a feedback questionnaire via tablet computer. The protocol took up to 1.5 hours to complete.

Results

Analysis

Data were analysed using R version 3.2.4 (R Core Team, 2016). In regards to driving performance measures, only participants who completed the driving tasks were included in analyses (i.e. $n = 73$ for the Brake RT task and $n = 53$ for the Traffic Task).

Descriptive statistics

Table 8.2 displays descriptive statistics for cognitive, visual, physical activity, and driving performance measures. As registered by the screening questionnaire and PASE, participants were generally healthy and fit. Participants reported a median of one medical condition (range $0-6$), and a median Self-Rated Health of "Very Good". PASE scores indicated that the sample was very physically active. The greatest contribution to PASE scores in our sample, according to Washburn et al.'s (1993) item weights, were yard work, housework, and walking and cycling. MMSE scores indicated that all participants were in the normal range (scores below 24 indicate possible impairment; all participants in the present sample scored between 26 and 30; 98.5% scored 28 or above). Fifty-one percent of the participants achieved the maximum score of 30.

IT, Crowding, and UFOV™ Subtest 2 scores were in the normal range. IT and

Crowding Total scores were very similar to those reported previously in a sample with similar characteristics (Crowding Total mean 13.1 errors, IT mean 66.4ms; Matas et al., 2014). UFOV™ Subtest 2 performance was better than reported previously; Matas et al. (2014) reported a mean of 115.5ms for this subtest, and Edwards et al. (2006) reported norms of between 84.14ms (age $65 - 69$, > 12 years education) and 232.00ms (age $80 - 84$, < 12 years education). Even considering that our sample was relatively young (mean age 72.93) and well-educated (84% had post-secondary school education), UFOV™ Subtest 2 score was still quicker than expected. J. D. Edwards et al. (2006) reported a norm of 102.80ms (SD = 97.48, $N = 534$) for people aged $70 - 74$ with more than 12 years education. The difference between our sample and this norm is highly significant, $t(613) = 3.41$, $p < .001$.

Regarding vision, 10 participants (12%) had corrected visual acuity worse than logMAR 0.3, the minimum standard required to hold a driver's license in Australia and many other countries. All participants had log contrast sensitivity of 1.65 or above, indicating no contrast sensitivity impairment in the sample.

The average Brake RT of 0.61s corresponded well to previous performance reported for this task (Matas et al., 2016) and for Brake RTs reported for similar simulated driving tasks (Hollis et al., 2013; Martin et al., 2010; Zhang et al., 2007).

Traffic Task Total Score suggested good overall performance, with a median score of 71.5 out of 80. However, the median score for Traffic Grade was 5 out of 10. Fifteen participants had a Traffic Grade of less than 5. This indicates that the average participant achieved a poor score in at least one Grade Category. As seen in Table 8.2, the worst Grade Category was Lane Position with a mean score of 7.63 out of 10 (median 7.5), and 7 participants scoring less than 5 out of 10. The best categories were Car Following, Priority, and Signalling, with no participants scoring less than 5 out of 10 for these categories.

Simulator experience

Of the 82 participants who attended the study, $n = 24$ reported prior simulator experience. There was no significant difference between participants with simulator experience and participants without simulator experience on any of the variables of interest (cognitive, visual, physical activity, and driving performance measures). There was also no significant differences in Mini-SSQ scores or dropout.

Table 8.2: Descriptive statistics for cognitive, visual, physical activity, and driving performance measures

Variable	N	Mean	SD	Range	Fail Grade (< 5)
Age	82	72.9	4.4	65 to 84	
PASE	82	147.0	56.2	10.46 to 350.04	
Visual Acuity (logMAR	81	0.1	0.13	−0.12 to 0.43	
Contrast Sensitivity (log contrast)	81	1.86	0.13	1.65 to 1.95	
MMSE	82	29.3	0.8	26 to 30	
IT (ms)	82	61.5	18.1	30.25 to 147.13	
Crowding Total	81	12.2	4.9	2 to 24	
UFOV™ Subtest 2 (ms)	81	65.0	55.3	17 to 250	
Brake RT (s)	73	0.61	0.12	0.34 to 1.06	
Traffic Task Total Score	53	70.09	6.54	52 to 80	
Traffic Task Grade	53	5.25	2.67	0 to 10	15
General	53	8.34	2.56	0 to 10	4
Lane Position	53	7.63	2.89	0 to 10	7
Speed Control	53	8.51	2.28	0 to 10	3
Steering	53	8.42	1.73	4 to 10	1
Signs	53	8.24	1.99	0.5 to 10	2
Car Following	53	9.97	0.21	8.50 to 10	0
Priority	53	9.74	0.75	7.67 to 10	0
Signalling	53	9.25	1.02	6 to 10	0

Simulator sickness and dropout

SS was measured using the Mini-SSQ and was assessed after the Practice Drive (SSQ1), after the Brake RT assessment (SSQ2) and after the Traffic Task (SSQ3). Eighty participants attempted the driving tasks. Three participants (4%) dropped out during or immediately after the Practice Drive. A further 8 participants (10%) dropped out during or immediately after the Brake RT assessment, and a further 15 (19%) participants dropped out during the Traffic Task (total dropouts: $n = 26$, 33%). Descriptive statistics for SSQ scores are shown in Table 8.3.

Table 8.3: Descriptive statistics for SSQ scores

SS Measure	N	Mean	Mdn	SD	Range
SSQ1	80	1.27	1	1.44	0 to 6
Complete	53	0.91	1	1.20	0 to 5
Dropout	27	2.00	2	1.62	0 to 6
SSQ2	77	0.30	0	0.69	0 to 4
Complete	53	0.25	0	0.52	0 to 2
Dropout	24	0.42	0	0.97	0 to 4
SSQ3	67	2.40	2	2.75	0 to 10
Complete	52[a]	1.93	1	2.51	0 to 10
Dropout	15[a]	3.93	3	3.08	0 to 9

[a] SSQ3 Scores for one completer and one dropout were not recorded

The Brake RT task (SSQ2) was associated with the fewest SS symptoms, and the Traffic Task (SSQ3) was associated with most SS; paired samples t-tests indicated that SSQ2 was significantly lower than SSQ1 ($t(76) = 5.705$, $p < .001$) and SSQ3 ($t(66) = 6.644$, $p < .001$), and SSQ3 was significantly higher than SSQ1 ($t(66) = 4.58$, $p < .001$). Welch independent samples t-tests indicated that dropouts reported significantly higher SS scores than completers for SSQ1 ($t(40.95) = 3.109$, $p = .003$) and SSQ3 ($t(19.685) = 2.270$, $p = .035$), but not SSQ2 ($t(28.992) = 0.811$, $p = .424$).

Overall, SSQ scores were low; 1 point reflects "slight" experience of one symptom, 2 points reflects "moderate experience" of one symptom (or "slight" experience of 2 symptoms), and 3 points reflects "severe" experience one symptom (or a combination of "slight" or "moderate" experience of two or more conditions). "General discomfort" was the most commonly reported symptom for all three SS measurements, accounting for approximately 30% of the total SS score. "Faintness/dizziness" and "Blurred vision" were the next most commonly reported symptoms, accounting for approximately 21% and 20% of the total SS score, respectively.

Dropout, driving performance, and cognitive function. Cognitive and driving performance for completers and dropouts is shown in Table 8.4. Welch independent samples t-tests showed that there were no significant differences in performance between participants who completed the driving tasks and participants who dropped out due to simulator sickness.

Table 8.4: Comparison between dropouts and completers on cognitive test performance and driving performance

	Completers		Dropouts			
Variable	N	M (SD)	N	M (SD)	t (df)	p
MMSE	54	29.33 (0.93)	26	29.31 (0.62)	0.15 (70.1)	.88
IT	54	61.96 (18.56)	26	62.69 (16.27)	0.18 (55.8)	.86
UFOV™ Subtest 2	54	62.8 (54.32)	25[a]	71.76 (59.34)	0.64 (43.3)	.52
Crowding	54	12.46 (4.64)	26	11.77 (5.25)	0.57 (44.4)	.57
Brake RT	54	0.60 (0.11)	19[b]	0.65 (0.15)	1.18 (25.0)	.25

Note. [a] One participant did not complete UFOV™ Subtest 2 due to difficulty understanding the task requirements. [b] $n = 19$ participants completed the Brake RT Task, but dropped out prior to completing the Traffic Task.

Relationship between predictor variables and driving outcomes

The Shapiro-Wilk test indicated that two of the outcome measures (Traffic Task Total and Traffic Task Grade) were not normally distributed, therefore the relationships between variables were assessed using Spearman correlation coefficients. The results are shown in Figure 8.3. Age was not significantly associated with performance on the simulated driving tasks, or to performance on the cognitive assessments.

Cognitive measures. It was hypothesised that better performance on the driving assessments is related to better performance on UFOV™ Subtest 2, IT, and Crowding. Consistent with this, faster processing speed (IT and UFOV™ Subtest 2) and fewer crowding errors showed significant small correlations with better performance on the Brake RT test. There was a small correlation between fewer crowding errors and better performance on the Traffic Task (higher Traffic Total and Traffic Grade), although only the correlation with Traffic Grade reached statistical significance. There were small correlations between IT and UFOV™ Subtest 2 and the Traffic Task, which were not statistically significant. There was a small significant correlation between MMSE and Brake RT, such that higher scores on the MMSE were associated with faster Brake RT. There was no association between MMSE and performance on the Traffic Task.

Physical activity. It was hypothesised that higher PASE scores are associated with better driving performance. The Spearman correlation coefficients (see Figure 8.3) indicated non-significant weak-to-moderate negative correlations between PASE scores and Brake RT, Traffic Grade, and Traffic Total. Higher PASE scores were associated with faster Brake RT, as expected, but poorer performance of the Traffic Task, opposite to our prediction.

Visual measures. The Spearman correlation coefficients (Figure 8.3) showed a weak-moderate positive correlation between contrast sensitivity and performance of the Traffic Task (Total Score and Grade), although these correlations did not reach statistical significance. Contrast sensitivity was not associated with Brake RT performance. Visual acuity showed very weak, non-statistically significant associations with Brake RT and performance on the Traffic Task.

Predictors of driving performance

We used linear regression to determine the best combination of predictors to predict Brake RT performance. Models were not constructed for Traffic Total and Traffic Grade

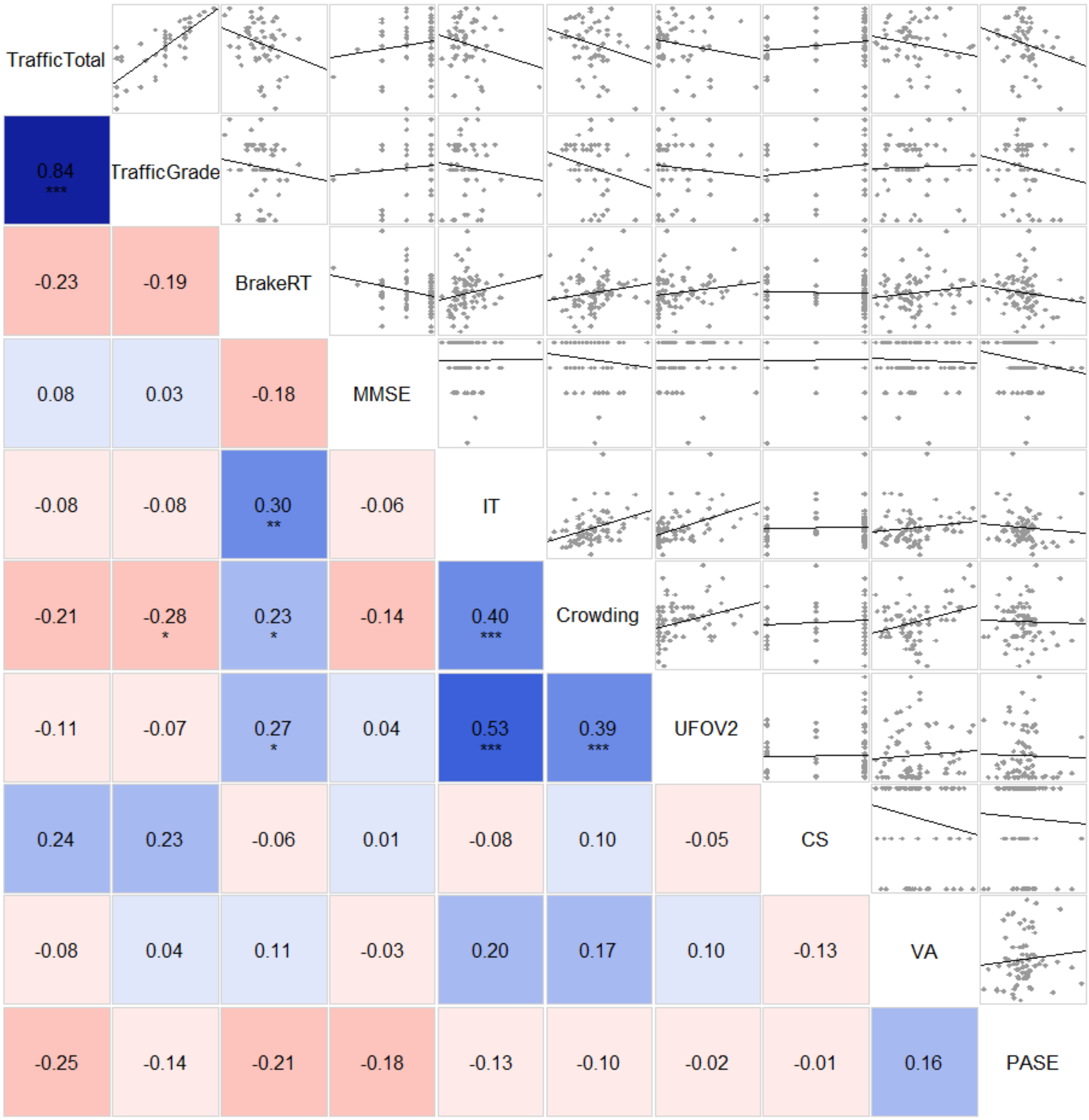

Figure 8.3: Correlations between driving performance measures and predictor variables (cognitive, visual, and physical activity measures). The lower diagonal displays the Spearman correlation between the variables (* $p < .05$, ** $p < .01$, *** $p < .001$). The upper diagonal shows scatterplots of the relationships between variables with a line of best fit. Figure produced using R package `corrgram` (Wright, 2015).

due to the lack of significant bivariate correlations between the predictor variables and the outcome measures from the Traffic Task.

Models were selected via backward elimination from an initial full model including MMSE, IT, UFOV™ Subtest 2, Crowding Total, PASE Total, Visual Acuity, and Contrast Sensitivity. At each step, a single predictor was eliminated based on the largest reduction in the Akaike Information Criterion (AIC; Akaike, 1974). Elimination ceased when removal of additional predictors no longer reduced the AIC and therefore did not improve the goodness-of-fit of the model (R package `car`; Fox & Weisberg, 2011). The relative importance of each predictor in the final models was calculated using R package `relaimpo` (Grömping, 2006). Relative importance is used to determine the contribution of each variable to the model, taking into account the direct effect, and the effect when combined with the other variables in the regression equation (J. W. Johnson & LeBreton, 2004). We used the LMG method for calculating relative importance; the LMG method calculates the contribution of each variable to the model across all possible orderings of regressors.

The selected models are shown in Table 8.5. Regression assumptions were checked using global validation of linear model assumptions (R package `gvlma`; Peña & Slate, 2006). Assumptions were initially not satisfied for all of the Brake RT models. Inspection of the models revealed the presence of an influential outlier case, a participant with a very high Brake RT score (1.056s). When this case was removed, the parameters for the models remained approximately equal, and the assumptions for the models were met.

For Brake RT, the backward selection model selected from the initial full model included PASE, UFOV™ Subtest 2, and MMSE as predictors, and all three predictors were statistically significant. The model explained 14.2% of the variance in Brake RT, $F(3, 68) = 3.752$, $p = .015$. Relative importance analysis (LMG) indicated that, of the total R^2, UFOV™ Subtest 2 accounted for 36% of variance explained, MMSE accounted for 33%, and PASE accounted for 31%. The predictors were related to the outcome in the expected direction, that is higher PASE scores, higher MMSE score, and faster UFOV™ performance were related to faster Brake RT.

The analysis was then repeated without two participants who recorded low scores on the MMSE, relative to the rest of the sample; it was suspected that these two cases were influencing the relationship between Brake RT and MMSE. These two participants scored 26 and 27 out of 30, compared to the rest of the participants who scored 28 or above. When these two cases were removed, MMSE no longer contributed to the model. The model containing PASE and UFOV™ Subtest 2 was statistically significant overall ($F(2, 67) = 3.201$, $p = .047$) and explained 8.72% of the variance in Brake RT, but neither of the predictors reached significance (PASE: $t(67) = -1.703$, $p = .093$; UFOV™ Subtest 2: $t(67) = 1.841$, $p = .070$).

