Predicting cognitive fitness to drive

with Touchscreen DriveSafe DriveAware

Beth J Cheal

Supervisors

Professor Anita Bundy
Faculty of Health Sciences
The University of Sydney, Sydney, Australia

Professor and Department Head,
Department of Occupational Therapy
Colorado State University, Fort Collins, US

Dr Ann-Helen Patomella
Senior Lecturer,
Division of Occupational Therapy
Karolinska Institutet, Solna, Sweden

Dr Justin Scanlan
Senior Lecturer and Course Co-Director
Faculty of Health Sciences
The University of Sydney, Sydney, Australia

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Health Sciences
The University of Sydney
Sydney, Australia

04.09.2018

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purpose.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

I certify that if my candidature is successful, this thesis will be lodged with the Director of University Libraries and made available for immediate use.

Beth Cheal

04.09.2018

Table of Contents

Figures

Tables

Acknowledgements

I would like to thank my supervisors, Professor Anita Bundy, Dr Ann-Helen Patomella and Dr Justin Scanlan for their guidance and support throughout my candidature. Dr Patomella was so generous with her time and support in the critical study design and site setup stages and kindly allowed me to use Performance Drive (P-Drive), which she modified for the Australian and New Zealand context. Dr Scanlan provided much needed help with the data analysis and write up phase. Professor Bundy was there for every step with an amazing ability to distil complex challenges into simple and invaluable advice and direction. She provided a positive and supportive learning environment that motivated independent thinking and encouraged me to achieve.

I would like to thank Anne Madden, Vice President Pearson Australia, New Zealand and Singapore, for employing me as project manager for the touchscreen conversion of DriveSafe DriveAware (DSDA) and kindly allowing me to complete the project as a PhD at The University of Sydney. I would like to thank the many staff at Pearson who provided tireless support and expertise including Haijiang Kuang, Senior Psychometrician, Melinda Cooper, Occupational Therapy Consultant, and members of the Research and Development Team including Christopher Wilson, Sabrina Mooney, Dr Nicki Joshua, and Dr Fiona Brown. I would also like to thank the digital design studio Very Livingston, for their professionalism and high quality software design.

I would like to acknowledge the staff and research participants from the driving clinics across Australia and New Zealand who supported the development of the touchscreen version of DSDA. Individuals within these clinics willingly shared their expertise, collected data, and provided valuable feedback on the test. Their commitment to developing better clinical tools for use in driver assessment and rehabilitation is indicative of their professionalism in clinical practice. These services were: All Ages Occupational Therapy, Victoria; Australian Unity Retirement Living Services, Sydney; Calvary Health Care Sydney Driving Program, Sydney; St. Joseph's Hospital Driving Program, Sydney; Evolution OT, Brisbane; Home & Driving Occupational Therapy Services, Perth; Independent Living Centre WA, Perth; Kevin O'Leary & Associates, Wellington, New Zealand; and, Organisation of Therapy and Rehabilitation Services (OTRS), Hamilton & Auckland, New Zealand.

I would like to thank staff and fellow PhD students at The University of Sydney, who generously assisted with any research dilemmas via the Participation in Everyday Life (PIEL) research meetings. I would also like to thank The University of Sydney Post-Graduate Research Support Scheme (PRSS) for providing funding for my presentation with Dr Patomella at the World Federation of Occupational Therapy Conference in Japan and for funding some equipment used in the research. I would like to thank Dr Lynn Kay, DSDA author. I have certainly stood on her shoulders both in my career as a driver-trained occupational therapist and as a researcher in DSDA.

Finally, I would like to thank my husband Rodney and children, Kirrily, Rowan and Imogen, for being patient and supportive over 5 years of significant time and effort dedicated to achieving a long-held goal. I would also like to thank my lovely mother Ruth Allan who has always encouraged and supported me, and my father Dr Jon Allan, a general practitioner who often described the difficulty of addressing patient fitness-to-drive in medical practice.

Abstract

Driving is a highly valued daily living activity easily disrupted by illness, injury, or age-related changes. General practitioners (GPs) are ultimately responsible for determining medical fitness to drive but lack valid and reliable tools. The desktop (original) version of DriveSafe DriveAware (DSDA) is a promising, valid and reliable test but it is not practical for medical practice. The concern of this thesis was the conversion of original DSDA into a touchscreen test of cognitive fitness to drive for GPs and occupational therapists to use in predicting patient driving performance without on-road testing.

Because we were transitioning from a test administered and scored by a trained assessor to one where patient touch responses were scored in-app, we needed to develop an automatic data collection and scoring system that reflected the decision that would otherwise have been made by an expert-rater. We tested usability of the system with older adults then examined set scoring parameters to determine if these discriminated at-risk from comparison drivers. Results indicated the system we designed reflected the decisions that would have been made by a trained assessor.

Next, we conducted a study to examine the internal validity, reliability, and predictive validity of data gathered with touchscreen DSDA. The criterion measure was outcome of a standardised occupational therapy on-road assessment. Rasch analysis provided evidence that touchscreen DSDA had retained the strong psychometric properties of original DSDA. However, results of a discontinued feasibility study indicated potential barriers to uptake of the test by physicians. Touchscreen DSDA may rather be a tool for occupational therapists to use in driver screening and addressing the community mobility need of their clients. Research indicates no there is no one best tool for screening fitness to drive. However, the thesis findings indicate touchscreen DSDA is one useful tool.

List of Abbreviations

APP	Application (iPad)
DDFT	Deceleration Detection Flicker Test
DSDA	DriveSafe DriveAware
FTDS	Fitness-to-Drive Screening Measure
GDE-Framework	Goals for Driver Education Framework
GP	General Practitioner
HCDT	Hazard Change Detection Test
IADL	Instrumental Activities of Daily Living
MFM	Modified Flicker Method test
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MVPT	Motor-Free Visual Perceptual Test
NPV	Negative Predictive Value
PPV	Positive Predictive Value
SDSA	Stroke Drivers Screening Assessment
SMCTests	Sensory-Motor and Cognitive Tests
TCI	Task-Capability Interface model
TMT-A and -B	Trail Making Test (Part A and B)
UCD	User-Centred Design
UFOV ^{TM}	Useful Field of View test
VRST	Visual Recognition Slide Test (precursor test to DSDA)

CHAPTER 1

Background to the Research

People in many cultures consider driving as one of their most-valued instrumental activities of daily living (IADLs) (i.e., the often-complex activities that support daily life in the home and community) (Al-Hassani & Alotaibi, 2014; Dickerson, Reistetter, & Gaudy, 2012; Fricke & Unsworth, 2001). Driving is important for maintaining independence and connectedness to the community and for providing a sense of self-worth and self-determination (Donorfio, D'Ambrosio, Coughlin, & Mohyde, 2009; White et al., 2012). Gaining a license is considered a rite of passage for the young and older drivers want to drive for as long as possible, often with no plan for cessation (Coxon & Keay, 2015; Kostyniuk & Shope, 2003). However, driving is complex and therefore easily disrupted by illness, injury, or age-related changes. The presence of chronic and multiple medical conditions, particularly among older drivers, is associated with increased crash risk and driving errors (Barco et al., 2015; Charlton et al., 2010; Dobbs, Heller, & Schopflocher, 1998; Kay, Bundy, Clemson, & Jolly, 2008; Marshall, 2008; Marshall & Man-Son-Hing, 2011; Molnar, Patel, Marshall, Man-Son-Hing, & Wilson, 2006; Papa et al., 2014). Despite this, researchers recommend fitness to drive is determined on an individual level, with a focus on functional status rather than diagnosis (Charlton et al., 2010; Dickerson et al., 2017; Dickerson et al., 2007; Laycock, 2011; Marshall, 2008; Marshall & Man-Son-Hing, 2011). 'Fitness to drive' is a driver characteristic defined by the absence of a medical condition or functional deficit (physical, cognitive or sensory-perceptual) that significantly increases crash risk or significantly impairs ability to control a vehicle and conform to road laws (Transportation Research Board, 2016).

The Need for Driver Screening

"Assessing Fitness to Drive" (Austroads & National Transport Commission, 2016) provides physicians in Australia and New Zealand with a decision-making guideline regarding fitness to drive for patients with medical conditions. These guidelines indicate that a practical on-road assessment may be required to determine fitness to drive (Austroads & National Transport Commission, 2016). On-road assessments, usually conducted by occupational therapists with specialised training, are generally considered optimal for determining fitness to drive because functional driving performance is assessed in real traffic. Therefore, the test has high acceptability (Kay, Bundy, Clemson, et al., 2008; Langford, 2008; Laycock, 2011; Wheatley & Di Stefano, 2008). On-road assessment is ideal for patients where there is uncertainty regarding fitness to drive, where vehicle modifications may be required, or where it is clear the patient will benefit from a driver rehabilitation programme (Dickerson, Reistetter, Davis, & Monahan, 2011; Kay, 2008). However, it is not practical for all drivers who have a medical condition to undergo an on-road assessment due to cost, waiting times, and a limited number of qualified assessors (Dickerson & Bédard, 2014; Dickerson et al., 2011; Kay, Bundy, & Clemson, 2009a). A first-level screen is required to allow identification of patients who are either clearly fit or unfit to drive, avoiding unnecessary referral to driving clinics for these groups (Bédard & Dickerson, 2014; Dickerson, 2014b; Dickerson & Bédard, 2014; Kay, Bundy, & Clemson, 2008; Korner-Bitensky, Toal-Sullivan, & von Zweck, 2007b; Molnar, Byszewski, Marshall, & Man-Son-Hing, 2005). If a first-level screen with good predictive ability existed, patients identified as not fit could be advised to discontinue driving and may use their time and monetary resources in other ways (e.g., for alternative transport). Patients identified as fit could return to driving without unnecessary and expensive testing (Bédard & Dickerson, 2014). Only drivers with inconclusive results would require referral for on-road testing, ensuring limited driving clinic resources are dedicated to those with uncertain outcomes or who will benefit from driver rehabilitation (Gibbons et al., 2017).

Driver screening is primarily of benefit for determining cognitive fitness to drive because drivers with physical impairments often still require referral to driving clinics to determine if vehicle modifications or license restrictions are needed (Kay, 2008). For more than 25 years, researchers have examined clinical tests to identify a cognitive screen that can accurately predict driving performance without testing drivers on-road (Kay, Bundy, Clemson, Cheal, & Glendenning, 2012). The desk top (original) version of DriveSafe DriveAware (DSDA) is one test that has shown sufficient sensitivity and specificity to accurately predict on-road performance (Allan, Coxon, Bundy, Peattie, & Keay, 2015; Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2008; Kay et al., 2009a; Kay, Bundy, & Clemson, 2009b; Kay et al., 2012) and test-retest reliability (O'Donnell, Morgan, & Manuguerra, 2018). Scores trichotomise drivers into "unsafe", "further testing", and "safe" categories. In this research, safe driving was operationalised as meeting license authority standards for driving with no intervention required; unsafe driving represented failure to meet license authority guidelines for driving and assessor judgement that the participant had no potential for improvement (Kay, Bundy, & Clemson, 2008; Kay et al., 2009a). The original DSDA identifies safe drivers with a sensitivity of between 91% and 93%, unsafe drivers with a specificity of 96%, and correctly classifies 90% of drivers (Hines & Bundy, 2014; Kay et al., 2009a). Driver-trained occupational therapists have used original DSDA, and its precursor test Visual Recognition Slide Test (VRST), as part of a clinical assessment of cognitive fitness to drive in Australia for more than 25 years (Kay, Bundy, & Clemson, 2008).

General practitioner context. Administration of original DSDA is limited to driver-trained occupational therapists because verbal responses need to be interpreted by a trained professional (Kay & Bundy, 2009). In most countries, general practitioners (GPs), rather than occupational therapists, are the professionals ultimately responsible for determining fitness to drive for older and medically impaired drivers. GPs are in an ideal position to screen drivers because patients usually present to them in the first instance; they are required to fill out license authority medical forms (Dobbs et al., 1998; Sims, Rouse-Watson, Schattner, Beveridge, & Jones, 2012); and there is mandatory reporting of medically 'at risk' drivers in jurisdictions of many countries including the US, Canada, and Australia (Austroads & National Transport Commission, 2016; Jang et al., 2007). Surveys show that GPs believe they should be responsible for making determinations about medical fitness to drive but lack valid and reliable driver screens that are practical for use in medical practice (Classen et al., 2016; Dobbs et al., 1998; Fildes, 2008; Jang et al., 2007; Marshall, Demmings, Woolnough, Salim, & Man-Son-Hing, 2012; Molnar et al., 2006; Sims et al., 2012; Wilson & Kirby, 2008; Woolnough et al., 2013; Yale, Hansotia, Knapp, & Ehrfurth, 2003).