Finally, we repeated the analysis using only those predictors that were significantly individually correlated with Brake RT (IT, UFOV™ Subtest 2, and Crowding). A parsimonious model containing only IT was identified; this model was statistically significant and explained 5.39% of the variance in Brake RT, $F(1, 70) = 3.984$, $p = .040$.

Discussion

Correlation between functional predictors and driving outcomes

We aimed to investigate the relationship of several cognitive, visual, and physical activity factors to performance on two simulated driving tasks in a sample of healthy older adults. We also aimed to determine the best combination of these variables for predicting driving performance. The selected variables were theoretically linked to driving performance in older adults, for example through the Multifactorial Model of Driving Safety (Anstey et al., 2005), which specifies that safe driving capacity depends on cognitive, visual, and physical function, and numerous reviews confirming the relationship between various cognitive abilities and driving outcomes (e.g. Anstey et al., 2005; Janke, 1994; Mathias & Lucas, 2009) We measured driving performance using two simulated driving tasks: a Brake RT test, and a Traffic Task involving negotiating traffic in a city environment.

Cognitive ability. Of our included cognitive assessments (MMSE, reflecting general cognitive function; UFOV™ Subtest 2, reflecting divided attention/efficiency of focussed attention; IT, assessing visual processing speed; and ProPerVis Crowding, assessing crowding across the visual field), the results indicated that IT, Crowding, and UFOV™ Subtest 2 were significantly correlated with performance on the Brake RT task, although the magnitude of all three correlations was low. Only Crowding was significantly associ-

Table 8.5: Regression of predictors on Brake RT

Model	R^2	B	Sig	SE B	95% CI B	β	lmg (% R^2)
Brake RT Full	.142		*				
Intercept		1.561	**	0.450	0.627, 2.460		
MMSE		-0.031	*	2.225×10^{-4}	-0.061, -0.001	-0.235	33%
PASE		-4.465×10^{-4}	*	2.238×10^{-4}	-8.904×10^{-4}, -2.540×10^{-6}	-0.229	31%
UFOV™ Subtest 2		4.507×10^{-4}	*	0.015	4.142×10^{-6}, 8.972×10^{-4}	0.226	36%
Brake RT No MMSE	.087		*				
Intercept		0.633	**	0.039	0.553, 0.711		
PASE		-3.801×10^{-4}		2.233×10^{-4}	-8.258×10^{-4}, 6.553×10^{-5}	-0.199	46%
UFOV™ Subtest 2		4.223×10^{-4}		2.293×10^{-4}	-3.551×10^{-5}, 8.800×10^{-4}	0.215	54%
Brake RT Cognitive	.054		*				
Intercept		0.517	**	0.047	0.423, 0.612		
IT		1.452×10^{-3}	*	7.275×10^{-4}	1.146×10^{-4}, 2.903×10^{-3}	0.232	100%

* $p < .05$, ** $p < .01$

ated with performance on the Traffic Task, showing a low correlation with Traffic Grade. The remaining (non-significant) correlations reflected very weak to weak associations in the expected direction.

The correlations indicated that faster IT, faster UFOV™ Subtest 2, and fewer crowding errors were associated with better performance on the simulated driving tasks (although only Crowding was associated with the Traffic Task). This is consistent with evidence that aspects of attention and processing speed are important for the driving task, and several studies reporting associations between visual attention and various measures of simulated driving performance (Alosco, Spitznagel, Cleveland, & Gunstad, 2013) (Foley et al., 2013; Mullen, Chattha, Weaver, & Bédard, 2008; S. W. Park et al., 2011) and processing speed and simulated driving performance (S. W. Park et al., 2011; Szlyk, Myers, Zhang, Wetzel, & Shapiro, 2002).

UFOV™ performance specifically is frequently reported to be predictive of simulated driving performance (Bélanger et al., 2010; Hoffman et al., 2005; Mullen et al., 2008; Rogé, Pébayle, Campagne, & Muzet, 2005). UFOV™ Subtest 2 is often used on its own, is predictive of crash outcomes, and is highly correlated with UFOV™ Total Score (Ball et al., 2006; Owsley et al., 1998). However, the results demonstrated no significant correlation between UFOV™ Subtest 2 and the Traffic Task, and only a small correlation between UFOV™ Subtest 2 and Brake RT. It is possible that this subtest was not sensitive enough in our sample; approximately 25% of the sample achieved a score of 17ms, the best possible score for this test, and overall performance was significantly quicker than the norm (J. D. Edwards et al., 2006). The more complex Subtest 3 may have been more appropriate for this sample; Subtest 3 requires localisation of stimuli amongst distractors across the visual field, and may reflect aspects of Crowding (Burns et al., 2005; Matas et al., 2014). UFOV™ Subtest 3 would likely have resulted in a broader range of scores, and may have been a more relevant assessment in relation to the complex environment and demands of the Traffic Task.

Crowding was the only predictor variable that was significantly correlated with the Traffic Task; crowding was weakly associated with Traffic Task Grade. Crowding affects the ability to identify objects in a cluttered scene; a crowded object will appear indistinct or jumbled (Levi, 2008). The Traffic Task took place in a complex city environment including buildings, pedestrians, traffic, and various roadside object. Therefore drivers who were more susceptible to the effects of crowding may have had difficulty processing and identifying important objects, such as other cars and pedestrians, in the crowded city scene. The effects of crowding increase with age, especially when the stimuli of interest are located in peripheral vision, which may explain its relevance to driving: drivers must focus on the road ahead, while also processing important peripheral information.

Visual function. The two visual assessments (contrast sensitivity and visual acuity) showed negligible to weak non-significant correlations with the driving outcome measures. This was despite the fact that 12% of the participants recorded visual acuity scores below the standard required to hold a driver's license in Australia (logMAR 0.3). This result is consistent with reviews indicating that low-level visual functions such as visual acuity and contrast sensitivity generally show no association, or are only very weakly associated, with driving outcomes such as crash risk, and are largely irrelevant for assessment of older drivers (Anstey et al., 2005; Owsley & McGwin Jr, 2010; Wood, 2002).

Physical activity. The physical activity measure (PASE self-report questionnaire) was non-significantly weakly correlated with the driving outcome measures ($r_s = -0.14$ to -0.25). Although the correlations were non-significant, the results were somewhat surprising in that higher PASE scores (indicating greater physical activity) tended to be associated with lower scores on the Traffic Task. On the other hand, the relationship between PASE and Brake RT was in the expected direction. This suggests that PASE reflected aspects of physical activity and fitness that may be beneficial for the lower-level demands of the Brake RT Task, but did not provide an advantage for the more complex Traffic Task.

Functional predictors of Brake RT

Linear regression models were used to determine the contribution of the functional predictors to Brake RT. The results showed that a combination of UFOV™ Subtest 2, MMSE, and PASE significantly explained 14.2% of the variance in Brake RT. All three predictors were significant and relative importance analysis indicated that the predictors explained approximately equal amounts of the total variance (UFOV™ 36%, MMSE 33%, PASE 31%). An explanation for this result is that the cognitive abilities (UFOV™ Subtest

2 and MMSE) enabled quick and efficient processing of the Brake RT stimulus, while the physical activity measure (PASE) reflected the physical speed of reaction once the stimulus had been processed.

The contributions of UFOV™ Subtest 2 and MMSE suggested that intact general cognitive function and efficient processing of the visual field were important for this task. UFOV™ Subtest 2, which reflects "divided attention" or efficiency of processing across the visual field, may have been particularly useful for this task because participants were required to remain vigilant to the Brake RT stimuli appearing in central vision, while also maintaining their speed and position on the road, requiring attention to the speedometer and lane markings. The environment for the Brake RT was graphically very sparse, consisting of flat countryside and no roadside distractions or other vehicles.

MMSE ranged from 26 to 30; that a relationship was observed in this narrow range indicates that even very minimal evidence of cognitive impairment impacted on Brake RT. There were two participants who recorded MMSE scores of 26 and 27, compared to the rest of the participants who scored 28 to 30. When the two low-scoring participants were removed from the analysis, MMSE no longer contributed to Brake RT performance; that is, for the participants scoring 28 or above on the MMSE, there was no association with Brake RT. The reduced model indicated that the combination of UFOV™ Subtest 2 and PASE significantly explained 8.72% of the variance on Brake RT. It is possible that the two low-scoring participants were representative of a real effect of MMSE on Brake RT; however, the analysis would need to be repeated with a broader range of scores to confirm this. If this is the case, it indicated that even small reductions in general cognitive functioning, which would still be considered non-impaired or in the "normal" range, may affect driving reaction speeds.

The contribution of PASE suggested that higher physical activity was associated with faster Brake RT. An explanation for this is that the PASE reflects aspects of physical activity and physical fitness that are beneficial for the lower-level demands of a simple RT task and enabled drivers to quickly react to the Brake RT stimuli. However, contrary to the literature suggesting a link between physical activity and cognitive ability (Bherer et al., 2013; Erickson et al., 2015; Marmeleira, 2012; Stine-Morrow & Basak, 2011), the results showed no significant association between PASE and the cognitive measures. This may be due to the characteristics of the sample of participants; the participants were very cognitively and physically fit and recorded above-average scores on the PASE and the cognitive measures. The relationship between physical activity and cognitive ability may only become evident when a broader range of scores are included. Despite this, the relationship between PASE and Brake RT shows that even in this high-functioning group of participants, physical activity was beneficial for speed of reaction.

Finally, a model was investigated using only the cognitive predictors that were significantly correlated with Brake RT (i.e. IT, Crowding, and UFOV™ Subtest 2). Of these, IT emerged as the best individual predictor of Brake RT. The IT-only model significantly explained 5.39% of the variance in Brake RT. IT is a measure of processing speed; drivers with faster IT were therefore able to process the Brake RT stimuli more quickly than those with slower IT. Processing speed is an important contributor to driving safety, and various tests of processing speed have consistently been associated with driving outcomes (Anstey et al., 2005; Lundberg et al., 1998). IT may be a particularly useful measure to include in driver screening procedures, because it is also a biomarker for general cognitive decline associated with aging (Deary et al., 2010; Gregory et al., 2008).

Limitations

There are a number of possible explanations for the small correlations and low explained variance observed in the present study. Firstly, our sample consisted of healthy, high-functioning older adults. Performance on the cognitive assessments was above average, and participants were generally very physically active. Ceiling effects were evident for UFOV™ Subtest 2 and the MMSE. It is possible that the high level of performance on the predictor variables meant that the range of scores observed was not sufficient to detect a relationship with the driving outcomes. If this is this case, it is consistent with the recommendation that license screening and assessment for older drivers targets only those who are already displaying evidence of functional impairment (Fildes et al., 2008; Langford, 2008; OECD, 2001).

Secondly, the driving simulator and driving tasks may not have been an appropriate measure of driving performance in this sample. Some evidence for the usability and acceptability of the simulator and tasks has been reported (Matas et al., 2016). However, further evidence is likely needed to determine the extent of the relationship between performance on the simulator, and real-world driving outcomes. Furthermore, scores on the simulated driving task were high, with the simulator recording few errors or violations for each participant. It is possible that the outcome measures generated by the simulator are not sensitive enough to detect the true variance in driving performance. An alternative explanation is that the driving outcome using in the present study is different to previous studies, and the simulator assessment is arguably a lesser assessment than the gold-standard on-road driving assessment.

A further problem encountered in the study was dropout due to SS. Although participants were screened to ensure they were fit, healthy, and had no history of SS or extensive history of motion sickness, 36% of the sample were unable to complete the driving tasks due to the onset of SS symptoms. Though symptoms were generally minor, our safety protocols dictated that participants displaying any signs of SS were unable to continue the task. The dropout rate was approximately equal to the dropout rate of 31% previously reported for this simulator with a similar sample, and is within the range reported in a review of SS in older drivers (between 35% and 75%, with an average around 40%; Trick & Caird, 2011). One concern with SS is that dropouts may differ systematically in their driving ability or cognitive function (Freund & Green, 2006); however, results in the present sample indicated that this was not the case, consistent with our findings from previous simulator studies (Matas et al., 2015, 2016).

There appeared to be some limitations with the use of PASE in the present sample of participants. The mean PASE total score was higher than has been reported previously, with one exception. Washburn et al. (1993) reported a mean total PASE Score of 102.9 (SD 64.1; age 65 – 100 [mean 75 years], USA); Schuit et al. (1997) reported a mean total PASE score of 85.5 (age 60 – 80 [mean age 70.6 years], Netherlands); (Washburn et al., 1999) reported a mean total PASE score of 131.3 (SD 70.4; age 55 – 75 [mean age 66.5 years], USA); Ismail et al. (2015) reported a mean total PASE score of 94.96 (SD 62.82; age 60+ [mean age 66.4] years, Malaysia), and Logan, Gottlieb, Maitl, Meegan, and Spriet (2013) reported mean total PASE Score of 155 (SD 66; age 60 years or older [mean 72 years], Canada). PASE has been reported to be higher for males and those with higher education (Chad et al., 2005; Washburn et al., 1999), but we found no differences according to gender or education level in the present study.

The results indicated that higher PASE scores tended to be associated with faster Brake RT, but were associated with poorer performance on the Traffic task (although the correlations were not statistically significant). However, the direction of the relationship for the Traffic Task was opposite to our prediction. One possible explanation for this is that older adults who are more physically active are more likely to utilise alternative forms of transport, such as cycling or walking, or conversely, may choose these alternative methods as a way of self-regulating their driving. At least one large study provides some support for this, with non-drivers reporting walking more often than drivers (Ross et al., 2011). However, in the present study, PASE was not significantly correlated with either of our driving exposure measures (self-reported distance driven and time spent driving).

Implications

In terms of current assessment procedures, the results suggest that visual acuity may not be useful for identifying at-risk drivers. If this is the case, visual acuity is therefore not a useful measure for inclusion in licensing assessments, as is current policy in most jurisdictions. The lack of association between the visual function measures and simulated driving performance measures adds further evidence to the literature which generally finds negligible or very low associations between these measures and driving (Anstey et al., 2005; Owsley & McGwin Jr, 2010; Wood, 2002). Regarding the cognitive measures, it is suggested that these measures may be useful as screening measures for older drivers, especially those who are most at-risk. To test this suggestion, assessment in terms of real-world driving would be required. IT, Crowding and UFOV™ could be administered quickly and efficiently via PC computer in an environment such as a doctor's office or licensing authority to identify drivers exhibiting severe declines in these areas which may put them at risk for dangerous driving; this would be consistent with the alternative targeted License Assessment model suggested by Fildes et al. (2008). Further research with a broader range of participants would be needed to establish appropriate cut-points on these tests, and to investigate the relationship between test performance and on-road driving.

The results were consistent with the Multifactorial Model of Driving Safety proposed by Anstey et al. (2005), which suggests that safe driving depends on visual function, physical function, and cognitive ability. Although neither of the vision measures were associated with simulated driving performance, an adequate level visual function is likely a threshold measure for driving safety, whereby the threshold is lower than the scores recorded by participants in the present study, and is likely to be lower than the current specified requirements for licensing. The results relating to physical function were somewhat consistent with the model; PASE contributed to Brake RT, but was not associated with the Traffic Task. The results relating to the cognitive measures were consistent with the model, with better performance on the cognitive measures being associated with better performance on the driving performance tasks, particularly Brake RT.

Although the correlations reported in the present study were small, our sample comprised of a unique, high-functioning and highly educated subset of the older driver population. The characteristics of the sample of participants meant that they performed very well on the screening tests and performed better than would be expected in the general population of older drivers. The fact that significant correlations were observed even in this unique sample suggests that these processes of visual attention and processing speed are important for the driving performance of all older drivers, not just those who appear impaired or at-risk. A broader sample including impaired drivers, or those suspected of being at-risk, would be likely to produce stronger results.

Conclusion

Overall, the results demonstrated negligible to small correlations between simulated driving performance and the measures of cognitive ability, visual function, and physical activity. The sample consisted of high-functioning, physically active older drivers. The amount of variance in the outcomes that could be explained by the functional measures was low. As argued by (Bédard et al., 2008), a lack of predictive value (despite statistically significant associations) is an issue for driver screening tests often considered in the literature. Although some significant associations between the predictors and simulated driving outcomes were observed, it must be noted that a large proportion of variance remained unexplained by the models. A combination of PASE, IT, and Crowding Total significantly explained 17.21% of the variance in Traffic Task Total Score, although in this model, lower Physical Activity was associated with better driving performance. A combination of UFOV™ Subtest 2, MMSE, and PASE significantly explained 14.2% of the variance in Brake RT. The results relating to Physical Activity require further investigation. It is suggested that future research in this area focuses on those drivers who are already showing signs on functional impairment or reduced driving safety.

Chapter 9.
Conclusions

Summary of findings

The aim of the thesis was to identify cognitive predictors of safe driving ability in older adults. Chapters 1 and 2 have provided an overview of the background, rationale and aims for the thesis, and reviewed the cognitive and functional abilities that have been associated with driving ability and the materials commonly used for assessment.