The original DSDA is a promising screen for GPs because it is face valid, sufficiently predictive for trichotomising patients (i.e., "safe", "requires further testing", and "unsafe") via two evidence-based cut-off scores and was developed against the optimum criterion measure of a standardised occupational therapy on-road assessment. However, the original version of DSDA is not suitable for medical practice because it takes 20-30 minutes and requires a trained administrator using a data projector and screen. The DSDA authors, Dr Lynn Kay and Professor Anita Bundy, recognised the need to convert the test to touchscreen so it was suitable for general practice (Kay, 2008). The process of conversion of original DSDA into a brief, valid, and practical driver screen for GPs is the subject of this thesis. GPs, medical specialists, and licensing authorities often refer to occupational therapists for further assessment of patient fitness to drive where the outcome of driver screening is uncertain. Therefore, occupational therapists also have an important role in predicting patient fitness to drive.

Occupational therapy context. Driver assessment and rehabilitation is recognized as an important practice domain for occupational therapists (American Occupational Therapy Association, 2014; Canadian Association of Occupational Therapists, 2009; Occupational Therapy Australia, 2015). The profession is widely recognized as a leader in the field: contributing significantly to the development of clinical practice guidelines, assessment protocols, position and consensus statements, standardised assessments, and theoretical constructs of driving behaviour (American Occupational Therapy Association, 2014; Bédard & Dickerson, 2014; Canadian Association of Occupational Therapists, 2009; Carr, Schwartzberg, Manning, & Sempek, 2010; Di Stefano & Macdonald, 2010; Dickerson, 2013, 2014b; Kay et al., 2009a; Kay, Bundy, Clemson, et al., 2008; Occupational Therapy Australia, 2015; Patomella, Tham, Johansson, & Kottorp, 2010; Vrkljan, Myers, Crizzle, Banchard, & Marshall, 2013).

Occupational therapists in Australia and New Zealand are required to attend additional training for certification to conduct on-road assessments (designated as "driver-trained occupational therapists") (Austroads & National Transport Commission, 2016; Occupational Therapy Australia, 2015). However, it is important for all occupational therapists to address driving regardless of their area of expertise because driving is a critical IADL (Dickerson, 2013; Dickerson & Bédard, 2014; Dickerson, Meuel, Ridenour, & Cooper, 2014; Vrkljan, McGrath, & Letts, 2011). Occupational therapists who are not driver-trained are frequently asked to address driving with their patients. This includes advising medical teams regarding fitness to drive, identifying which patients require further testing, advising patients and family regarding community mobility options, and providing patient support in cessation of driving (Dickerson, 2013, 2014b; Korner-Bitensky et al., 2007b). Researchers

(Bédard & Dickerson, 2014; Dickerson, 2014b; Korner-Bitensky, Toal-Sullivan, & von Zweck, 2007a; Vrkljan et al., 2011) asserted that occupational therapists who are not specialists in driving have the skill and competence to apply screen results to determine fitness to drive for those at the extremities (i.e., clearly fit or unfit to drive); and can determine if, and when, to refer to specialist services. All occupational therapists need to understand how to appropriately screen fitness to drive (Dickerson, 2013; Dickerson & Bédard, 2014; Dickerson et al., 2011) and require access to suitable and valid cognitive fitness-to-drive tools so reliable advice is given in this high-stakes area and patients are not subjected to extensive cognitive testing unsuitable for predicting driving performance (Dickerson, 2013; Vrkljan et al., 2011).

Survey results have indicated that many driver-trained occupational therapists are not confident using clinical tools alone to determine fitness to drive, preferring to take patients on-road regardless of the outcome of off-road testing (Korner-Bitensky, Bitensky, Sofer, Man-Son-Hing, & Gelinas, 2006; Korner-Bitensky, Sofer, Gelinas, & Mazer, 1998; Vrkljan et al., 2013). This is consistent with clinical practice in Australia and New Zealand where all drivers referred for a driving assessment are typically taken on-road. Respondents in one study indicated that on-road assessment was always conducted because it was not possible to predict driving performance regardless of how poorly patients performed off-road and because patients and family would not accept license cancellation unless based on results of a practical driving test (Korner-Bitensky et al., 1998). However, researchers (Dickerson & Bédard, 2014; Kay et al., 2009a; Korner-Bitensky & Sofer, 2009) have proposed that on-road testing should not be automatic due to the risk of collision, and the pressure to ensure health services reflect best practice and are cost effective. Surveys have indicated significant variation in off-road assessment clinical practice among clinicians, with low use of computerised assessments and selection of tests that may not be evidence-based or driving related (Dickerson, 2013; Korner-Bitensky et al., 2006). Whilst the focus of this thesis is on the development of a driver screen for GPs, I hope the research will also provide driver-trained occupational therapists with a valid, driving related clinical screen to minimize unnecessary and expensive on-road testing: thus ensuring on-road assessment and rehabilitation are available for those who will benefit most. This is particularly important in Australia where health service providers are increasingly reluctant to subsidise the cost of driver assessment and require driving clinics to be self-funding.

Personal statement. The research reported in this thesis developed out of my experience as an occupational therapist specialising in driver assessment and rehabilitation. I had used the original version of DSDA and VRST in my clinical assessment of patient fitness to drive since 1999. I had participated in DSDA and VRST research as a staff member in the Driver Rehabilitation & Fleet Safety Services Clinic at The University of Sydney. Pearson published the original version of DSDA as a cognitive fitness-to-drive assessment for driver-trained occupational therapists in 2009. Pearson and the DSDA authors recognised the need to convert the test to touchscreen so GPs could use it as a screen to guide decision making regarding patient fitness to drive (Kay, 2008).

I was employed by Pearson Australia to project manage the touchscreen DSDA conversion in 2012. My brief was to convert original DSDA into a short, valid, accurate, touchscreen test suitable for a broad range of diagnostic groups, not tied to the Australian context, and user-friendly for GPs, occupational therapists and other health professionals. This involved changing the test from one requiring a trained administrator using a computer, data projector and screen, to one that was largely self-administered via iPad by health professionals without specialised training. The challenge was to keep the touchscreen version as similar as possible to the original version, whilst retaining the test's predictive validity and reliability. Pearson agreed for this project to be conducted as a PhD under the supervision of Professor Anita Bundy at The University of Sydney. Professor Bundy is an author of DSDA and receives royalties from sale of the test. While she had a vested interest in the completion of the project, pass or fail of the thesis was determined by external examiners. Further I was hired by Pearson to develop a valid and reliable iPad application but Pearson were not invested in me receiving a pass mark on the thesis per se.

As an employee of Pearson and project manager for the conversion of original DSDA to touchscreen, I had vested interest in the success of the project. However, I was not involved in the data collection at the participating driving clinics based around Australia and New Zealand. I collected data for four cases via my Sydney-based driving clinic, Rehab on Road. As with all assessors, I was blind to the results of touchscreen DSDA at the time of the assessment. The statistical analysis was initially conducted by Haijiang Kuang, Senior Psychometrician at Pearson US, to comply with Pearson's policy that all statistical analysis for all products is conducted internally prior to publication to ensure accuracy and validity. Statistical accuracy and validity were also checked by my three university supervisors.

Thesis Overview

Chapter 2 comprises a critical review of existing driving models that I conducted to identify key concepts and mechanisms potentially underlying DriveSafe performance so I could retain these in the digital DSDA test conversion. Generally, tests are developed to reflect a model. However, VRST was developed over 25 years ago by an occupational therapist trying to develop a clinical assessment to reflect real-world driving situations. In retrospect, I wished to examine existing driving models to place the touchscreen DriveSafe subtest within the context of an overall driving model.

Chapter 3 comprises a critical review of the literature that I conducted to determine if GPs needed a fitness-to-drive screen to justify project funding, and to identify how such a screen should be designed. I identified existing fitness-to-drive tools, evaluated whether these were sufficiently accurate to predict driving performance, and determined whether these tests were feasible for medical practice. I also identify suitable touchscreen test design criteria for GPs to inform digital DSDA design and research project design.

Chapter 4 comprises an extended methodology in which I describe the process of developing touchscreen DSDA, including identification of suitable test design criteria to ensure the screen was optimally designed for the intended context (general practice), the intended administrators (GPs and other health professionals), and the group most likely to take touchscreen DSDA (older adults).

Chapter 5 comprises a manuscript titled “Usability testing of touchscreen DriveSafe DriveAware with older adults: A cognitive fitness-to-drive screen”. This study was conducted to test usability of touchscreen DSDA with older adults concurrently with the iPad application (app) design, programming, and evaluation. I sought to make the touchscreen version of DSDA as similar as possible to the original version in order to retain validity. However, I was transitioning from a test administered by a trained assessor who interpreted and scored patient verbal responses, to a test where patient touchscreen responses were scored automatically in-app. Therefore, I needed to develop an automatic data collection and scoring system that reflected the decision that would otherwise have been made by an expert rater. I took a user-centred design approach: involving older adults in each stage of the development to ensure touchscreen DSDA achieved the project aims and was user friendly for the intended context and user groups.

Chapter 6 comprises a manuscript titled “Converting the DriveSafe subtest of DriveSafe DriveAware for touchscreen administration”. This study was conducted as part of the development of the automatic data collection and scoring system for DriveSafe, the more complex of the two DSDA subtests. I applied a structured process to determine what constituted correct scores. I then examined the resulting scoring parameters to determine if these discriminated at-risk drivers from a comparison sample, prior to psychometric evaluation.

Chapter 7 presents a manuscript titled “Predicting fitness to drive for medically at-risk drivers using touchscreen DriveSafe DriveAware”. I conducted this study to examine the internal validity, reliability, and predictive validity of data gathered with touchscreen DSDA, and to set cut-off scores for trichotomising drivers based on the likelihood of passing an on-road assessment. The outcome of a standardised occupational therapy on-road assessment was the criterion measure.

In Chapter 8 I discuss a study I commenced to test the feasibility of touchscreen DSDA for medical practice. I discontinued this study for the various reasons discussed. Two doctors implemented touchscreen DSDA into their clinics and completed an interview regarding their experiences. Whilst these results only represent the views of two doctors, they raise important issues that require further exploration in future research.

Finally, in Chapter 9, I synthesise the conclusions of the literature reviewed and the findings reported in the three manuscripts. I discuss the limitations of the research, implications for practice, and directions for future research.

CHAPTER 2

Critical Review of Existing Driving Models

Becky Zropf, an occupational therapist who established driver assessment and rehabilitation training for occupational therapists at The University of Sydney in 1989, developed VRST, the precursor test to the DriveSafe subtest (Cheal & Kuang, 2015). Ms Zropf reported there was no suitable driving-related cognitive assessment available at the time (personal communication via telephone and email, February 23, 2015). Therefore, she developed a test to simulate the need to quickly and accurately identify hazards that a driver should notice for safe decision-making. She introduced a time pressure as in real driving, along with areas of low contrast (e.g., shadows and pedestrians wearing dark clothes), to help identify drivers likely to have difficulty on-road. In short, she developed VRST to reflect the visual search ability required in real world driving. In retrospect, as I began this thesis, I wished to examine existing driving models to: (a) place the touchscreen DriveSafe subtest within the context of an established overall driving model; and (b) identify key concepts and mechanisms underlying DriveSafe performance. I needed this knowledge to ensure that I retained critical test mechanisms in the touchscreen DSDA conversion. Such knowledge can also be applied in future revision of the test (e.g., if Pearson wished to introduce moving scenes, increase item complexity, or transition to alternative modes of digital administration).

Researchers (Carsten, 2007; de Winter & Happee, 2012; Michon, 1985; Ranney, 1994) have argued that a successful driving model is one that is simple, explicit, usable, validated, and predictive. In addition to these criteria, I sought a model consistent with DriveSafe theoretical concepts and able to identify how components of DriveSafe mechanisms interact to produce behaviour (Glennan, 2005). A mechanism for a behaviour can be defined as “a complex system that produces that behaviour by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations” (Glennan, 2005, p. 445). Additionally, I sought a model capable of explaining and predicting the driving behaviour of both healthy and functionally impaired drivers (Fox, Bowden, & Smith, 1998; Kay, 2008). The focus of this portion of the literature review was on models that would potentially explain the DriveSafe subtest.