Study 1 focussed on the UFOV™, a well-known screening test for driving performance. Despite its frequent use and established relationships with driving outcomes, what the UFOV™ actually measures was not well understood. Relative importance regression analysis was used to investigate whether performance on the three UFOV™ Subtests could be explained in terms of combinations of six cognitive and visual predictor variables: general mental competence (assessed by MMSE), processing speed (assessed by IT), crowding across the visual field (assessed by ProPerVis Crowding), change detection (assessed by DriverScan), contrast sensitivity, and visual acuity. The results contributed to an improved understanding of the underlying cognitive factors that contribute to UFOV™ performance. It was found that UFOV™ Subtest 1 primarily reflected low-level visual function (visual acuity and contrast sensitivity). UFOV™ Subtest 1 often displays floor effects and lack of range in normal samples, and this result showed that it is dependent on low-level visual processing rather than higher-level cognitive processes and is therefore unlikely to be useful for non-impaired users. UFOV™ Subtest 2 reflected aspects of change detection, processing speed (as measured by IT), and general cognitive function (as measured by MMSE). UFOV™ Subtest 3 reflected crowding across the visual field (as measured by ProPerVis Crowding), processing speed (as measured by IT), contrast sensitivity, and general cognitive function (as measured by MMSE).

These results suggested that the measures of IT and Crowding, which had not previously been assessed in relation to driving, may be useful in predicting driving performance, based on their importance for UFOV™ subtests 2 and 3, which are known to be predictive of driving outcome including crash risk and performance in on-road driving assessments. The fact that IT and Crowding explained part of UFOV™ suggested that they might also explain variance in driving performance, and that driving performance is dependent on aspects of visual processing as measured broadly by UFOV™, and measured more specifically by IT and Crowding. Based on these results, IT and Crowding were examined further in the following studies, which investigated their contribution to driving performance as measured by two tasks on a driving simulator which assessed reaction time while driving, and general driving performance.

Studies 2 to 4 investigated methodological issues relating to the use of driving simulators for assessment of older drivers, including simulator sickness, validity, reliability, and usability. The results from Study 2, which investigated risk factors for simulator sickness, showed that older adults in general are a high-risk group, as are females and those with a history of motion sickness. This study provided a discussion of risk factors for simulator sickness, and other factors inherent in the custom-made simulator used in the study that probably contributed to the high simulator sickness rate. As a consequence of the high proportion of elderly drivers experiencing motion sickness during this study, the decision was taken to discontinue using the in-house custom-designed simulator developed for the study and to replace it with a commercially purchased simulator for future work.

Studies 3 and 4 used a variety of methods to show that the new driving simulator used in these studies, and in Study 5, demonstrated reliability, face validity, content validity, discriminant validity, and convergent validity, and was perceived as acceptable by participants. Feedback from older participants in Study 3 indicated that they were generally positive towards the simulator and perceived that the scenarios were an appropriate assessment of their driving ability. The study found support for the content and convergent validity of two of the driving tasks (a Distracted Driving task and a Traffic Participation Task). The Brake RT task showed good test-retest reliability, and in Study 4, the younger drivers recorded faster Brake RTs than the older drivers, providing evidence for the discriminant validity of the task. Study 4 also showed that there was no statistically significant difference between the age groups on overall performance on the Traffic Task, but there were some differences in the specific grade categories; the older drivers scored significantly higher in the category of Speed Control (i.e. speeding infringements), and the younger drivers scored significantly higher in the categories General, Lane Position, and Signalling. This provided support for the hypothesis that younger drivers would perform worse in the grade categories reflecting intentional violations (e.g. speeding), and that older drivers would perform worse in the grade categories reflecting errors in information processing and decision making. Feedback was generally positive from both age groups, but the younger drivers provided significantly more positive feedback and were also significantly less likely to experience simulator sickness. Overall, it was concluded that the simulator and tasks were suitable for use with older drivers. Studies 2 to 4 provided a discussion of methodological considerations around the use of driving simulators in older drivers, and informed the methodology for Study 5.

Study 5 investigated the contribution of a range of functional measures (cognitive, visual, and physical activity) in relation to performance on two simulated driving tasks: a Brake RT task and a Traffic Participation task. Of particular interest were IT and ProPerVis Crowding. These two tests have not previously been investigated in relation to simulated driving performance. The relationship between physical activity and driving performance was also of interest; evidence suggests that physical activity and exercise may be related to cognitive function and driving ability in older adults. Physical activity was measured using a self-report questionnaire (PASE). The results showed small, but statistically significant, correlations between the cognitive measures (IT, ProPerVis Crowding, and UFOV™ Subtest 2) and the performance on the simulated Brake RT task. However, none of the visual, physical, or visual functional predictor variables were significantly correlated with performance on the simulated Traffic Task. Overall, the results indicated that the cognitive measures, reflecting aspects of speed of visual processing, explained a small but significant amount of variance in the Brake RT measure in a sample of high-functioning healthy older drivers.

The following sections draw together the results of the five studies and provide a discussion of practical and theoretical implications, limitations, and future directions.

Implications of findings

Relationship between cognitive abilities and driving performance

Chapter 1 briefly reviewed the Multifactorial Model of Driving Safety (Anstey et al., 2005). The model separates the “capacity to drive safely” from actual “driving behaviour”. The capacity to drive safely depends on inter-correlated cognitive, visual, and physical function factors, and actual driving behaviour depends on both the capacity to drive safely, and beliefs about driving capacity, accuracy of self-monitoring, and insight into functional ability.

Consistent with Anstey et al.’s (2005) review and the other literature reviewed in Chapter 1, the results indicated that driving performance was most related to the cognitive measures, which showed small but significant correlations with simulated driving performance for older adults. The visual function and physical activity measures were not significantly associated with driving performance.

The results also confirmed that the younger drivers performed significantly better than the older drivers on the cognitive assessments. This occurred even though the older participants were high-functioning and highly educated compared to the general population of older drivers. For the younger drivers, there was no significant association between the cognitive abilities and driving performance, but for the older drivers, all three cognitive measures showed small but statistically significant correlations with the simulated driving tasks. The results from Study 4 and Study 5 showed that IT, UFOV™ Subtest 2, were related to performance on the Brake RT task, and Crowding showed a small but significant correlation with the Traffic Task. These results suggested that the three cognitive measures (IT, Crowding, and UFOV™ Subtest 2) were sensitive to age-related cognitive declines that are relevant for driving performance in older adults, and the large to very large effect sizes for the age group difference between the older and younger drivers suggested that these measures may be useful as screening measures to detect age-related declines relating to driving risk. The following sections describe the findings, conclusions, and implications relating to the cognitive assessments used throughout the thesis.

ProPerVis Crowding. Of the cognitive variables, ProPerVis Crowding was the only variable that was significantly correlated with performance on the Traffic Task in Study 5. It was also significant correlated with performance on the Brake RT task and showed the largest age-group difference, with older adults recording significantly more errors than younger adults at all positions across the visual field.

The effect of crowding can affect performance on everyday tasks, including reading, interacting with the environment, and driving, because the visual scene is often cluttered (Whitney & Levi, 2011). In a cluttered scene, crowding affects the identification of objects, making them appear indistinct or jumbled with surrounding objects (Levi, 2008). The results reported here, consistent with previous results, showed that the effects of crowding increased with age and with the eccentricity of the object, indicating that older adults are particularly vulnerable to crowding for objects outside of central vision.

Crowding may be particularly relevant for driving performance in older adults because crash patterns have shown that older drivers are more likely to be involved in crashes at intersections. Intersection crashes are likely to involve a complex scenario with multiple vehicles and important objects in different areas of the visual field. For example, analysis of Australian crash data from 1996 − 1999 showed that older drivers were more likely to be involved in intersection crashes involving multiple vehicles, and were more at-risk when the intersection was controlled by stop or give-way signs. Particularly prominent intersection crashes involved attempting a right turn in front of a vehicle approaching from the opposite direction; two vehicles entering an intersection from adjacent paths; and attempting a right turn in front of a vehicle approaching from an adjacent path (Langford & Koppel, 2006). This is consistent with other studies from Australia, Europe, and the United States reporting increased involvement in collisions involving multiple vehicles, failure to give right of way, turning across traffic, failing to heed stop signs, and failing to notice other objects, people, or vehicles (Baldock & McLean, 2006; Clarke et al., 2010; L. E. Hakamies-Blomqvist, 1993; McGwin Jr & Brown, 1999; Skyving, Berg, & Laflamme, 2009)

This pattern of crashes suggests that older drivers have difficulty with detecting and monitoring important objects across the visual field. The effects of crowding may cause vehicles in peripheral vision to appear indistinct or jumbled, leading to difficulty navigating the intersection successfully. This may explain why Crowding was the only cognitive measure that showed an association with performance on the Traffic Task; this task took place in a cluttered city environment and drivers were required to continuously monitor the position of other vehicles and objects in the scene. By comparison, IT and UFOV™ Subtest 2 reflected visual processing speed and attention but without the presence of visual distraction or clutter.

The development of the ProPerVis task has been described by Burns et al. (2005); Corlett and White (2002); and Werner (2008). However, ProPerVis has not featured in published literature, except for the first three studies of this thesis which have now been published. This thesis is also novel in that it has been the first to investigate the relationship between Crowding and driving performance.

ProPerVis was developed as a potential alternative to the UFOV™ for screening fitness-to-drive in older adults. Previous results have shown that the effects of crowding are stronger at greater eccentricities, and that older adults are more susceptible to crowding than younger adults (Corlett & White, 2002; Werner, 2008). These results were replicated throughout the current series of studies. In particular, Study 4 compared the performance of younger and older adults and the results showed significant main effects for age and stimulus position, as well as a significant age $\times$ position interaction whereby older adults were more susceptible to the effects of crowding at the parafoveal and peripheral stimulus positions.

The effect size for the mean difference in total crowding errors between older and younger adults was very large (Cohen's $D = 2.44$) indicating very little overlap in the performance distribution of the older and younger age groups. Inspection of the results showed that the almost half of the younger adults scored zero or one total errors, whereas all of the older adults scored at least two errors.

Performance on the crowding task was consistent across studies, and the results for the Crowding task in the studies included here were similar to results reported previously (Burns et al., 2005; Werner, 2008). In comparison, performance on UFOV™ Subtest 2 (and to a lesser extent, IT) was inconsistent between studies, although it should be noted that the age range and mean age of participants was slightly different between studies. A summary of reported performance for older adults on IT, UFOV™, and ProPerVis crowding is shown in Table 9.1.

It has been suggested a combination of IT and Crowding may provide a useful

Table 9.1: Summary of results for UFOV™, IT, and ProPerVis Crowding

	Burns et al. (2005)		Werner (2008)		Study 1		Study 3		Study 4/5	
Variable	M	SD	M	SD	M	SD	M	SD	M	SD
IT[a]	67.0	18.2	56.0	15.0	66.3	26.0	71.6	29.2	61.5	18.0
UFOV™[a]										
Subtest 1	18.0	4.3	28.4	35.9	28.2	30.1	-		-	
Subtest 2	35.8	24.3	99.6	122.0	115.5	106.7	84.6	79.7	65.0	55.3
Subtest 3	176.8	91.4	209.1	125.8	223.6	103.4	-		-	
Crowding[b]										
Central	.92	.08	.92	.13	.92	.16	.93	.05	.93	.11
Parafoveal	.72	.17	.79	.16	.78	.19	.80	.15	.79	.14
Peripheral	.42	.18	.46	.18	.46	.21	.54	.22	.49	.17
Participants	N = 60, mean age 67.3 years (SD 6.5) for females, mean age 62.4 years (SD 3.9) for males		N = 30, mean age 62 years (SD 8.39, range 49 80)		N = 82, mean age 75.13 years (SD 7.79, range 62 – 92)		N = 111, mean age 73.87 years (SD 6.10, range 65 – 95)		N = 82, mean age 72.93 years (SD 4.39, range 65 – 84)	

[a] IT and UFOV™ measure are in ms. [b] Crowding measures are proportion of correct trials.

alternative to the UFOV™ for driver screening (Burns et al., 2005; Werner, 2008). IT and Crowding reflect similar cognitive processes (visual processing and speed and attention) to the UFOV™. IT provides a sensitive measure of processing speed in central vision, and the results reported here showed that IT and Crowding were significantly correlated with UFOV™ Subtests 2 and 3. Of the three UFOV™ Subtests, Crowding was most strongly correlated with UFOV™ Subtest 3, consistent with the results reported by Burns et al. (2005). The results from Study 1 showed that both IT and Crowding contributed to performance on UFOV™ Subtest 3. Similar to ProPerVis Crowding, UFOV™ Subtest 3 features flanked targets in peripheral vision. A combination of IT and Crowding is argued to have several advantages over the UFOV™.

Firstly, UFOV™ Subtest 1 (processing speed) is too easy for most older drivers and suffers from floor effects. It therefore is not sensitive enough to differences in processing speed, and it is argued that the use of the random noise mask in the UFOV™ contributes to this issue because an image of the target stimulus may persist after that mask has appeared (Burns et al., 2005). In contrast, IT did not display floor effects or range restriction in the studies reported here (see Figure 7.3). Furthermore, IT uses a "flash" backward pattern mask rather than a random noise mask, a procedure which appropriately masks the target stimuli by reducing apparent movement (Evans & Nettelbeck, 1993).

Secondly, Crowding and IT appear to be more reliable than the UFOV™. As seen in Table 9.1, performance on the Crowding measures was consistent across studies, while performance on the UFOV™ Subtests varied greatly by comparison. Additionally, Burns et al. (2005) reported that UFOV™ performance was susceptible to significant practice improvements over one week, while Crowding and UFOV™ showed weaker and non-significant practice effects.

Finally, Crowding has an advantage over UFOV™ Subtests 2 and 3 in the way that the targets are presented. Targets in the Crowding task are presented in five distinct positions (central, left and right parafoveal, and left and right peripheral). While previous versions of the UFOV™ presented targets at varying eccentricities, in the current version targets are presented at one of nine radial positions of the same eccentricity. This means that the Crowding task provides additional information about impairments at different areas in the visual field.

These results suggest that a combination of IT and Crowding may be useful for driver screening. IT and Crowding can both be administered quickly and easily via PC and would be easy to deliver in a setting such as a doctor's office or DMV office. It is suggested that further studies continue to investigate the relationship between IT, Crowding, and driving outcomes. Particularly, future research should investigate on-road driving performance and future crash risk, with a focus on those older drivers deemed to be at-risk. Additionally, if these tests were to be used for driver screening, appropriate cut-points for scores indicating a significant level of risk would need to be determined.

Inspection Time. As already suggested above, it is argued that a combination of IT and Crowding could be used as an alternative to the UFOV™. As a test of visual processing in central vision, IT appears to be more sensitive and reliable than UFOV™ Subtest 1. The results from Study 1 also showed that IT explained a significant amount of variance in the more complex UFOV™ Subtests 2 and 3, which assess visual attention across the visual field, but which also depend on speed of visual processing. IT showed an approximately normal distribution of scores for both older and younger adults.

IT was significantly correlated with performance on the simulated Brake RT task for older drivers, and of the functional predictor variables, IT showed the highest correlation with Brake RT. This association suggests that drivers with faster IT were able to process the Brake RT stimuli more efficiently than those with slower IT.

In addition to investigating further the relationship between IT, Crowding, and driving outcomes in older drivers, future research could also investigate IT as a potential biomarker for driving difficulties. A biomarker is an indicator that is able to predict future changes in cognitive and functional outcomes better than chronological age; and, short term changes in a biomarker should be able to predict future declines (Baker & Sprott, 1988). IT has been proposed as a good candidate as a biomarker for future cognitive performance, because it reflects speed of processing (which is implicated in age-related cognitive decline) and is also related to general cognitive functioning (Nettelbeck & Wilson, 2004). IT has been reported to be stable across generations and sensitive to age-related change (Nettelbeck & Wilson, 2004), and, in a study comparing IT to other measures of speed of processing as potential biomarkers, IT was least dependent on childhood cognitive abilities (Deary et al., 2010). Gregory et al. (2008) investigated the utility of IT as a biomarker for cognitive decline and reported that, in a sample of healthy older adults aged over 80, current IT and short-term changes in IT predicted future performance in tests of fluid reasoning, perceptual speed, and working memory. In a follow-up study, Gregory et al. (2009) reported that IT predicted performance four years later on Timed IADL tasks, such that older adults with poorer baseline IT made more errors and took more time to complete everyday activities. If it is established that slower IT or short-term changes in IT are predictive of future driving difficulties in older adults, then these drivers could potentially engage in interventions to reduce their crash risk and improve their future mobility.