I reviewed literature via four online databases: Medline via OvidSP (1946 to present), Cinahl via Ebsco (1982 to present); Scopus; and, PsycTESTS. My search terms included: automobile driving; driver behaviour models; motivational models; trait models; traffic environment; traffic psychology; and, cognition. I did not apply a year or locality limitation. I reviewed the reference lists of relevant articles along with articles citing these publications. By hand searching the reference lists, I identified several key conference presentations and frequently cited texts that I reviewed.

Classification of Models

Over many decades, researchers and theorists have proposed a vast number of driving models to examine the construct of safe driving. I applied Michon’s (1985) 2-way classification system to organise this analysis (see Figure 2.1).

	Taxonomic	Functional
Input-Output (Behavioural)	Task Analysis	Mechanistic Models Adaptive Control Models - Servo-control - Information flow control
Internal State (Psychological)	Trait Models	Motivational Models Cognitive (Process) Models

Figure 2.1. Michon's (1985) classification of driver behaviour model types (p. 490).

Classification of models by interaction of components. Michon (1985) used the labels "Functional" and "Taxonomic" to distinguish whether components of the model either do or do not interact, respectively (Michon, 1985). The components of functional models do dynamically interact (e.g., mechanistic and motivational models) (de Winter & Happee, 2012). In contrast, taxonomic models present an inventory of facts where elements of the model have no dynamic relationship (de Winter & Happee, 2012; Michon, 1985; Ranney, 1994). Task analysis and Trait Models typify taxonomic models (Michon, 1985).

Classification of models by orientation. Michon (1985) used the labels "Input-output" and "Internal State" to distinguish whether models are oriented either to observable behaviour or to driver motivations respectively. Input-output models are oriented to observable behaviour (e.g., task analysis and mechanistic models) (Michon, 1985; Ranney, 1994). Whereas internal state or psychological models are motivation oriented, making assumptions about the processes occurring in the driver's mind (e.g., trait models and motivational models) (de Winter & Happee, 2012; Michon, 1985).

Input / output models. I excluded models classified as input-output (i.e., both task analysis and mechanistic/adaptive control models) from this review because these models are based on analysis of observable driving behaviour (e.g., traffic following or lane keeping) which are unrelated to DSDA (Michon, 1985). Task analysis models generally provide a detailed description of the driving task in terms of task, performance, and ability requirements (Michon, 1985). Mechanistic / adaptive control models describe a sequence of stages of observable driving behaviour with various inputs and outputs, typically via complex computer simulations and mathematical models (de Winter & Happee, 2012; Michon, 1985; Ranney, 1994). The driver's motivations and cognitive processes are not taken into consideration (de Winter & Happee, 2012; Heikoop, de Winter, van Arem, & Stanton, 2015; Ranney, 1994). DriveSafe measures the construct of awareness of the driving environment, involves visual search, and assumes attention is critical to safe driving; driving is not broken down into component parts (Kay et al., 2009a). Therefore, DSDA does not fit within the input-output paradigm. I required a model capable of understanding human errors and difficulties.

Internal state models. Internal state models generally attempt to explain the whole driving task, taking into consideration the actions of drivers, their mental processes (such as motivations and risk acceptance), and the driving environments (de Winter & Happee, 2012; Lutzenberger & Albayrak, 2013; Ranney, 1994; Shinar & Oppenheim, 2011). Michon (1985) identified three categories of internal state models: motivational, cognitive, and, trait models.

Motivational models. Motivational model theorists view driving as a self-paced task impacted by internal threat-related emotions such as fear and anxiety that affect driving behaviour (Schmidt-Daffy, 2013). These emotions increase or decrease depending on task demands, speed, and capability of the driver (de Winter & Happee, 2012; Fuller, 2005; Ranney, 1994; Schmidt-Daffy, 2013). Motivational models are generally applied to the study of driver behaviour via computer simulations of dynamic driving environments, manipulated to produce changes in traffic density, speed, and risk (Schmidt-Daffy, 2013; Zhang & Kaber, 2013). Ranney (1994) noted that the study of driving in this context may be misdirected since the model assumes driving is determined by the motivations and goals of the driver, which are difficult to simulate in a laboratory (Fox et al., 1998). The most recent and well-cited motivational model is Fuller's Task-Capability Interface (TCI) model (See Figure 2.2) (de Winter & Happee, 2012; Fuller, 2005).

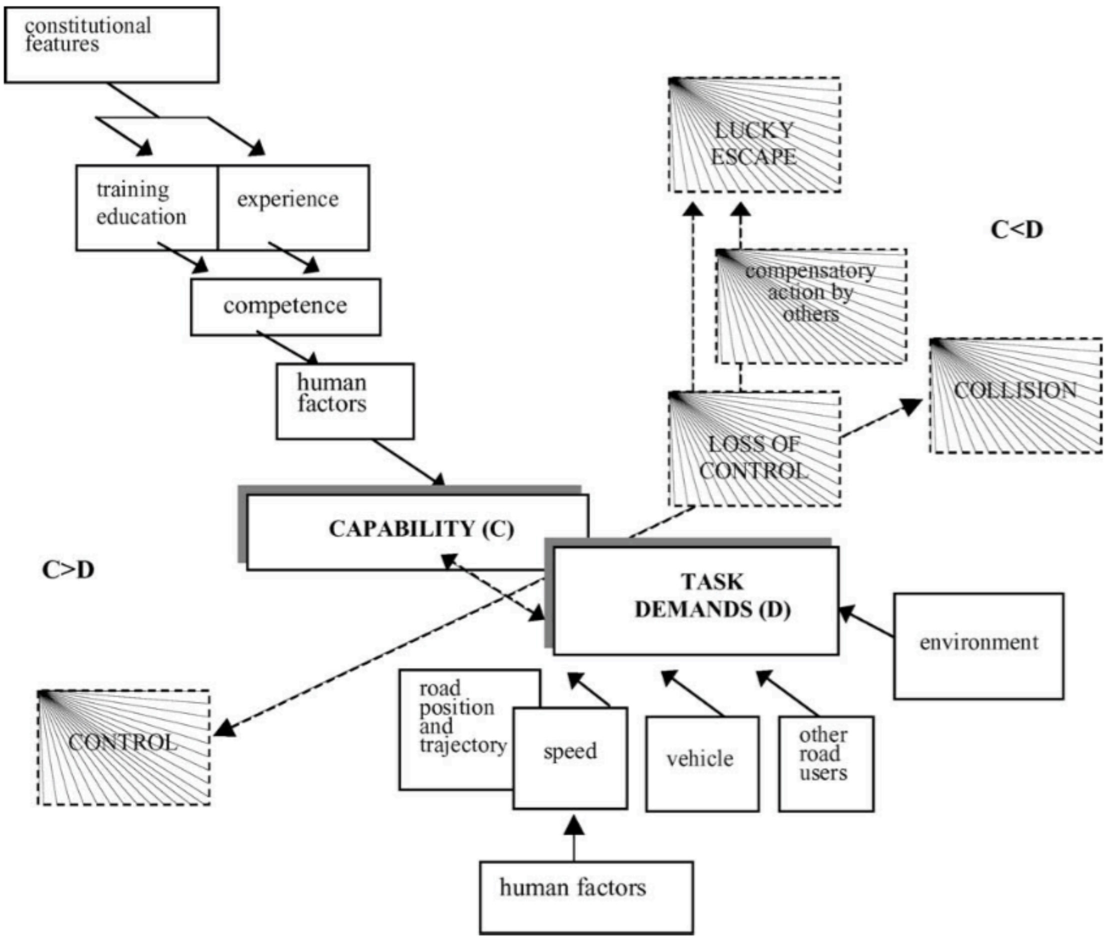


Figure 2.2. Fuller's (2005) Task Capability Interface model (p. 465).

Kay (2008) critiqued the TCI model for its congruence with DSDA and concluded that the emphasis on risk mediated by speed as the critical factor impacting safe driving, was too simplistic. Additionally, whilst concepts such as self-awareness and the difference between learned and conscious driving behaviour are represented, these are not adequately specified, and the model lacks sufficient empirical testing (Kay, 2008). This is consistent with other researchers' (Carsten, 2007; de Winter & Happee, 2012; Ranney, 1994) cautions about applying motivational models to driver assessment due to lack of specification, over-reliance on confirmation, lack of precision, failure to account for the complexity of driving, and lack of quantitative research support. I found that motivational models were not a good fit with DSDA due to an incompatible theoretical basis and insufficient specification of key mechanisms. The DriveSafe test design also precluded application (i.e., presentation of static not dynamic scenes involving visual search and recall of object features).

Cognitive models. Cognitive-model theorists view driving as a hierarchical task, involving automatic and conscious cognitive processes (Fox et al., 1998; Lutzenberger & Albayrak, 2013; Michon, 1985). Michon's (1985) Hierarchical Control Model was the first to conceptualise driving behaviour as a hierarchy of interacting skills influenced by environmental input: namely, Control Level (driving manoeuvres), Manoeuvring Level (adapting to current traffic situations) and Strategic Level (general plans) (Lützenburger & Albayrak, 2013; Michon, 1985). The Transportation Research Board (2016) redefined Michon's (1985) model terms for professionals engaged in driver evaluation and rehabilitation for consistency, and to reflect advances in clinical practice, measurement tools and technology. The operational level of driving relates to control of the vehicle through operation of car controls (based on skills that are over learned and largely automatic) (Transportation Research Board, 2016). The tactical level refers to the manoeuvring control executed over a vehicle to complete a goal directed trip (e.g., maintaining lane position, gap selection and obeying traffic signs). The strategic level relates to trip planning (e.g., goals, route and mode of travel), including accepting the related costs and risks. It also includes adapting the trip in response to obstacles (e.g., a route change due to road closure) (Transportation Research Board, 2016).

Michon's (1985) model was foundational for many subsequent works and contributed significantly to contemporary conceptualisations of driving behaviour (Fox et al., 1998; Lutzenberger & Albayrak, 2013). For example, Dickerson & Bédard (2014) adapted Michon's (1985) model to establish an occupational therapy framework for identifying driving risk and potential for return to driving. Clinical decision making questions were proposed for each level in the heirachy and a check list was designed to identify which client factors may impact safe driving (Dickerson & Bédard, 2014). The authors proposed that generalist occupational therapists could use the tool to guide clinical judgement regarding whether further assessment was required, and driver-trained occupational therapists could use the tool to organise information and identify information gaps needed to make a fitness to drive determination (Dickerson & Bédard, 2014).

One human factors psychology model that expanded Michon's (1985) model is the Goals for Driver Education Framework (GDE-framework), which conceptualises driving as a broad set of skills applied according to the driver's motives and goals; higher levels guide and control behaviour on lower levels (Lützenburger & Albayrak, 2013; Peraaho, Keskinen, & Hatakka, 2003). This model was updated to include a fifth level in 2010, and renamed Goals for Driver Education in the Social Perspective (GDE5SOC) (see Figure 2.3). The lower three levels relate to the technical aspects of the driving task (vehicle manoeuvring, mastery of traffic situations, and driving goals and context) (Keskinen, 2014). The fourth level relates to personality and life skills factors affecting driver choices and behaviour (Keskinen, 2014).

The fifth level was added because it was recognised that driver choices are also affected by the social environment (e.g., social norms, legislation and subculture).