Useful Field of View Test™. As reviewed in Chapter 1, UFOV™ is frequently used in driving research. The studies involved in the program of research included here all included the UFOV™. Study 1 showed that UFOV™ Subtest 1 primarily reflected low-level visual function; UFOV™ Subtest 2 reflected change detection, processing speed (as assessed by IT) and general cognitive function; and UFOV™ Subtest 3 reflected crowding (as assessed by ProPerVis), processing speed (as assessed by IT), contrast sensitivity, and general cognitive function. These results contributed to an understanding of the psychometric properties of the UFOV™, which have not previously been well defined in the literature. Studies 4 and 5 showed significant age-group differences in performance on UFOV™ Subtest 2, and that for older drivers UFOV™ Subtest 2 was significantly correlated with performance on the simulated Brake RT task.

It became clear from Study 1 that UFOV™ Subtest 1 was susceptible to floor effects. Based on the results of Study 1, it was concluded that UFOV™ Subtest 1 was too easy for most healthy older adults and was insufficiently sensitive to differences in processing speed. For the remaining studies, UFOV™ Subtest 2 was selected for inclusion rather than the full UFOV™, or the more difficult UFOV™ Subtest 3, because it has previously been reported to have the strongest relationship with driving outcomes. However, Subtest 2 also displayed some evidence of floor effects, with the majority of participants scoring close to 17.6ms, the best possible score for the test. If the more complex UFOV™ Subtest 3 had been included, it may have shown more variance and a stronger relationship with the simulated driving performance measures. Nonetheless, the significant correlation between UFOV™ Subtest 2 and Brake RT showed that the aspects of attention and processing speed picked up by this subtest were relevant for driving reaction speeds, even amongst a sample of participants with restricted range on the measure.

UFOV™ Subtest 2 was not significantly associated with performance on the Traffic Task. In addition to the range restriction already noted, this may have been because Subtest 2 lacks the aspects of selective attention and crowding that may have been picked up more effectively by Subtest 3, and that may have been more relevant to driving performance in the complex city driving task.

As reviewed briefly in Chapter 2, previous research has shown that UFOV™ performance can be improved with training, that such improvements have been maintained for up to five years, and that Speed of Processing Training (based on the UFOV™) has been associated with reduced crash risk, improvements in certain measures of simulated and on-road driving performance, better self-reported driving mobility, and improvements in Timed IADL tasks.

Future studies could investigate potential interventions and training programs to target UFOV™, IT and Crowding in older drivers. These training programs could be administered via PC in the driver's home, and, if effective, could provide a simple and accessible strategy for improving cognitive performance and driving outcomes for older drivers. Training programs for IT and Crowding could be developed based on the successful Speed of Processing Training programs. Such training programs would be particularly useful for at-risk drivers whose driving ability is affected by age-related cognitive decline. Drivers could be required or encouraged to undertake such programs in order to improve their confidence, mobility, and safety of themselves and other road users. Training may also be appealing to high-functioning older drivers wishing to maintain or improve their cognitive abilities and driving performance. The results of the studies reported here showed that, even in a sample of volunteer participants who, as a whole, were cognitively high-functioning and highly educated, better processing speed and fewer crowding errors were associated with improved Brake RT in Study 5.

MMSE. The MMSE was included as a screening measure of general cognitive functioning. Surprisingly, despite the fairly severe restriction of range in the measure (all participants scored at least 26 out of 30), the MMSE was shown to correlate significantly with performance on UFOV™ Subtests 2 and 3, and to contribute to a model predicting performance on the Brake RT task. These results invite further consideration.

As reviewed in Chapter 2, the MMSE has been reported to be associated with driving outcomes in some studies with cognitively impaired drivers (Adler et al., 2005) but generally shows no association with driving outcomes in healthy drivers (Crizzle et al., 2012; Joseph et al., 2014). Mathias and Lucas (2009) reported that where an association between mental status and driving does exist, the effect size is smaller than the effect size for other cognitive functions (e.g. attention, perception, and reasoning). Anstey et al. (2005) concluded that mental status tests have shown inconsistent relationships with driving outcomes, and are vulnerable to strong ceiling effects in normal samples. Crizzle et al. (2012) concluded that the MMSE is not sensitive enough for non-impaired older drivers.

The results from the present series of studies support Anstey et al.'s (2005) and Crizzle et al.'s (2012) conclusions. The MMSE showed strong ceiling effects and lacked sensitivity among healthy older drivers. Although the requirement for participation was a score of at least 24 out of 30 on the MMSE, all participants in fact scored 26 or above. The measure showed strong ceiling effects, with the majority of participants recording scores of 29 or 30 out of 30. Accordingly, the MMSE lacked sensitivity as a measure of general cognitive functioning.

However, even with the ceiling effects and restriction of range, performance on the MMSE was associated with performance on UFOV™ Subtests 2 and 3 in Study 1, and contributed to a model predicting performance on the Brake RT task in Study 5. Moreover, in Study 1, MMSE contributed to performance on UFOV™ Subtests 2 and 3 even when the other cognitive variables were included in the regression models. This suggested that global cognitive functioning, as measured by MMSE, had some importance for UFOV™ performance, in addition to the variance explained by processing speed and visual attention; particularly, questions in the MMSE reflecting aspects of short term memory and following basic instructions may have explained this contribution. The contribution of the MMSE in Study 5 to a model predicting performance on the Brake RT task was found to be the result of two participants who recorded low scores on the MMSE, relative to the rest of the sample, who also recorded slower Brake RTs. It is possible that in this analysis the two participants were representative of a real effect of MMSE on Brake RT. The results from Study 1 and Study 5 have therefore suggested that, even small reductions in general cognitive functioning within the "normal" range may have an effect on driving performance and cognitive test performance.

It is possible that this effect is representative of a real decline in general cognitive functioning on driving and cognitive performance outcomes even among non-impaired older drivers scoring within the upper normal range on the MMSE, but undetectable with this measure. To test this suggestion, it may be beneficial to investigate the effect further with a more sensitive measure of general cognitive functioning. For example, an assessment such as Timed IADL tasks (Gregory et al., 2009; Owsley, McGwin Jr, Sloane, Stalvey, & Wells, 2001) or the Everyday Problems Test (Willis, 1996) may be more appropriate for testing the day-to-day cognitive functioning of healthy older adults. Timed IADL tasks and the Everyday Problems Test both assess accuracy of performance on everyday tasks, such as finding a telephone number, finding and reading medicine instructions, and finding and counting correct change. Timed IADL tasks also measure speed of task performance, and have been reported to be related to IT (Gregory et al., 2009) and the UFOV™ (Owsley, Sloane, McGwin Jr, & Ball, 2002) in non-impaired older adults. Because Timed IADL tasks are scored based on both speed and accuracy, they are sensitive to small differences in efficiency of task performance and yield a wider range of scores. As they are more discriminating than the MMSE, the Everyday Problems Test and Timed IADL tasks may be useful for identifying minimal cognitive declines that can predict driving performance in healthy older adults.

Assessment and screening for older drivers

Current and proposed licensing and assessment procedures were described in Chapter 1. Most Australian jurisdictions have abolished mandatory age-based on-road testing, but most states require some form of compulsory age-based medical assessment (see Table 1.1). Fitness-to-drive is determined by a GP or other qualified health professional based on standards specified by Austroads (2012). Drivers must meet the specified visual acuity and visual fields standards to hold a license. Medical professionals and drivers themselves are expected to report medical conditions that can affect driving ability to the licensing authority, and diagnosis-specific guidelines are provided for determining fitness-to-drive. Currently, there are no specific standards relating to cognitive abilities, although declines in cognitive abilities are considered in the context of other conditions. A driver may also be required to take an on-road driving test on the recommendation of GP or other qualified health professional.

Two proposed driver assessment models were described in Chapter 1: the Australasian Model (Figure 1.1; Fildes et al., 2000) and the North American NHTSA Model (Figure 1.2; Staplin et al., 2003). Both models are consistent with the OECD recommendation that testing for older drivers will be most effective and equitable if only those drivers who pose a significant risk to others are targeted (OECD, 2001). Both models feature a graded series of assessments, with drivers being referred into the system from various community sources including medical and health practitioners, family members, police, and the licensing authority. At an early stage in the process, both models feature validated off-road functional screening tests, in areas of mental function, visual function, and physical function, to assist in determining fitness-to-drive. These functional screening tests must be able to validly predict driving outcomes.

In terms of current assessment procedures, the results suggest that visual acuity is not generally useful for identifying at-risk drivers, and therefore this varibale need not be included in licensing assessments, as is currently the case in most jurisdictions. The lack of association between the visual function measures and driving performance measures adds further evidence to the literature which generally finds negligible or very low associations between these measures and driving (Anstey et al., 2005; Owsley & McGwin Jr, 2010; Wood, 2002). Visual acuity likely represents a threshold measure for the ability to drive safety, whereby the threshold is well below the typical requirement of logMAR 0.5. To add support to this, a number of participants in the present series of studies recorded visual acuity estimates below logMAR 0.5 but were otherwise fit, healthy, and able to complete all study tasks with no obvious signs of difficulty.

Regarding the cognitive measures included in the present series of studies, it is suggested that IT and ProPerVis Crowding may be useful as screening measures for older drivers, especially those who are most at-risk. The cognitive measures showed stronger associations with driving outcomes than the visual measures and are more relevant to the complex visual and cognitive skills needed to drive safely. IT and ProPerVis Crowding could be administered quickly and efficiently via PC in an environment such as a doctor's office or licensing authority to identify drivers exhibiting severe declines in these areas which would put them at risk for dangerous driving. These tests could be considered for inclusion in driver screening models such as those described by Fildes et al. (2000) and

Staplin et al. (2003). For this to be considered, it would need to be determined whether IT and ProPerVis Crowding can validly predict driving outcomes such as crash involvement.

Assessment of driving performance with low-cost driving simulators

As described in Chapter 3, several challenges were experienced with the use of a low-cost driving simulator for older adults. The main concern was the rate of simulator sickness. The rate of simulator sickness was subsequently substantially lower with the replacement commercially purchased simulator compared to the in-house custom-designed simulator, and rates of simulator sickness were at the lower end of rates reported in the literature for similar studies. Nonethtless, simulator sickness still affected 33% of the older participants in Study 5; and, ideally, an attempt should be made to improve this outcome. For research studies, simulator sickness may be reduced by making modifications to the scenario, scene, and simulator setup can help to reduce SS (Cassavaugh et al., 2011), and use of screening procedures (Trick & Caird, 2011) and adaptation procedures (Domeyer et al., 2013). However, making these changes may limit the usefulness of the assessment. For example, reducing the field of view may reduce simulator sickness, but is it also likely to limit the ability to present objects and vehicles in peripheral vision; detecting and monitoring objects in peripheral vision is particularly relevant for older drivers. Limiting turns around corners may reduce simulator sickness, but cornering is an essential part of driving, and is highly relevant for older drivers given the reported risk of intersection crashes. No immediate solution to this problem can be offered here but it is likely that continued investigation of the balance between the manipulation of these essential driving variables and experience of SS could improve method.

Another concern was the validation of the driving simulator. The gold-standard for validation of driving simulators is to validate simulated driving performance against on-road driving performance on the same measures; that is, behaviour in the simulator should accurately reflect behaviour in a real-world driving situation. As this was considered to be outside the scope of the present series of studies, other methods were used to investigate the validity and reliability of the driving simulator. Feedback from participants indicated that they were generally positive towards the simulator and perceived that the scenarios were an appropriate assessment of their driving ability, and it was shown that the simulator and aspects of the selected tasks demonstrated evidence of content validity, discriminant validity, face validity, and test-retest reliability. However, feedback and comments from some of the older participants suggested that they did have some concerns about the realism and relevance of the driving simulator assessment. This was evident when comparing the feedback questionnaire responses from the older and younger participants in Study 4, where the younger drivers provided significantly more favourable responses to all questions. The validity of the SimWorx simulator for use with older adults would benefit from further investigation, preferably by validating performance against on-road driving.

The results reported here showed a significant association between cognitive test performance and simulated driving performance. This result is only meaningful if simulated driving performance is established as reflective of actual real-world driving ability. Future studies should seek to replicate the relationship between IT and ProPerVis Crowding and driving ability using alternative outcome measures, for example an on-road driving assessment, or state-recorded retrospective or prospective at-fault crash involvement. It would also be useful to investigate the relationship with a higher-fidelity driving simulator. Such simulators, although significantly more expensive - and beyond the scope of the studies reported here - provide a more realistic experience through features such as a motion platform, real car controls, a full car cabin, and 360° displays.

Limitations and future directions

Sample characteristics for studies with older drivers

All studies included in the thesis used volunteer participants, who in general were healthy, cognitively high-functioning, highly educated, and physically active. The convenience sample of participants in each study were recruited from various community sources and through snowballing. The aims of the studies were clearly identified and therefore attracted participants with an interest in driving ability, healthy ageing, and driving simulation. The studies using a driving simulator were also restricted to recruitment of participants in excellent health who had no history of motion sickness or simulator sickness. Clinical participants, for example drivers with cognitive impairment or dementia, were outside of the scope of the thesis. The characteristics of the sample of participants meant that they performed very well on the screening tests and performed better than would be expected in the general population of older adults. In short, they would not be considered to be typical "older drivers" - or at least should be regarded as representing older drivers with high levels of competence.

These participants were appropriate for the goal of determining the contribution of underlying cognitive processes to driving performance. This goal was only limited by the floor/ceiling effects of some of the functional predictor variables and the driving performance outcomes, and the validity of the driving simulator. The fact that significant correlations were observed even in this unique sample suggests that the processes of visual attention and processing speed are important for the driving performance of all older drivers, not just those who appear impaired or at-risk.

A broader sample including impaired drivers, or those suspected of being at-risk, would be likely to produce stronger results. As suggested by the OECD, and the NHTSA and Australasian screening models described in Chapter 1, driver screening procedures will be most useful when targeted to those older drivers who are already showing signs of functional impairment or high-risk driving behaviour. The results from a sample including at-risk drivers would be essential for determining appropriate cut-points on the various screening tools.

Physical activity measurement

Study 5 used a self-report physical activity questionnaire, the PASE, to investigate the relationship between physical activity and driving. A measure of physical activity was included based on evidence suggesting a link between physical fitness, cognitive abilities, and driving performance. The Multifactorial Model of Driving Safety also includes physical function as a determinant of Safe Driving Ability (Anstey et al., 2005). However, the results showed that PASE scores were not significantly correlated with any of the cognitive ability measures or driving performance measures. PASE did contribute to a regression model predicting Brake RT, suggesting that higher self-reported physical activity tended to be associated with faster Brake RT, however this predictor did not reach statistical significance.

The relationships between PASE scores and other study variables raised questions regarding the validity of the measure. PASE scores were not significantly associated with any of the cognitive measures, as was expected based on previous research suggesting a correlation between physical activity and cognitive ability in older adults. PASE scores also showed no association with age, gender, or education level; previous studies have reported that higher PASE scores are associated with male gender and higher education, and that PASE scored tend to decrease with age (Chad et al., 2005; Washburn et al., 1999).

As mentioned in the discussion of Study 5, PASE scores were higher than has generally been reported previously. The high scores on the PASE may have reflected a volunteer bias, whereby only those older drivers who considered themselves to be healthy, active, and physically fit volunteered to take part in the study. The participants in all studies were well-educated and cognitively healthy, and it is likely that they were more physically fit and active than the general population of older drivers.

The high scores for some participants may also have been caused by response errors by participants when filling out the questionnaire; for example, inspection of individual responses indicated some participants' responses implied that they spent eight or more hours each day engaged in light, moderate, or vigorous physical activities, a figure which appears unlikely for even the most active participants. It is possible that these responses were caused by misinterpretation of the questionnaire items whereby participants reported the total duration of activity per week, rather than the average duration per day, leading to overestimation of physical activity and a higher than expected PASE score.

It is also worth noting that self-report physical activity questionnaires in general tend to be associated with overestimation of actual activity (Sallis & Saelens, 2000; Shephard, 2003). Direct, objective measures of physical activity, such as from a pedometer or accelerometer, tend to provide more accurate and reliable results (Colbert, Matthews, Havighurst, Kim, & Schoeller, 2011; Falck, McDonald, Beets, Brazendale, & Liu-Ambrose, 2016). However, these direct measures have a separate set of complications, including compliance with use, familiarity with the technology, accuracy of measurement for different types of activity, and validity for older adults who tend to walk slower and may have problems with gait compared to other age groups (Kowalski, Rhodes, Naylor, Tuokko, & MacDonald, 2012).

If using the PASE, future studies should aim to ensure accuracy of responding. Problems of this kind were not foreseen in setting up present study, and participants completed the questionnaire in their own time via computer. Changes could be made to the presentation or wording of the questions to emphasise the correct time periods. The questionnaire could also be programmed to disallow or highlight implausible responses (e.g. reporting 6+ hours of daily strenuous activity) and require that the participant confirm their responses or provide more information. Another option could be to administer the questionnaire over the phone or in person.

The relationship between physical activity, physical fitness, and driving performance in older adults has been investigated in only a few studies. Several questions remain to be answered about the relationship between physical activity and driving. For example, which type of physical activity is most beneficial for driving? What duration and intensity is required? And if there is a association between physical activity and driving, what is the mechanism of the relationship? Future research should address these questions with valid and reliable measures of physical activity. Intervention studies could be developed to determine whether certain types of physical activity are beneficial for driving, and longer-term studies could investigate whether physical activity over the lifetime is associated with better driving outcomes in old age.