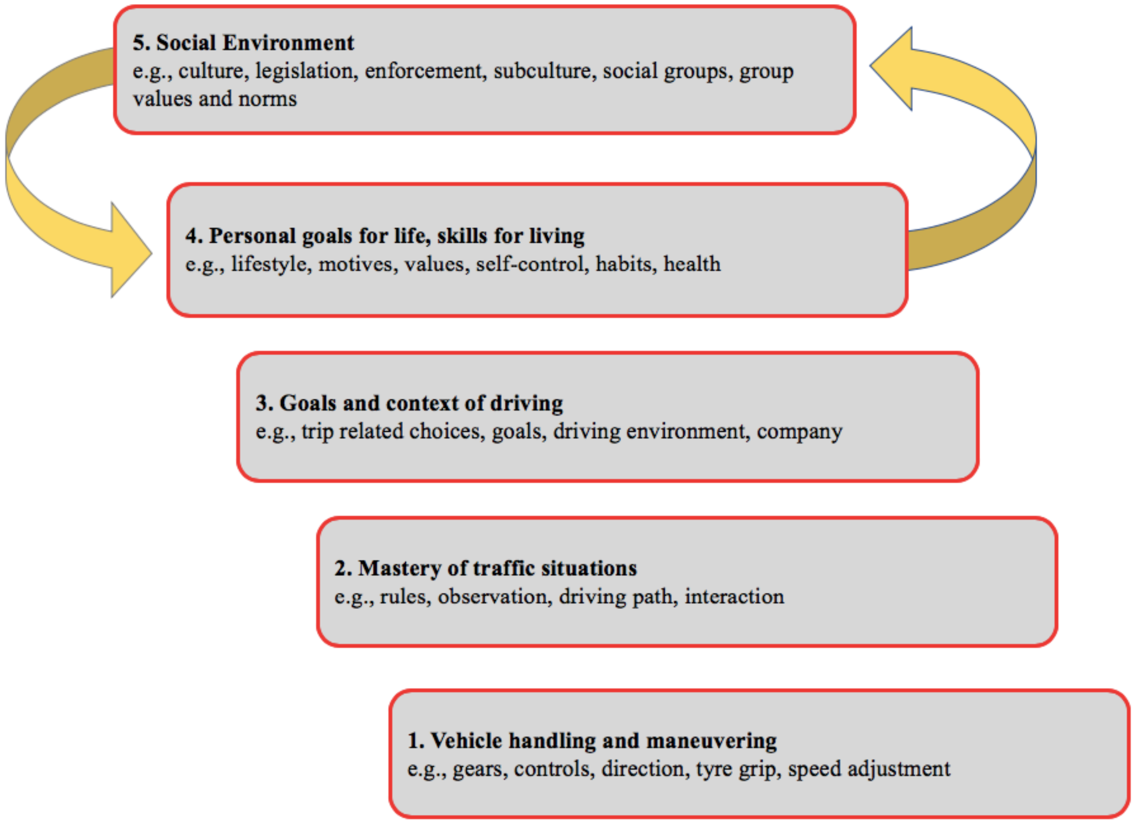


Figure 2.3. Goals for Driver Education in the Social Perspective (GDE5SOC) five-level driving hierarchy (Keskinen, Peräaho, Laapotti, Hernetkoski, & Katila, 2010).

Kay (2008) selected the GDE-framework as the best fit for DSDA because the model was sufficiently complex to represent the driving task, attempted to explain the pattern of crashes common among novice drivers, and incorporated skills, experience, and self-awareness. These factors allow application to both healthy and impaired drivers: required for predicting fitness to drive among patients with functional impairments (Kay, 2008).

Nevertheless, the GDE-framework is difficult to apply to DriveSafe because it was developed to guide novice driver education with a focus on the driver; DriveSafe was developed to assess awareness of the driving environment. Inclusion of self-evaluation fits with the DriveAware conceptualisation of awareness as critical for safe driving. However, DriveAware operationalizes awareness as a lack of discrepancy between the participant's responses and an agreed standard (Kay et al., 2009b). Whereas, the GDE-framework only broadly discusses the concept and does not define or operationalise it so that it can be empirically tested. The framework touches on concepts that may explain driver attention (e.g., reality is stored as constantly changing mental representations that guide attention, perception, and decision making) (Peraaho et al., 2003). Yet, the underlying mechanisms are not specified, limiting validation or use for prediction, and resulting in failure to generate a significant body of research findings (de Winter & Happee, 2012; Michon, 1985; Ranney, 1994; Shinar & Oppenheim, 2011). Still, I found the model too complex for practical application, in agreement with other researchers (Carsten, 2007; de Winter & Happee, 2012) who considered the adding of more and more components to Michon's (1985) hierarchical control model counter-productive and unnecessary for measurement.

Trait models. Trait-based models describe relationships between driver characteristics, focusing on a broad range of stable driver traits such as visual acuity, visual fields, memory, or selective attention (de Winter & Happee, 2012). Traits that are transient or that change over time (e.g., motivations of the driver or environmental factors) are generally not considered (Ranney, 1994). An example of a trait model is the Multifactorial Model of Driver Safety proposed by Anstey et al. (2005) which identifies predictors of on-road performance and accident risk among older drivers (see Figure 2.4). Anstey et al. (2005) identified three "enabling factors" required for safe driving, with additional dimension comprising self-monitoring and belief about driving capacity.