Accurate screening tools, assessment and licensure for older drivers

Both the Australasian Model License Re-assessment Procedure (Fildes et al., 2000) and the NHTSA Driver Screening and Evaluation Model (Staplin et al., 2003) featured functional screening tests as key components for making fitness to drive decisions. To be useful to the models, these tests must be able to reliably and validly predict driving outcomes.

Langford (2008) described some of the challenges involved with the development and application of screening tests. According to Langford (2008, pp. 332) the aim of a screening test is to "to identify those older drivers whose levels of functional impairment have deteriorated to such an extent as to be associated with an "unacceptably high" probability of causing a crash". Langford considered two screening tests which have been associated with the prediction of at-fault crashes, the UFOV™ and the MaryPODS test battery. On the basis that failing these tests has been associated with a 2× increase in at-fault crash risk, Langford calculated that given reported rates of performance on the screening tests, and rates of crash-involvement in older drivers, applying the screening tests to the whole population of older drivers would result in large numbers of drivers failing the test, even though 99% of them would have continued to drive without crashing in the subsequent year. It was argued that the validity of screening tests would be improved if they produced three-level outcomes (pass, fail, or uncertain) and were targeted towards a sub-population of older drivers already determined to be at heightened risk for crashing. In this case, far fewer drivers would be deemed to fail the test when they were otherwise safe to continue driving.

A necessary first step for a screening test is that is shows a significant association with real-world driving outcomes. ProPerVis Crowding and IT demonstrated significant (but small) associations with simulated driving performance. However, the sample of participants was high-functioning and probably not representative of the general population of older drivers, and especially not representative of at-risk older drivers. In the present series of studies, predictive validity of the measures was not considered, and cut-points were not calculated. If IT and ProPerVis Crowding were to be considered for inclusion in driver screening procedures, future research would need to establish their predictive validity for important driving outcomes (ideally at-fault crash involvement), and establish whether appropriate cut-points could be determined for these tests to categorise drivers as safe, unsafe, or uncertain. This would need to be done with a sample which includes at-risk drivers, for example drivers who are showing signs of functional declines, or drivers who have recently been involved in crashes or near-misses.

Concluding remarks

This thesis has investigated cognitive and functional predictors of driving performance in older adults. Current licensing procedures for older drivers typically rely on medical and visual assessments, but the validity of existing procedures is questionable. Thus, as the proportion of older drivers continue to grow, there is a need to develop accurate screening tools and fair licensing procedures for older drivers, particularly those showing signs of functional declines or risky driving behaviours.

The results have suggested that cognitive abilities measures are more strongly associated with driving performance measures than are low-level visual function measures. Even among high-functioning active drivers, measures of visual attention, visual processing speed, and crowding across the visual field were significantly correlated with driving reaction time and overall performance on a traffic participation task; however, it should be noted that there was large proportion of unexplained variance in the models, and further research would be required to investigate the predictive validity of the measures. It has been suggested that IT and ProPerVis Crowding have potential for use as driver screening and assessment tools, particularly when considered as an alternative to the UFOV™. For this suggestion to be tested, further research would be required to investigate the relationship between test performance and on-road driving. In the framework of a targeted driver licensing model for at-risk older drivers, IT and ProPerVis Crowding could be administered quickly and efficiently in a setting such as a doctor's office or DMV office to identify drivers exhibiting severe declines in these areas which would put them at risk for dangerous driving.

References

Adler, G., Rottunda, S., & Dysken, M. (2005). The older driver with dementia: An updated literature review. *Journal of Safety Research*, *36*(4), 399-407. doi: 10.1016/j.jsr.2005.07.005

Akaike, H. (1974). New look at the statistical model identification. *IEEE Transactions on Automatic Control*, *AC-19*(6), 716-723. doi: 10.1109/TAC.1974.1100705

Allen, R. W., Park, G. D., Cook, M., Rosenthal, T. J., Fiorentino, D., & Viirre, E. (2003). Novice driver training results and experience with a PC-based simulator. In *Proceedings of the 2nd international driving symposium on human factors in driver assessment, training and vehicle design.*

Allen, R. W., Park, G. D., Cook, M. L., & Fiorentino, D. (2012). Simulator training of novice drivers: A longitudinal study. *Advances in Transportation Studies*, *27*, 51-68. doi: 10.4399/97888548486724

Allen, R. W., Rosenthal, T. J., & Cook, M. L. (2011). A short history of driving simulation. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (chap. 2). Boca Raton, FL: CRC Press. doi: 10.1201/b10836-27

Alosco, M. L., Spitznagel, M. B., Cleveland, M. J., & Gunstad, J. (2013). Cognitive deficits are associated with poorer simulated driving in older adults with heart failure. *BMC Geriatrics*, *13*(1), 58. doi: 10.1186/1471-2318-13-58

Anstey, K. J., Horswill, M. S., Wood, J. M., & Hatherly, C. (2012). The role of cognitive and visual abilities as predictors in the multifactorial model of driving safety. *Accident Analysis and Prevention*, *45*, 766-774. doi: 10.1016/j.aap.2011.10.006

Anstey, K. J., & Wood, J. (2011). Chronological age and age-related cognitive deficits are associated with an increase in multiple types of driving errors in late life. *Neuropsychology*, *25*(5), 613-621. doi: 10.1037/a0023835

Anstey, K. J., Wood, J., Caldwell, H., Kerr, G., & Lord, S. R. (2009). Comparison of self-reported crashes, state crash records and an on-road driving assessment in a population-based sample of drivers aged $69 - 95$ years. *Traffic Injury Prevention*, *10*(1), 84-90. doi: 10.1080/15389580802486399

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review*, *25*(1), 45-65. doi: 10.1016/j.cpr.2004.07.008

Asimakopulos, J., Boychuck, Z., Sondergaard, D., Poulin, V., Ménard, I., & Korner-Bitensky, N. (2012). Assessing executive function in relation to fitness to drive: A review of tools and their ability to predict safe driving. *Australian Occupational Therapy Journal*, *59*(6), 402-427. doi: 10.1111/j.1440-1630.2011.00963.x

Australian Bureau of Statistics. (2008). *Population projections Australia 2006 to 2101.* Retrieved from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3222.02006%20to%202101

Austroads. (2012). *Assessing fitness to drive for commercial and private vehicle drivers* (No. AP-G56-13). Austroads/National Transport Commission Australia. Retrieved from https://www.onlinepublications.austroads.com.au/items/AP-G56-13

Austroads. (2016). *Assessing fitness to drive for commercial and private vehicle drivers* (No. AP-G56-16). Austroads/National Transport Commission Australia. Retrieved from https://www.onlinepublications.austroads.com.au/items/AP-G56-16

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, *8*(2), 129-148. doi: 10.1037/1082-989X.8.2.129

Bach, M. (1996). The Freiburg Visual Acuity Test - Automatic measurement of visual acuity. *Optometry and Vision Science*, *73*(1), 49-53. doi: 10.1097/00006324-199601000-00008

Bach, M. (2007). The Freiburg Visual Acuity Test - Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *245*(7), 965-971. doi: 10.1007/s00417-006-0474-4

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation - advances in research and theory* (Vol. 8, p. 47-89).

Baker, G. T., & Sprott, R. L. (1988). Biomarkers of aging. *Experimental Gerontology*, *23*(4-5), 223-239. doi: 10.1016/0531-5565(88)90025-3

Baldock, M. R. J., & McLean, A. J. (2006). *Older drivers: Crash involvement rates and causes* (Report No. CASR015). Adelaide, Australia: Centre for Automotive Safety

Research. Retrieved from `http://casr.adelaide.edu.au/staff/?pid=516`

Ball, K. K. (1997). Attentional problems and older drivers. *Alzheimer Disease and Associated Disorders*, *11*(SUPPL. 1), 42-47. doi: 10.1097/00002093-199706001-00010

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America. A, Optics and image science*, *5*(12), 2210-2219. doi: 10.1364/JOSAA.5.002210

Ball, K. K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., ... Willis, S. L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association*, *288*(18), 2271-2281. doi: 10.1001/jama.288.18.2271

Ball, K. K., Edwards, J. D., Ross, L. A., & McGwin Jr, G. (2010). Cognitive training decreases motor vehicle collision involvement of older drivers. *Journal of the American Geriatrics Society*, *58*(11), 2107-2113. doi: 10.1111/j.1532-5415.2010.03138.x

Ball, K. K., & Owsley, C. (1993). The useful field of view test: A new technique for evaluating age-related declines in visual function. *Journal of the American Optometric Association*, *64*(1), 71-79.

Ball, K. K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology and Visual Science*, *34*(11), 3110-3123.

Ball, K. K., & Rebok, G. (1994). Evaluating the driving ability of older adults. *Journal of Applied Gerontology*, *13*(1), 20-38. doi: 10.1177/073346489401300103

Ball, K. K., Roenker, D. L., & Bruni, J. R. (1990). Developmental changes in attention and visual search throughout adulthood. In J. T. Enns (Ed.), *Advances in psychology* (Vol. 69, p. 489-508). North-Holland. doi: 10.1016/S0166-4115(08)60472-0

Ball, K. K., Roenker, D. L., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin Jr, G., ... Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? *Journal of the American Geriatrics Society*, *54*(1), 77-84. doi: 10.1111/j.1532-5415.2005.00568.x

Ball, K. K., Ross, L. A., Roth, D. L., & Edwards, J. D. (2013). Speed of processing training in the ACTIVE study: how much is needed and who benefits? *Journal of aging and health*, *25*(8 Suppl), 65S-84S.

Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *American Journal of Occupational Therapy*, *64*(2), 336-340. doi: 10.5014/ajot.64.2.336

Bédard, M., Weaver, B., Darzinš, P., & Porter, M. M. (2008). Predicting driving performance in older adults: We are not there yet! *Traffic Injury Prevention*, *9*(4), 336-341. doi: 10.1080/15389580802117184

Bélanger, A., Gagnon, S., & Yamin, S. (2010). Capturing the serial nature of older drivers' responses towards challenging events: A simulator study. *Accident Analysis and Prevention*, *42*(3), 809-817. doi: 10.1016/j.aap.2009.07.010

Bentley, S. A., LeBlanc, R. P., Nicolela, M. T., & Chauhan, B. C. (2012). Validity, reliability, and repeatability of the useful field of view test in persons with normal vision and patients with glaucoma. *Investigative Ophthalmology and Visual Science*, *53*(11), 6763-6769. doi: 10.1167/iovs.12-9718

Bherer, L., Erickson, K. I., & Liu-Ambrose, T. (2013). A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *Journal of Aging Research*, *2013*. doi: 10.1155/2013/657508

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177-178. doi: 10.1038/226177a0

Bowers, A. R., Anastasio, R. J., Sheldon, S. S., O'Connor, M. G., Hollis, A. M., Howe, P. D., & Horowitz, T. S. (2013). Can we improve clinical prediction of at-risk older drivers? *Accident Analysis and Prevention*, *59*, 537-547. doi: 10.1016/j.aap.2013.06.037

Bron, A. M., Viswanathan, A. C., Thelen, U., de Natale, R., Ferreras, A., Gundgaard, J., ... Buchholz, P. (2010). International vision requirements for driver licensing and disability pensions: Using a milestone approach in characterization of progressive eye disease. *Clinical Opthamology*, *4*, 1361-1369. doi: 10.2147/OPTH.S15359

Brooks, J. O., Goodenough, R. R., Crisler, M. C., Klein, N. D., Alley, R. L., Koon, B. L., ... Wills, R. F. (2010). Simulator sickness during driving simulation studies. *Accident Analysis and Prevention*, *42*(3), 788-796. doi: 10.1016/j.aap.2009.04.013

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, *114*(3), 542-

551. doi: 10.1037/0033-2909.114.3.542

Bureau of Infrastructure, Transport and Regional Economics. (2011). *Road deaths Australia 2010 statistical summary.* Canberra, Australia.

Burns, N. R., Kremer, S. M., & Baldock, M. R. J. (2005). *Development of peripheral vision tests for driver assessment* (Report No. CASR109). Adelaide, Australia: Centre for Automotive Safety Research.

Burns, N. R., & Nettelbeck, T. (2003). Inspection Time in the structure of cognitive abilities: Where does IT fit? *Intelligence*, *31*(3), 237-255. doi: 10.1016/S0160-2896(02)00120-4

Burns, N. R., & White, M. A. (2007). *ProPerVis.* Adelaide, Australia: School of Psychology, University of Adelaide.

Caird, J. K., Chisholm, S. L., Edwards, C. J., & Creaser, J. I. (2007). The effect of yellow light onset time on older and younger drivers' perception response time (PRT) and intersection behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *10*(5), 383-396. doi: 10.1016/j.trf.2007.03.002

Caird, J. K., Johnston, K. A., Willness, C. R., & Asbridge, M. (2014). The use of meta-analysis or research synthesis to combine driving simulation or naturalistic study results on driver distraction. *Journal of Safety Research*, *49*, 91-96. doi: 10.1016/j.jsr.2014.02.013

Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis and Prevention*, *71*, 311-318. doi: 10.1016/j.aap.2014.06.005

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention*, *40*(4), 1282-1293. doi: 10.1016/j.aap.2008.01.009

Caragata, G. E., Tuokko, H., & Damini, A. (2009). Fit to drive: A pilot study to improve the physical fitness of older drivers. *Activities, Adaptation and Aging*, *33*(4), 240-255. doi: 10.1080/01924780903349080

Carr, D. B., Barco, P. P., Wallendorf, M. J., Snellgrove, C. A., & Ott, B. R. (2011). Predicting road test performance in drivers with dementia. *Journal of the American Geriatrics Society*, *59*(11), 2112-2117. doi: 10.1111/j.1532-5415.2011.03657.x

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484-1525. doi: 10.1016/j.visres.2011.04.012

Carvalho, A., Rea, I. M., Parimon, T., & Cusack, B. J. (2014). Physical activity and cognitive function in individuals over 60 years of age: A systematic review. *Clinical Interventions in Aging*, *9*, 661-682. doi: 10.2147/CIA.S55520

Cassavaugh, N. D., Domeyer, J. E., & Backs, R. W. (2011). Lessons learned regarding simulator sickness in older adult drivers. In *Universal access in human-computer interaction. context diversity* (Vol. 6767, p. 263-269). Springer Berlin Heidelberg. doi: 10.1007/978-3-642-21666-4_29

Casutt, G., Theill, N., Martin, M., Keller, M., & Jäncke, L. (2014). The drive-wise project: Driving simulator training increases real driving performance in healthy older drivers. *Frontiers in Aging Neuroscience*, *6*, 85. doi: 10.3389/fnagi.2014.00085

Chad, K. E., Reeder, B. A., Harrison, E. L., Ashworth, N. L., Sheppard, S. M., Schultz, S. L., ... Lawson, J. A. (2005). Profile of physical activity levels in community-dwelling older adults. *Medicine and Science in Sports and Exercise*, *37*(10), 1774-1784. doi: 10.1249/01.mss.0000181303.51937.9c

Chalacas, E. (2015). *Driving simulator validity, usability and user acceptability for young drivers* (Unpublished Honours Psychology Thesis). The University of Adelaide, Adelaide, Australia.

Chao, Y. C. E., Zhao, Y., Kupper, L. L., & Nylander-French, L. A. (2008). Quantifying the relative importance of predictors in multiple linear regression analyses for public health studies. *Journal of Occupational and Environmental Hygiene*, *5*(8), 519-529. doi: 10.1080/15459620802225481

Charlton, J., Langford, J., Fildes, B. N., Oxley, J., Bohensky, M., Koppel, S., & Congiu, M. (2008). *Trial of a model licence re-assessment procedure for older drivers: Stage 3 research* (Report No. AP-R322/08). Sydney, Australia: Austroads Incorporated.

Charman, W. N. (1997). Vision and driving - a literature review and commentary. *Ophthalmic and Physiological Optics*, *17*(5), 371-391. doi: 10.1016/S0275-5408(97)00014-8

Chevan, A., & Sutherland, M. (1991). Hierarchical partitioning. *The American Statistician*, *45*(2), 90-96. doi: 10.2307/2684366

Clarke, D. D., Ward, P., Bartle, C., & Truman, W. (2010). Older drivers' road traffic crashes in the uk. *Accident Analysis and Prevention*, *42*(4), 1018-1024. doi: 10.1016/j.aap.2009.12.005

Classen, S., Bewernitz, M., & Shechtman, O. (2011). Driving simulator sickness: An evidence-based review of the literature. *American Journal of Occupational Therapy*, *65*(2), 179-188. doi: 10.5014/ajot.2011.000802

Classen, S., & Brooks, J. (2014). Driving simulators for occupational therapy screening, assessment, and intervention. *Occupational Therapy in Health Care*, *28*(2), 154-162. doi: 10.3109/07380577.2014.901590

Clay, O. J., Wadley, V. G., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, K. K. (2005). Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: Current and future implications. *Optometry and Vision Science*, *82*(8), 724-731. doi: 10.1097/01.opx.0000175009.08626.65

Colbert, L. H., Matthews, C. E., Havighurst, T. C., Kim, K., & Schoeller, D. A. (2011). Comparative validity of physical activity measures in older adults. *Medicine and Science in Sports and Exercise*, *43*(5), 867-876. doi: 10.1249/MSS.0b013e3181fc7162

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, *14*(2), 125-130. doi: 10.1111/1467-9280.t01-1-01430

Corlett, E. Y., & White, M. A. (2002). *Age related decrements in peripheral vision and crowding* (Unpublished Masters Psychology Thesis). The University of Adelaide, Adelaide, Australia.

Cosman, J. D., Lees, M. N., Lee, J. D., Rizzo, M., & Vecera, S. P. (2012). Impaired attentional disengagement in older adults with useful field of view decline. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *67 B*(4), 405-412. doi: 10.1093/geronb/gbr116

Cox, D. J., Gonder-Frederick, L. A., Kovatchev, B. P., Julian, D. M., & Clarke, W. L. (2000). Progressive hypoglycemia's impact on driving simulation performance: Occurrence, awareness, and correction. *Diabetes Care*, *23*(2), 163-170. doi: 10.2337/diacare.23.2.163

Crisler, M. C., Brooks, J. O., Venhovens, P. J., Healy Jr, S. L., Hirth, V. A., McKee, J. A., & Duckworth, K. (2011). Seniors' and physicians' attitudes toward using driving simulators in clinical settings. *Occupational Therapy in Health Care*, *26*(1), 1-15. doi: 10.3109/07380577.2011.634889

Crizzle, A. M., Classen, S., Bédard, M., Lanford, D., & Winter, S. (2012). MMSE as a predictor of on-road driving performance in community dwelling older drivers. *Accident Analysis and Prevention*, *49*, 287-292. doi: 10.1016/j.aap.2012.02.003

Daigneault, G., Joly, P., & Frigon, J. Y. (2002). Executive functions in the evaluation of accident risk of older drivers. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 221-238. doi: 10.1076/jcen.24.2.221.993

Dawson, J. D., Uc, E. Y., Anderson, S. W., Johnson, A. M., & Rizzo, M. (2010). Neuropsychological predictors of driving errors in older adults. *Journal of the American Geriatrics Society*, *58*(6), 1090-1096. doi: 10.1111/j.1532-5415.2010.02872.x

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., ... Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135-152. doi: 10.1093/bmb/ldp033

Deary, I. J., Johnson, W., & Starr, J. M. (2010). Are processing speed tasks biomarkers of cognitive aging? *Psychology and Aging*, *25*(1), 219-228. doi: 10.1037/a0017750

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the united kingdom health and lifestyle survey. *Psychology and Aging*, *21*(1), 62-73. doi: 10.1037/0882-7974.21.1.62

Devlin, A., McGillivray, J., Charlton, J., Lowndes, G., & Etienne, V. (2012). Investigating driving behaviour of older drivers with mild cognitive impairment using a portable driving simulator. *Accident Analysis and Prevention*, *49*, 300-307. doi: 10.1016/j.aap.2012.02.022

de Waard, D., Dijksterhuis, C., & Brookhuis, K. A. (2009). Merging into heavy motorway traffic by young and elderly drivers. *Accident Analysis and Prevention*, *41*(3), 588-597. doi: 10.1016/j.aap.2009.02.011

de Winter, J. C. F., de Groot, S., Mulder, M., Wieringa, P. A., Dankelman, J., & Mulder, J. A. (2009). Relationships between driving simulator performance and driving test results. *Ergonomics*, *52*(2), 137-153. doi: 10.1080/00140130802277521

de Winter, J. C. F., & Dodou, D. (2010). The driver behaviour questionnaire as a predictor of accidents: A meta-analysis. *Journal of Safety Research*, *41*(6), 463-470. doi: 10.1016/j.jsr.2010.10.007

Dickerson, A. E. (2013). Driving assessment tools used by driver rehabilitation specialists:

Survey of use and implications for practice. *American Journal of Occupational Therapy*, *67*(5), 564-573. doi: 10.5014/ajot.2013.007823

Dickerson, A. E., Meuel, D. B., Ridenour, C. D., & Cooper, K. (2014). Assessment tools predicting fitness to drive in older adults: A systematic review. *American Journal of Occupational Therapy*, *68*(6), 670-680. doi: 10.5014/ajot.2014.011833

Dickerson, A. E., Reistetter, T. A., Burhans, S., & Apple, K. (2016). Typical brake reaction times across the life span. *Occupational Therapy in Health Care*, *30*(2), 115-123. doi: 10.3109/07380577.2015.1059971

Dinger, M. K., Oman, R. F., Taylor, E. L., Vesely, S. K., & Able, J. (2004). Stability and convergent validity of the Physical Activity Scale for the Elderly (PASE). *Journal of Sports Medicine and Physical Fitness*, *44*(2), 186-192.

Domeyer, J. E., Cassavaugh, N. D., & Backs, R. W. (2013). The use of adaptation to reduce simulator sickness in driving assessment and research. *Accident Analysis and Prevention*, *53*, 127-132. doi: 10.1016/j.aap.2012.12.039

Eckert, K. A., Shi, Z., Taylor, A. W., Wittert, G., Price, K., & Goldney, R. D. (2013). Learning from an epidemiological, population-based study on prescribed medicine use in adults. *Pharmacoepidemiology and Drug Safety*, *22*(3), 271-277. doi: 10.1002/pds.3380

Edwards, C. J., Creaser, J. I., Caird, J. K., Lamsdale, A. M., & Chisholm, S. L. (2004). Older and younger driver performance at complex intersections: Implications for using perception-response time and driving simulation. In *Proceedings of the second international driving symposium on human factors in driver assessment, training and vehicle design* (p. 33-38). Park City, UT.

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L., & Ball, K. K. (2006). The useful field of view test: Normative data for older adults. *Archives of Clinical Neuropsychology*, *21*(4), 275-286. doi: 10.1016/j.acn.2006.03.001

Edwards, J. D., Valds, E. G., Peronto, C., Castora-Binkley, M., Alwerdt, J., Andel, R., & Lister, J. J. (2015). The efficacy of insight cognitive training to improve useful field of view performance: A brief report. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *70*(3), 417-422. doi: 10.1093/geronb/gbt113

Edwards, J. D., Vance, D. E., Wadley, V. G., Cissell, G. M., Roenker, D. L., & Ball, K. K. (2005). Reliability and validity of useful field of view test scores as administered by personal computer. *Journal of Clinical and Experimental Neuropsychology*, *27*(5), 529-543. doi: 10.1080/13803390490515432

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a speed of processing intervention to near and far cognitive functions. *Gerontology*, *48*(5), 329-340.

Edwards, J. D., Wadley, V. G., Vance, D. E., Wood, K., Roenker, D. L., & Ball, K. K. (2005). The impact of speed of processing training on cognitive and everyday performance. *Aging and Mental Health*, *9*(3), 262-271. doi: 10.1080/13607860412331336788

Eggermont, L. H. P., Milberg, W. P., Lipsitz, L. A., Scherder, E. J. A., & Leveille, S. G. (2009). Physical activity and executive function in aging: The MOBILIZE Boston study. *Journal of the American Geriatrics Society*, *57*(10), 1750-1756. doi: 10.1111/j.1532-5415.2009.02441.x

Elliott, D. B., Sanderson, K., & Conkey, A. (1990). The reliability of the Pelli-Robson contrast sensitivity chart. *Ophthalmic and Physiological Optics*, *10*(1), 21-24. doi: 10.1016/0275-5408(90)90124-H

Emerson, J. L., Johnson, A. M., Dawson, J. D., Uc, E. Y., Anderson, S. W., & Rizzo, M. (2012). Predictors of driving outcomes in advancing age. *Psychology and Aging*, *27*(3), 550-559. doi: 10.1037/a0026359

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2 SPEC. ISS.), 97-120. doi: 10.1016/j.trf.2005.04.012

Erickson, K. I., Hillman, C. H., & Kramer, A. F. (2015). Physical activity, brain, and cognition. *Current Opinion in Behavioral Sciences*, *4*, 27-32. doi: 10.1016/j.cobeha.2015.01.005

Evans, G., & Nettelbeck, T. (1993). Inspection Time: A flash mask to reduce apparent movement effects. *Personality and Individual Differences*, *15*(1), 91-94. doi: 10.1016/0191-8869(93)90045-5

Falck, R. S., McDonald, S. M., Beets, M. W., Brazendale, K., & Liu-Ambrose, T. (2016). Measurement of physical activity in older adult interventions: A systematic review. *British Journal of Sports Medicine*, *50*(8), 464-470. doi: 10.1136/bjsports-2014-094413

Fildes, B. N. (2008). Future directions for older driver research. *Traffic Injury Prevention*, *9*(4), 387-393. doi: 10.1080/15389580802272435

Fildes, B. N., Charlton, J., Pronk, N., Langford, J., Oxley, J., & Koppel, S. (2008). An Australasian model license reassessment procedure for identifying potentially unsafe drivers. *Traffic Injury Prevention*, *9*(4), 350-359. doi: 10.1080/15389580802091199

Fildes, B. N., Charlton, J., Pronk, N., Oxley, J., Newstead, S., Oxley, P., ... Keall, M. (2004). *Model licence re-assessment procedure for older drivers: Stage 2 research* (Report No. AP-R259/04). Sydney, Australia: Austroads Incorportated.

Fildes, B. N., Pronk, N., Langford, J., Hull, M., Frith, B., & Anderson, R. (2000). *Model licence re-assessment procedure for older and disabled drivers* (Report No. AP-R176). Sydney, Australia: Austroads Incorporated.

Flanagan, M. B., May, J. G., & Dobie, T. G. (2005). Sex differences in tolerance to visually-induced motion sickness. *Aviation Space and Environmental Medicine*, *76*(7), 642-646.

Foley, J. M., Gooding, A. L., Thames, A. D., Ettenhofer, M. L., Kim, M. S., Castellon, S. A., ... Hinkin, C. H. (2013). Visuospatial and attentional abilities predict driving simulator performance among older HIV-infected adults. *American Journal of Alzheimer's Disease and other Dementias*, *28*(2), 185-194. doi: 10.1177/1533317512473192

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). 'Mini mental state'. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198. doi: 10.1016/0022-3956(75)90026-6

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (Second ed.). Thousand Oaks, CA: SAGE Publications.

Freund, B., & Green, T. R. (2006). Simulator sickness amongst older drivers with and without dementia. *Advances in Transportation Studies*(SPEC.ISSUE), 71-74.

Frittelli, C., Borghetti, D., Iudice, G., Bonanni, E., Maestri, M., Tognoni, G., ... Iudice, A. (2009). Effects of Alzheimer's disease and mild cognitive impairment on driving ability: A controlled clinical study by simulated driving test. *International Journal of Geriatric Psychiatry*, *24*(3), 232-238. doi: 10.1002/gps.2095

Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 179-185.

doi: 10.1016/j.trf.2008.11.002

Gentzler, M. D., & Smither, J. A. (2012). A literature review of major perceptual, cognitive, and/or physical test batteries for older drivers. *Work*, *41*(SUPPL.1), 5381-5383. doi: 10.3233/wor-2012-0825-5381

Gibbons, C., Mullen, N., Weaver, B., Reguly, P., & Bédard, M. (2014). One- and three-screen driving simulator approaches to evaluate driving capacity: Evidence of congruence and participants' endorsement. *American Journal of Occupational Therapy*, *68*(3), 344-352. doi: 10.5014/ajot.2014.010322

Golding, J. F. (2006). Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences*, *41*(2), 237-248. doi: 10.1016/j.paid.2006.01.012

Grabowski, D. C., Campbell, C. M., & Morrisey, M. A. (2004). Elderly licensure laws and motor vehicle fatalities. *Journal of the American Medical Association*, *291*(23), 2840-2846. doi: 10.1001/jama.291.23.2840

Graeber, D. A. (2001). Use of incremental adaptation and habituation regimens for mitigating optokinetic side effects. In *Proceedings of the human factors and ergonomics society 45th annual meeting* (p. 1901-1905). Minneapolis/St.Paul, MN.

Graeber, D. A., & Stanney, K. M. (2002). Gender differences in visually induced motion sickness. In *Proceedings of the human factors and ergonomics society 46th annual meeting* (Vol. 46, p. 2109-2113). SAGE Publications. doi: 10.1177/154193120204602602

Greenwood, P. M., & Parasuraman, R. (2004). The scaling of spatial attention in visual search and its modification in healthy aging. *Perception and Psychophysics*, *66*(1), 3-22. doi: 10.3758/BF03194857

Gregory, T., Callaghan, A., Nettelbeck, T., & Wilson, C. (2009). Inspection Time predicts individual differences in everyday functioning among elderly adults: Testing discriminant validity. *Australasian Journal on Ageing*, *28*(2), 87-92. doi: 10.1111/j.1741-6612.2009.00366.x

Gregory, T., Nettelbeck, T., Howard, S., & Wilson, C. (2008). Inspection Time: A biomarker for cognitive decline. *Intelligence*, *36*(6), 664-671. doi: 10.1016/j.intell.2008.03.005

Grömping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, *17*(1).

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *American Statistician*, *61*(2), 139-147. doi: 10.1198/000313007X188252

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and Inspection Time. *Intelligence*, *29*(6), 523-535. doi: 10.1016/S0160-2896(01)00078-2

Gu, Q., Dillon, C. F., & Burt, V. L. (2010). Prescription drug use continues to increase: US prescription drug data for 2007-2008. *NCHS data brief*(42), 1-8.

Hakamies-Blomqvist, L., Raitanen, T., & O'Neill, D. (2002). Driver ageing does not cause higher accident rates per km. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(4), 271-274.

Hakamies-Blomqvist, L. E. (1993). Fatal accidents of older drivers. *Accident Analysis and Prevention*, *25*(1), 19-27. doi: 10.1016/0001-4575(93)90093-C

Harada, N. D., Chiu, V., King, A. C., & Stewart, A. L. (2001). An evaluation of three self-report physical activity instruments for older adults. *Medicine and Science in Sports and Exercise*, *33*(6), 962-970.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334-337. doi: 10.1038/383334a0

Hoffman, L., Atchley, P., McDowd, J. M., & Dubinsky, R. (2005). The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging*, *20*(4), 610-622. doi: 10.1037/0882-7974.20.4.610

Hoffman, L., & McDowd, J. M. (2010). Simulator driving performance predicts accident reports five years later. *Psychology and Aging*, *25*(3), 741-745. doi: 10.1037/a0019198

Hoffman, L., Yang, X., Bovaird, J. A., & Embretson, S. E. (2006). Measuring attentional ability in older adults: Development and psychometric evaluation of driverscan. *Educational and Psychological Measurement*, *66*(6), 984-1000. doi: 10.1177/0013164406288170

Hollis, A. M., Lee, A. K. W., Kapust, L. R., Phillips, L. K., Wolkin, J., & O'Connor, M. G. (2013). The driving competence of 90-year-old drivers: From a hospital-based driving clinic. *Traffic Injury Prevention*, *14*(8), 782-790. doi: 10.1080/15389588

.2013.777957

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis and Prevention*, *38*(1), 185-191. doi: 10.1016/j.aap.2005.09.007

Ismail, N., Hairi, F., Choo, W. Y., Hairi, N. N., Peramalah, D., & Bulgiba, A. (2015). The Physical Activity Scale for the Elderly (PASE): Validity and reliability among community-dwelling older adults in malaysia. *Asia-Pacific Journal of Public Health*, *27*, 62S-72S. doi: 10.1177/1010539515590179

Jäger, M., Gruber, N., Müri, R., Mosimann, U. P., & Nef, T. (2014). Manipulations to reduce simulator-related transient adverse health effects during simulated driving. *Medical and Biological Engineering and Computing*, *52*(7), 601-610. doi: 10.1007/s11517-014-1162-x

Janke, M. K. (1994). *Age-related disabilities that may impair driving and their assessment: Literature review* (Report). Sacramento, CA: Research and Development Section, Division of Program and Policy Administration, California Department of Motor Vehicles.

Jay, G. M., & Willis, S. L. (1992). Influence of direct computer experience on older adults' attitudes toward computers. *Journals of Gerontology*, *47*(4), 250-257.

Jobe, J. B., Smith, D. M., Ball, K. K., Tennstedt, S. L., Marsiske, M., Willis, S. L., … Kleinman, K. (2001). ACTIVE: A cognitive intervention trial to promote independence in older adults. *Controlled Clinical Trials*, *22*(4), 453-479.

Johnson, D. M. (2005). *Introduction to and review of simulator sickness research* (Report No. 1832). Arlington, VA: US Army Research Institute for the Behavioural and Social Sciences.

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods*, *7*(3), 238-257. doi: 10.1177/1094428104266510

Joseph, P. G., O'Donnell, M. J., Teo, K. K., Gao, P., Anderson, C., Probstfield, J. L., … Yusuf, S. (2014). The mini-mental state examination, clinical factors, and motor vehicle crash risk. *Journal of the American Geriatrics Society*, *62*(8), 1419-1426. doi: 10.1111/jgs.12936

Kaber, D., Zhang, Y., Jin, S., Mosaly, P., & Garner, M. (2012). Effects of hazard exposure and roadway complexity on young and older driver situation awareness and performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15*(5), 600-611. doi: 10.1016/j.trf.2012.06.002

Kaptein, N. A., Theeuwes, J., & van der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record*, *1550*, 30-36. doi: 10.3141/1550-05

Kawano, N., Iwamoto, K., Ebe, K., Aleksic, B., Noda, A., Umegaki, H., … Ozaki, N. (2012). Slower adaptation to driving simulator and simulator sickness in older adults. *Aging Clinical and Experimental Research*, *24*(3), 285-289.

Kelly, M. E., Loughrey, D., Lawlor, B. A., Robertson, I. H., Walsh, C., & Brennan, S. (2014). The impact of exercise on the cognitive functioning of healthy older adults: A systematic review and meta-analysis. *Ageing Research Reviews*, *16*(1), 12-31. doi: 10.1016/j.arr.2014.05.002

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203-220. doi: 10.1207/s15327108ijap0303_3

Keshavarz, B., & Hecht, H. (2014). Pleasant music as a countermeasure against visually induced motion sickness. *Applied Ergonomics*, *45*(3), 521-527. doi: 10.1016/j.apergo.2013.07.009

Klosterhalfen, S., Kellermann, S., Pan, F., Stockhorst, U., Hall, G., & Enck, P. (2005). Effects of ethnicity and gender on motion sickness susceptibility. *Aviation Space and Environmental Medicine*, *76*(11), 1051-1057.

Kolasinski, E. M. (1995). *Simulator sickness in virtual environments* (Report No. 1027). Virginia, U.S.A.: U.S. Army Research Institute for the Behavioural and Social Sciences. Retrieved from `www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA295861`

Koppel, S., & Berecki-Gisolf, J. (2015). Car licensing trends of the Babyboomer cohort (b. 1946−1965) compared to earlier birth cohorts: Effects on the driving population in the state of Victoria, Australia. *Traffic Injury Prevention*, *16*(7), 657-663. doi: 10.1080/15389588.2014.1003817

Koppel, S., Charlton, J. L., Langford, J., Di Stefano, M., MacDonald, W., Vlahodimitrakou, Z., ... Marshall, S. C. (2016). Driving task: How older drivers on-road driving performance relates to abilities, perceptions, and restrictions. *Canadian Journal on Aging*, 1-17. doi: 10.1017/S0714980816000015

Kotecha, A., Spratt, A., & Viswanathan, A. (2008). Visual function and fitness to drive. *British Medical Bulletin*, *87*(1), 163-174. doi: 10.1093/bmb/ldn028

Kowalski, K., Rhodes, R., Naylor, P. J., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: A systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, *9*. doi: 10.1186/1479-5868-9-148

Krishnasamy, C., & Unsworth, C. A. (2011). Normative data, preliminary inter-rater reliability and predictive validity of the Drive Home Maze Test. *Clinical Rehabilitation*, *25*(1), 88-95. doi: 10.1177/0269215510382846

Langford, J. (2008). Usefulness of off-road screening tests to licensing authorities when assessing older driver fitness to drive. *Traffic Injury Prevention*, *9*(4), 328-335. doi: 10.1080/15389580801895178

Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008). Do age-based mandatory assessments reduce older drivers' risk to other road users? *Accident Analysis and Prevention*, *40*(6), 1913-1918. doi: 10.1016/j.aap.2008.08.010

Langford, J., Charlton, J., Bohensky, M., Irving, J., Martin, L., & Fildes, B. (2009). *Older driver model assessment program: Stage 4 validation of screening tests to assess at-risk older drivers* (Report No. AP-R336/09). Sydney, Australia: Austroads Incorporated.

Langford, J., Fitzharris, M., Koppel, S., & Newstead, S. (2004). Effectiveness of mandatory license testing for older drivers in reducing crash risk among urban older Australian drivers. *Traffic Injury Prevention*, *5*(4), 326-335. doi: 10.1080/15389580490509464

Langford, J., Fitzharris, M., Newstead, S., & Koppel, S. (2004). Some consequences of different older driver licensing procedures in Australia. *Accident Analysis and Prevention*, *36*(6), 993-1001. doi: 10.1016/j.aap.2003.11.003

Langford, J., & Koppel, S. (2006). Epidemiology of older driver crashes - identifying older driver risk factors and exposure patterns. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(5), 309-321. doi: 10.1016/j.trf.2006.03.005

Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk - a replication of low mileage bias. *Accident Analysis and Prevention*, *38*(3), 574-578. doi: 10.1016/j.aap.2005.12.002

Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention*, *35*(5), 797-803. doi: 10.1016/s0001-4575(02)00083-0

Lee, H. C., Lee, A. H., & Cameron, D. (2003). Validation of a driving simulator by measuring the visual attention skill of older adult drivers. *American Journal of Occupational Therapy*, *57*(3), 324-328.

Lee, H. C., Lee, A. H., Cameron, D., & Li-Tsang, C. (2003). Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. *Journal of Safety Research*, *34*(4), 453-459. doi: 10.1016/j.jsr.2003.09.007

Lemieux, L., Stinchcombe, A., Gagnon, K., & Bédard, M. (2014). Comparison of simulated driving performance across platforms: from 'Low-Cost Desktop' to 'Mid-Level' driving simulators. *Advances in Transportation Studies*, *34*.

Leversen, J. S. R., Hopkins, B., & Sigmundsson, H. (2013). Ageing and driving: Examining the effects of visual processing demands. *Transportation Research Part F: Traffic Psychology and Behaviour*, *17*, 1-4. doi: 10.1016/j.trf.2012.11.003

Levi, D. M. (2008). Crowding - An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635-654. doi: 10.1016/j.visres.2007.12.009

Lew, H. L., Poole, J. H., Lee, E. H., Jaffe, D. L., Huang, H. C., & Brodd, E. (2005). Predictive validity of driving-simulator assessments following traumatic brain injury: A preliminary study. *Brain Injury*, *19*(3), 177-188. doi: 10.1080/02699050400017171

Lewandowsky, S., Oberauer, K., Yang, L. X., & Ecker, U. K. H. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, *42*(2), 571-585. doi: 10.3758/BRM.42.2.571

Li, G., Braver, E. R., & Chen, L. H. (2003). Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident Analysis and Prevention*, *35*(2), 227-235. doi: 10.1016/S0001-4575(01) 00107-5

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis.* Harlow: Longman Higher Education.

Logan, S. L., Gottlieb, B. H., Maitl, S. B., Meegan, D., & Spriet, L. L. (2013). The Physical Activity Scale for the Elderly (PASE) questionnaire; Does it predict physical health? *International Journal of Environmental Research and Public Health*, *10*(9), 3967-3986. doi: 10.3390/ijerph10093967

Lundberg, C., Hakamies-Blomqvist, L., Almkvist, O., & Johansson, K. (1998). Impairments of some cognitive functions are common in crash-involved older drivers. *Accident Analysis and Prevention*, *30*(3), 371-377. doi: 10.1016/S0001-4575(97)00111-5

Lunsman, M., Edwards, J. D., Andel, R., Small, B. J., Ball, K. K., & Roenker, D. L. (2008). What predicts changes in useful field of view test performance? *Psychology and Aging*, *23*(4), 917-927. doi: 10.1037/a0013466

Lyman, S., Ferguson, S. A., Braver, E. R., & Williams, A. F. (2002). Older driver involvements in police reported crashes and fatal crashes: Trends and projections. *Injury Prevention*, *8*(2), 116-120. doi: 10.1136/ip.8.2.116

Makishita, H., & Matsunaga, K. (2008). Differences of drivers' reaction times according to age and mental workload. *Accident Analysis and Prevention*, *40*(2), 567-575. doi: 10.1016/j.aap.2007.08.012

Marmeleira, J. (2012). An examination of the mechanisms underlying the effects of physical activity on brain and cognition - A review with implications for research. *European Review of Aging and Physical Activity*, *10*(2), 1-12. doi: 10.1007/s11556-012-0105-5

Marmeleira, J., De Melo, F. M. S., Tlemcani, M., & Godinho, M. A. B. (2011). Exercise can improve speed of behavior in older drivers. *Journal of Aging and Physical Activity*, *19*(1), 48-61.

Marmeleira, J., Ferreira, I., Melo, F., & Godinho, M. (2012). Associations of physical activity with driving-related cognitive abilities in older drivers: An exploratory study. *Perceptual and Motor Skills*, *115*(2), 521-533. doi: 10.2466/10.06.25.PMS.115.5.521-533

Marottoli, R. A., Allore, H., Araujo, K. L. B., Iannone, L. P., Acampora, D., Gottschalk, M., … Peduzzi, P. (2007). A randomized trial of a physical conditioning program to enhance the driving performance of older persons. *Journal of General Internal Medicine*, *22*(5), 590-597. doi: 10.1007/s11606-007-0134-3

Marshall, S. C., Man-Son-Hing, M., Bdard, M., Charlton, J., Gagnon, S., Glinas, I., …

Woolnough, A. (2013). Protocol for candrive ii/ozcandrive, a multicentre prospective older driver cohort study. *Accident Analysis and Prevention*, *61*, 245-52. doi: 10.1016/j.aap.2013.02.009

Martin, P. L., Audet, T., Corriveau, H., Hamel, M., D'Amours, M., & Smeesters, C. (2010). Comparison between younger and older drivers of the effect of obstacle direction on the minimum obstacle distance to brake and avoid a motor vehicle accident. *Accident Analysis and Prevention*, *42*(4), 1144-1150. doi: 10.1016/j.aap.2010.01.001

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2014). Cognitive and visual predictors of UFOV performance in older adults. *Accident Analysis and Prevention*, *70*, 74-83. doi: 10.1016/j.aap.2014.03.011

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2015). Simulator sickness and dropout during a driving simulator study: A survival analysis. *Journal of Safety Research*, *55*, 159-69. doi: 10.1016/j.jsr.2015.08.004

Matas, N. A., Nettelbeck, T., & Burns, N. R. (2016). Assessment of driving simulator validity and acceptability for older adult drivers. *Advances in Transportation Studies*, *39*, 69-88.

Mathias, J. L., & Lucas, L. K. (2009). Cognitive predictors of unsafe driving in older drivers: A meta-analysis. *International Psychogeriatrics*, *21*(4), 637-653. doi: 10.1017/S1041610209009119

Mayhew, D. R., Simpson, H. M., Wood, K. M., Lonero, L., Clinton, K. M., & Johnson, A. G. (2011). On-road and simulated driving: Concurrent and discriminant validation. *Journal of Safety Research*, *42*(4), 267-275. doi: 10.1016/j.jsr.2011.06.004

McCauley, M. E. (1984). *Research issues in simulator sickness: Proceedings of a workshop.* Washington, D.C.: National Academy Press.

McGwin Jr, G., & Brown, D. B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis and Prevention*, *31*(3), 181-198.

Menard, S. W. (1995). *Applied logistic regression analysis.* Thousand Oaks, CA: Sage Publications.

Molloy, D. W., Alemayehu, E., & Roberts, R. (1991). Reliability of a standardized mini-mental state examination compared with the traditional mini-mental state examination. *American Journal of Psychiatry*, *148*(1), 102-105. doi: 10.1097/

00002093-199100530-00020

Molloy, D. W., & Standish, T. I. M. (1997). A guide to the standardized mini-mental state examination. *International Psychogeriatrics*, *9*(SUPPL. 1), 87-94. doi: 10.1017/S1041610297004754

Molnar, F. J., Marshall, S. C., Man-Son-Hing, M., Wilson, K. G., Byszewski, A. M., & Stiell, I. (2007). Acceptability and concurrent validity of measures to predict older driver involvement in motor vehicle crashes: An emergency department pilot case-control study. *Accident Analysis and Prevention*, *39*(5), 1056-1063. doi: 10.1016/j.aap.2007.02.003

Molnar, F. J., Patel, A., Marshall, S. C., Man-Son-Hing, M., & Wilson, K. G. (2006). Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: A systematic review. *Journal of the American Geriatrics Society*, *54*(12), 1809-1824. doi: 10.1111/j.1532-5415.2006.00967.x

Mourant, R. R., Rengarajan, P., Cox, D., Lin, Y., & Jaeger, B. K. (2007). The effect of driving environments on simulator sickness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *51*, 1232-1236. doi: 10.1177/154193120705101838

Mourant, R. R., & Thattacherry, T. R. (2000). Simulator sickness in a virtual environments driving simulator. In *Proceedings of the XIVth triennial congress of the international ergonomics association and 44th annual meeting of the human factors and ergonomics association, 'ergonomics for the new millennnium'* (p. 534-537). doi: 10.1177/154193120004400513

Mullen, N. W., Charlton, J., Devlin, A., & Bédard, M. (2011). Simulator validity: Behaviours observed on the simulator and on the road. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (chap. 13). Boca Raton, FL: CRC Press. doi: 10.1201/b10836-10

Mullen, N. W., Chattha, H. K., Weaver, B., & Bédard, M. (2008). Older driver performance on a simulator: Associations between simulated tasks and cognition. *Advances in Transportation Studies*(SPEC.ISSU), 31-42.

Mullen, N. W., Weaver, B., Riendeau, J. A., Morrison, L. E., & Bédard, M. (2010). Driving performance and susceptibility to simulator sickness: Are they related?

American Journal of Occupational Therapy, *64*(2), 288-295. doi: 10.5014/ajot.64.2.288

Nettelbeck, T. (2001). Correlation between Inspection Time and psychometric abilities: A personal interpretation. *Intelligence*, *29*(6), 459-474. doi: 10.1016/S0160-2896(01)00072-1

Nettelbeck, T., & Wilson, C. (2004). The Flynn effect: Smarter not faster. *Intelligence*, *32*(1), 85-93. doi: 10.1016/S0160-2896(03)00060-6

Organisation for Economic Co-operation and Development. (2001). *Ageing and transport: Mobility needs and safety issues*. Paris, France: OECD Publishing. doi: 10.1787/9789264195851-en.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., … Brouwer, R. (2004). *Deliverable 2 - HMI and safety-related driver performance* (Report No. GRD1/2000/25361 S12.319626). Institute for Transport Studies.

Ott, B. R., Davis, J. D., Papandonatos, G. D., Hewitt, S., Festa, E. K., Heindel, W. C., … Carr, D. B. (2013). Assessment of driving-related skills prediction of unsafe driving in older adults in the office setting. *Journal of the American Geriatrics Society*, *61*(7), 1164-1169. doi: 10.1111/jgs.12306

Ott, B. R., Festa, E. K., Amick, M. M., Grace, J., Davis, J. D., & Heindel, W. C. (2008). Computerized maze navigation and on-road performance by drivers with dementia. *Journal of Geriatric Psychiatry and Neurology*, *21*(1), 18-25. doi: 10.1177/0891988707311031

Ott, B. R., Heindel, W. C., Whelihan, W. M., Caron, M. D., Piatt, A. L., & Di-Carlo, M. A. (2003). Maze test performance and reported driving ability in early dementia. *Journal of Geriatric Psychiatry and Neurology*, *16*(3), 151-155. doi: 10.1177/0891988703255688

Owsley, C. (1994). Vision and driving in the elderly. *Optometry and Vision Science*, *71*(12), 727-735. doi: 10.1097/00006324-199412000-00002

Owsley, C., Ball, K. K., McGwin Jr, G., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association*, *279*(14), 1083-1088. doi: 10.1001/jama.279.14.1083

Owsley, C., & McGwin Jr, G. (1999). Vision impairment and driving. *Survey of Oph-*

thalmology, *43*(6), 535-550. doi: 10.1016/S0039-6257(99)00035-1

Owsley, C., & McGwin Jr, G. (2010). Vision and driving. *Vision Research*, *50*(23), 2348-2361. doi: 10.1016/j.visres.2010.05.021

Owsley, C., McGwin Jr, G., Sloane, M. E., Stalvey, B. T., & Wells, J. (2001). Timed instrumental activities of daily living tasks: Relationship to visual function in older adults. *Optometry and Vision Science*, *78*(5), 350-359.

Owsley, C., Sloane, M., McGwin Jr, G., & Ball, K. (2002). Timed instrumental activities of daily living tasks: Relationship to cognitive function and everyday performance assessments in older adults. *Gerontology*, *48*(4), 254-265. doi: 10.1159/000058360

Oxley, J. A., Charlton, J. L., Koppel, S. N., Scully, J., & Fildes, B. N. (2005). Crash risk of older female drivers - contributing factors. In *Annual proceedings - association for the advancement of automotive medicine* (p. 345-360).

Park, G. D., Allen, R. W., Fiorentino, D., Rosenthal, T. J., & Cook, M. L. (2006). Simulator sickness scores according to symptom susceptibility, age, and gender for an older driver assessment study. In *Proceedings of the human factors and ergonomics society annual meeting* (p. 2702-2706). doi: 10.1177/154193120605002607

Park, S. W., Choi, E. S., Lim, M. H., Kim, E. J., Hwang, S. I., Choi, K. I., ... Jung, H. E. (2011). Association between unsafe driving performance and cognitive-perceptual dysfunction in older drivers. *Physical Medicine and Rehabilitation*, *3*(3), 198-203. doi: 10.1016/j.pmrj.2010.12.008

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136-1169. doi: 10.1167/4.12.12

Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, *2*(3), 187-199.

Peña, E. A., & Slate, E. H. (2006). Global validation of linear model assumptions. *Journal of the American Statistical Association*, *101*(473), 341-354. doi: 10.1198/016214505000000637

Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica*, *86*(2-3), 227-272. doi: 10.1016/0001-6918(94)90004-3

Pollatsek, A., Romoser, M. R., & Fisher, D. L. (2012). Identifying and remediating failures of selective attention in older drivers. *Current Directions in Psychological Science*, *21*(1), 3-7. doi: 10.1177/0963721411429459

Pollatsek, A., Vlakveld, W., Kappe, B., Pradhan, A. K., & Fisher, D. L. (2011). Driving simulators as training and evaluation tools. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (chap. 30). Boca Raton, FL: CRC Press. doi: 10.1201/b10836-31

Pringle, H. L., Irwin, D. E., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin and Review*, *8*(1), 89-95. doi: 10.3758/BF03196143

Pringle, H. L., Kramer, A. F., & Irwin, D. E. (2004). Individual differences in the visual representation of scenes. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (p. 165-185). Cambridge, MA: MIT Press.

R Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Retrieved from `http://www.R-project.org/`

R Core Team. (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Retrieved from `http://www.R-project.org/`

Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the roads: A real distinction? *Ergonomics*, *33*(10-11), 1315-1332. doi: 10.1080/00140139008925335

Reger, M. A., Welsh, R. K., Watson, G. S., Cholerton, B., Baker, L. D., & Craft, S. (2004). The relationship between neuropsychological functioning and driving ability in dementia: A meta-analysis. *Neuropsychology*, *18*(1), 85-93. doi: 10.1037/0894-4105.18.1.85

Rensink, R. A. (2000). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, *7*(1-3), 345-376. doi: 10.1080/135062800394847

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245-277. doi: 10.1146/annurev.psych.53.100901.135125

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373. doi: 10.1111/j.1467-9280.1997.tb00427.x

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., & Vecera, S. P. (2009). Change blindness, aging, and cognition. *Journal of Clinical and Experimental Neu-*

ropsychology, *31*(2), 245-256. doi: 10.1080/13803390802279668

Roenker, D. L., Cissell, G. M., Ball, K. K., Wadley, V. G., & Edwards, J. D. (2003). Speed-of-processing and driving simulator training result in improved driving performance. *Human Factors*, *45*(2), 218-233. doi: 10.1518/hfes.45.2.218.27241

Rogé, J., Pébayle, T., Campagne, A., & Muzet, A. (2005). Useful visual field reduction as a function of age and risk of accident in simulated car driving. *Investigative Ophthalmology and Visual Science*, *46*(5), 1774-1779. doi: 10.1167/iovs.04-0540

Ross, L. A., Browning, C., Luszcz, M. A., Mitchell, P., & Anstey, K. J. (2011). Age-based testing for driver's license renewal: Potential implications for older Australians. *Journal of the American Geriatrics Society*, *59*(2), 281-285. doi: 10.1111/j.1532-5415.2010.03241.x

Ross, L. A., Edwards, J. D., O'Connor, M. L., Ball, K. K., Wadley, V. G., & Vance, D. E. (2016). The transfer of cognitive speed of processing training to older adults' driving mobility across 5 years. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *71*(1), 87-97. doi: 10.1093/geronb/gbv022

Roth, D. L., Goode, K. T., Clay, O. J., & Ball, K. K. (2003). Association of physical activity and visual attention in older adults. *Journal of Aging and Health*, *15*(3), 534-547. doi: 10.1177/0898264303253512

Rubin, G. S., Ng, E. S. W., Bandeen-Roche, K., Keyl, P. M., Freeman, E. E., West, S. K., ... Wilson, L. (2007). A prospective, population-based study of the role of visual impairment in motor vehicle crashes among older drivers: The SEE study. *Investigative Ophthalmology and Visual Science*, *48*(4), 1483-1491. doi: 10.1167/Iovs.06-0474

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: Status, limitations, and future directions. *Research Quarterly for Exercise and Sport*, *71*(2 SUPPL.), 1-14.

Samejima, F. (1970). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, *35*(1), 139. doi: 10.1007/BF02290599

Sandlin, D., McGwin, G., & Owsley, C. (2013). Association between vision impairment and driving exposure in older adults aged 70 years and over: A population-based examination. *Acta Ophthalmologica*, *92*(3), e207-12. doi: 10.1111/aos.12050

Schuit, A. J., Schouten, E. G., Westerterp, K. R., & Saris, W. H. M. (1997). Validity of the

Physical Activity Scale for the Elderly (PASE): According to energy expenditure assessed by the doubly labeled water method. *Journal of Clinical Epidemiology*, *50*(5), 541-546. doi: 10.1016/S0895-4356(97)00010-3

Schultheis, M. T., Rebimbas, J., Mourant, R., & Millis, S. R. (2007). Examining the usability of a virtual reality driving simulator. *Assistive Technology*, *19*(1), 1-8. doi: 10.1080/10400435.2007.10131860

Schultheis, M. T., Simone, L. K., Roseman, E., Nead, R., Rebimbas, J., & Mourant, R. (2006). Stopping behavior in a VR driving simulator: A new clinical measure for the assessment of driving. In *28th annual international conference of the ieee engineering in medicine and biology society, EMBS'06* (p. 4921-4924). doi: 10.1109/IEMBS.2006.260389

Schulze-Bonsel, K., Feltgen, N., Burau, H., Hansen, L., & Bach, M. (2006). Visual acuities "hand motion" and "counting fingers" can be quantified with the Freiburg Visual Acuity Test. *Investigative Ophthalmology and Visual Science*, *47*(3), 1236-1240. doi: 10.1167/iovs.05-0981

Sekuler, A. B., Bennett, P. J., & Mamelak, M. (2000). Effects of aging on the useful field of view. *Experimental Aging Research*, *26*(2), 103-120. doi: 10.1080/036107300243588

Shanmugaratnam, S., Kass, S. J., & Arruda, J. E. (2010). Age differences in cognitive and psychomotor abilities and simulated driving. *Accident Analysis and Prevention*, *42*(3), 802-808. doi: 10.1016/j.aap.2009.10.002

Shechtman, O. (2010). Validation of driving simulators. *Advances in Transportation Studies*, *2010 Special Issue*, 53-62. doi: 10.4399/97888548416596

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: A validation study. *Traffic Injury Prevention*, *10*(4), 379-385. doi: 10.1080/15389580902894989

Shechtman, O., Classen, S., Stephens, B., Bendixen, R., Belchior, P., Sandhu, M., ... Davis, E. (2007). The impact of intersection design on simulated driving performance of young and senior adults. *Traffic Injury Prevention*, *8*(1), 78-86. doi: 10.1080/15389580600994321

Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, *37*(3), 197-206.

Simons, D. J., & Ambinder, M. S. (2005). Change blindness: Theory and consequences.

Current Directions in Psychological Science, *14*(1), 44-48. doi: 10.1111/j.0963-7214.2005.00332.x

Sklar, A. L., Boissoneault, J., Fillmore, M. T., & Nixon, S. J. (2014). Interactions between age and moderate alcohol effects on simulated driving performance. *Psychopharmacology*, *231*(3), 557-566. doi: 10.1007/s00213-013-3269-4

Skyving, M., Berg, H. Y., & Laflamme, L. (2009). A pattern analysis of traffic crashes fatal to older drivers. *Accident Analysis and Prevention*, *41*(2), 253-258. doi: 10.1016/j.aap.2008.11.008

Smolensky, M. H., Di Milia, L., Ohayon, M. M., & Philip, P. (2011). Sleep disorders, medical conditions, and road accident risk. *Accident Analysis and Prevention*, *43*(2), 533-548. doi: 10.1016/j.aap.2009.12.004

Snellgrove, C. A. (2005). *Cognitive screening for safe driving competence of people with mild cognitive impairment or early dementia.* Canberra, Australia: Australian Transport Safety Bureau.

Staplin, L., Gish, K. W., Lococo, K. H., Joyce, J. J., & Sifrit, K. J. (2013). The maze test: A significant predictor of older driver crash risk. *Accident Analysis and Prevention*, *50*, 483-489. doi: 10.1016/j.aap.2012.05.025

Staplin, L., Lococo, K. H., Gish, K., & Decina, L. E. (2003). *Model driver screening and evaluation program* (Report No. DOT HS 809 582). Washington, DC: National Highway Transport Safety Administration.

Stewart, A. L., Mills, K. M., King, A. C., Haskell, W. L., Gillis, D., & Ritter, P. L. (2001). Champs physical activity questionnaire for older adults: Outcomes for interventions. *Medicine and Science in Sports and Exercise*, *33*(7), 1126-1141.

Stinchcombe, A., & Gagnon, S. (2013). Aging and driving in a complex world: Exploring age differences in attentional demand while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *17*, 125-133. doi: 10.1016/j.trf.2012.11.002

Stinchcombe, A., Gagnon, S., Zhang, J., Montembeault, P., & Bedard, M. (2011). Fluctuating attentional demand in a simulated driving assessment: The roles of age and driving complexity. *Traffic Injury Prevention*, *12*(6), 576-587. doi: 10.1080/15389588.2011.607479

Stine-Morrow, E. A. L., & Basak, C. (2011). Chapter 10 - cognitive interventions. In K. W. Schaie & L. W. Sherry (Eds.), *Handbook of the psychology of aging (seventh* *edition)* (p. 153-171). San Diego: Academic Press. doi: 10.1016/B978-0-12-380882-0.00010-3

Stoner, H. A., Fisher, D. L., & Mollenhauer Jr, M. (2011). Simulator and scenario factors influencing simulator sickness. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (chap. 14). Boca Raton, FL: CRC Press. doi: 10.1201/b10836-15

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 54, p. 29-58). Burlington: Academic Press. doi: 10.1016/B978-0-12-385527-5.00002-4

Szlyk, J. P., Myers, L., Zhang, Y., Wetzel, L., & Shapiro, R. (2002). Development and assessment of a neuropsychological battery to aid in predicting driving performance. *Journal of Rehabilitation Research and Development*, *39*(4), 483-496.

Tan, B. (2014). *Predicting distracted driving performance in young drivers* (Unpublished Honours Psychology Thesis). The University of Adelaide, Adelaide, Australia.

Therneau, T. (2014). *A package for survival analysis in S.* Retrieved from http://CRAN.R-project.org/package=survival

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922-935.

Trick, L. M., & Caird, J. (2011). Methodological issues when conducting research on older drivers. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (chap. 26). Boca Raton, FL: CRC Press. doi: 10.1201/b10836-27

Trick, L. M., Toxopeus, R., & Wilson, D. (2010). The effects of visibility conditions, traffic density, and navigational challenge on speed compensation and driving performance in older adults. *Accident Analysis and Prevention*, *42*(6), 1661-1671. doi: 10.1016/j.aap.2010.04.005

United Nations, Department of Economic and Social Affairs, Population Division. (2011). *World population prospects: The 2010 revision, volume II: Demographic profiles* (Report No. ST/ESA/SER.A/317). Retrieved from http://esa.un.org/wpp/Documentation/publications.htm

Unsworth, C. A., & Baker, A. (2014). Driver rehabilitation: A systematic review of the types and effectiveness of interventions used by occupational therapists to improve on-road fitness-to-drive. *Accident Analysis and Prevention*, *71*, 106-114. doi: 10.1016/j.aap.2014.04.017

Vance, D. E., Dawson, J., Wadley, V., Edwards, J., Roenker, D., Rizzo, M., & Ball, K. K. (2007). The Accelerate study: The longitudinal effect of speed of processing training on cognitive performance of older adults. *Rehabilitation Psychology*, *52*(1), 89-96. doi: 10.1037/0090-5550.52.1.89

Vance, D. E., Fazeli, P. L., Ball, D. A., Slater, L. Z., & Ross, L. A. (2014). Cognitive functioning and driving simulator performance in middle-aged and older adults with HIV. *Journal of the Association of Nurses in AIDS Care*, *25*(2), e11-e26. doi: 10.1016/j.jana.2013.12.001

Veiel, L. L., Storandt, M., & Abrams, R. A. (2006). Visual search for change in older adults. *Psychology and Aging*, *21*(4), 754-762. doi: 10.1037/0882-7974.21.4.754

Vertesi, A., Lever, J. A., Molloy, D. W., Sanderson, B., Tuttle, I., Pokoradi, L., & Principi, E. (2001). Standardized mini-mental state examination: Use and interpretation. *Canadian Family Physician*, *47*(Oct), 2018-2023.

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. *Journal of Safety Research*, *33*(1), 117-128.

Washburn, R. A., & Ficker, J. L. (1999). Physical Activity Scale for the Elderly (PASE): The relationship with activity measured by a portable accelerometer. *Journal of Sports Medicine and Physical Fitness*, *39*(4), 336-340.

Washburn, R. A., McAuley, E., Katula, J., Mihalko, S. L., & Boileau, R. A. (1999). The Physical Activity Scale for the Elderly (PASE): Evidence for validity. *Journal of Clinical Epidemiology*, *52*(7), 643-651. doi: 10.1016/S0895-4356(99)00049-9

Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The Physical Activity Scale for the Elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology*, *46*(2), 153-162. doi: 10.1016/0895-4356(93)90053-4

Werner, L. K. (2008). *A comparison of the ProPerVis and UFOV as potential fitness-to-drive tests* (Unpublished Honours Psychology Thesis). The University of Adelaide, Adelaide, Australia.

Whelihan, W. M., DiCarlo, M. A., & Paul, R. H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology*, *20*(2), 217-228. doi: 10.1016/j.acn.2004.07.002

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160-168. doi: 10.1016/j.tics.2011.02.005

Willis, S. L. (1996). Everyday cognitive competence in elderly persons: Conceptual issues and empirical findings. *Gerontologist*, *36*(5), 595-601.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495-501. doi: 1038/nrn1411

Wood, J. M. (2002). Aging, driving and vision. *Clinical and Experimental Optometry*, *85*(4), 214-220.

Wood, J. M., Anstey, K. J., Kerr, G. K., Lacherez, P. F., & Lord, S. (2008). A multidomain approach for predicting older driver safety under in-traffic road conditions. *Journal of the American Geriatrics Society*, *56*(6), 986-993. doi: 10.1111/j.1532-5415.2008.01709.x

Wood, J. M., Horswill, M. S., Lacherez, P. F., & Anstey, K. J. (2013). Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. *Accident Analysis and Prevention*, *50*, 1161-1168. doi: 10.1016/j.aap.2012.09.009

Wood, J. M., & Owsley, C. (2014). Useful field of view test. *Gerontology*, *60*(4), 315-318. doi: 10.1159/000356753

Wright, K. (2015). *corrgram: Plot a correlogram.* Retrieved from `https://CRAN.R-project.org/package=corrgram` (R package version 1.8)

Yager, C. E., Cooper, J. M., & Chrysler, S. T. (2012). The effects of reading and writing text-based messages while driving. In *Proceedings of the human factors and ergonomics society 56th annual meeting, HFES 2012* (p. 2196-2200). doi: 10.11771071181312561463

Yang, G., Jaeger, B., & Mourant, R. R. (2006). Driving performance of novice and experienced drivers in lane-change scenarios. In *50th annual meeting of the human factors and ergonomics society, HFES 2006.* doi: 10.1177/154193120605002609

Young, J., Angevaren, M., Rusted, J., & Tabet, N. (2015). Aerobic exercise to improve cognitive function in older people without known cognitive impairment. *Cochrane Database of Systematic Reviews*(4). doi: 10.1002/14651858.CD005381.pub4

Young, S. D., Adelstein, B. D., & Ellis, S. R. (2007). Demand characteristics in assessing motion sickness in a virtual environment: Or does taking a motion sickness questionnaire make you sick? *IEEE Transactions on Visualization and Computer Graphics*, *13*(3), 422-428. doi: 10.1109/TVCG.2007.1041

Zhang, L., Baldwin, K., Munoz, B., Munro, C., Turano, K., Hassan, S., … West, S. K. (2007). Visual and cognitive predictors of performance on brake reaction test: Salisbury eye evaluation driving study. *Ophthalmic Epidemiology*, *14*(4), 216-222. doi: 10.1080/09286580701502988