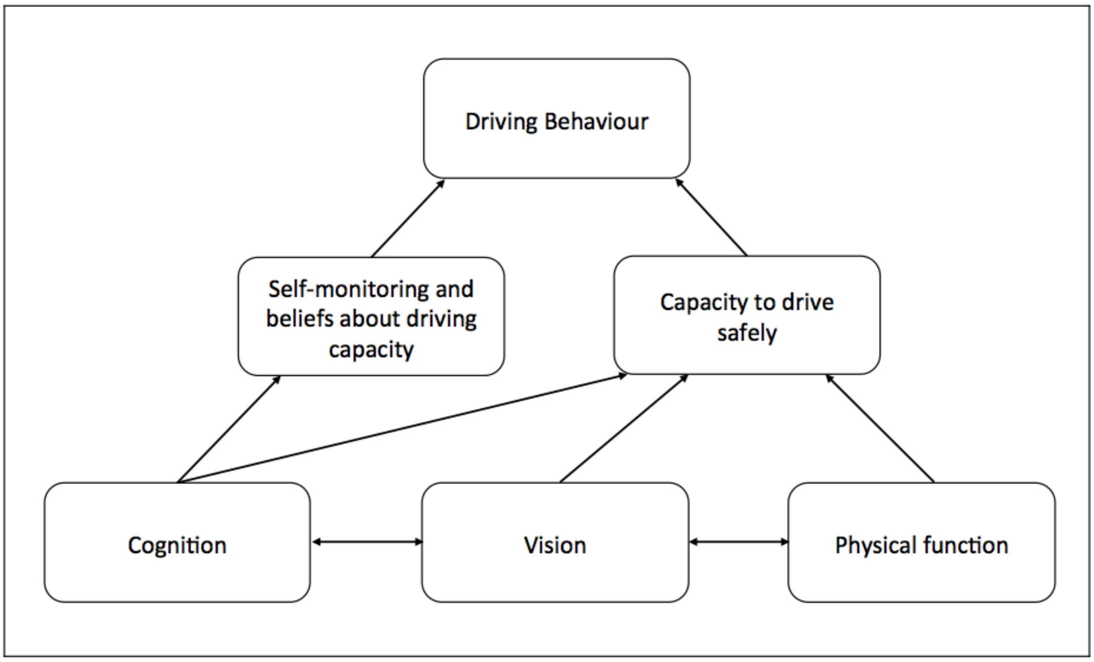


Figure 2.4. Example trait-based model: Multifactorial Model of Driver Safety (Anstey et al., 2005, p.60).

I was hesitant to explore application of trait models to DriveSafe because they are typically simplistic; Rabbit (1981) described them as lineal, independent processes that focus on identifying reliable predictors without consideration of psychological factors. They have failed to generate sufficient evidence they can predict driving performance despite a long history in driving research (Carsten, 2007; de Winter & Happee, 2012; Michon, 1985; Rabbitt, 1981; Ranney, 1994; Shinar & Oppenheim, 2011). Additionally, trait models generally lack well developed theoretical underpinnings and operational definitions of components, resulting in reliance on logic-based, post hoc explanations of relationships between the traits and the criterion measure (Adler & Silverstein, 2008; de Winter & Happee, 2012; Michon, 1985; Ranney, 1994). Trait models often lack clarity regarding what constructs are being measured, causing construct proliferation and redundancy among multiple studies (Heikoop et al., 2015; Ranney, 1994). This is particularly evident for tests of attention where researchers have not defined the specific psychological mechanisms being tapped into (e.g., selective versus divided attention) and have used overlapping constructs and terms (Heikoop et al., 2015; Ranney, 1994). Furthermore, Trait-model research typically relies on statistical correlations to explain empirical connections in studies, which do not necessarily explain causality or meaningful relationships among the components and the criterion measure, and have limited predictive power (Bédard, Weaver, Darzins, & Porter, 2008; Heikoop et al., 2015; Michon, 1985; Ranney, 1994; Shinar & Oppenheim, 2011).

Despite these criticisms, Ranney (1994) noted that selective attention was consistently identified as the strongest predictor of accident involvement in numerous trait-based studies, justifying further investigation. Thus, he suggested the visual search paradigm as a promising new direction for modelling driving behaviour, although still trait-based. The visual search paradigm focuses on a drivers' abilities to identify salient information in constantly changing driving scenes (Ranney, 1994). Thus, it is compatible with the DriveSafe assumption that visual attention is critical to safe driving. Additionally, test protocols most similar to DriveSafe arise from this paradigm: specifically the change blindness model. Therefore, I explored the change blindness, trait-model in detail, to determine if DriveSafe fitted within the model and if it could help me discover DriveSafe mechanism and how they behave to measure of awareness of the driving environment.

The Change Blindness Model – A Good Fit for DriveSafe?

I reviewed literature from the visual search paradigm and change blindness literature via four online databases: Medline via OvidSP (1946 to present), Cinahl via Ebsco (1982 to present); Scopus; and, PsycTESTS. I used the search terms: automobile driving; change blindness; change detection; attention; cognition; visual search; traffic accidents; risk awareness; driving errors; and, looked-but-failed-to-see. I did not use a year or locality limitation. I also reviewed reference lists from relevant articles along with articles citing these publications.

Background - visual attention. Given age-related increased crash risk at intersections and the high number of crashes attributed to missing or delayed hazard perception for all drivers, a significant body of research is dedicated to the study of failure to notice potential conflicts with other objects whilst driving (Baldock et al., 2016; Barco et al., 2015; Caird, Edwards, Creaser, & Horrey, 2005; Fildes, 2008; Preusser, Williams, Ferguson, Ulmer, & Weinstein, 1998; Rakotonirainy, Steinhardt, Delhomme, Darvell, & Schramm, 2012; Wood et al., 2009). A particular subject of study has been the common occurrence of drivers looking in the direction of other road users, particularly bike riders, but failing to perceive them. Accidents of this type are labelled "looked-but-failed-to-see" and typically result in drivers failing to take into account certain hazards (Koustanaï, Boloix, Van Elslande, & Bastien, 2008). Drivers have the information needed for perception but something prevents them from seeing an object in full view. Researchers (Bédard et al., 2006; Caird et al., 2005; Galpin, Underwood, & Crundall, 2009; Hoffman, Yang, Bovaird, & Embretson, 2006; Pringle, Irwin, Kramer, & Atchley, 2001; Rensink, 2005; Rensink, O' Regan, & Clark, 1997; Wood, Horswill, Lacherez, & Anstey, 2013) argued this key factor is attention. In fact, many studies have shown that failures in attention are the main cause of accidents (Eby & Kostyniuk, 2004; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; Koustanaï et al., 2008; Preusser et al., 1998; Ranney, 1994).

Hoffman et al. (2006) defined attention as, "the mechanism by which certain aspects of the environment are selected for further processing while others are inhibited" (p. 985). Attention is critical in most cognitive tasks (e.g., shopping or banking) that require searching for, and prioritising, information (Hoffman et al., 2006). In driving, the driver must monitor the visual scene to rapidly and accurately identify simultaneous information critical for safe decision-making, such as traffic light changes, upcoming road signs, and pedestrians crossing (Hoffman et al., 2006; Rizzo et al., 2009). Since failure to detect these types of changes is a significant cause of accident, numerous authors (Caird et al., 2005; Crundall, 2009; Hoffman et al., 2006; Koustanaï, Van Elslande, & Bastien, 2012) have suggested a measure of where, when, and how drivers allocate attention would be beneficial in road safety research and could perhaps be used to predict driving impairment. Attentional failures have been studied since the early history of psychology via a number of research traditions. The change blindness model emerged from the study of how a unified picture is formed from information obtained from separate glances (transsaccadic integration) (Jensen, Yao, Street, & Simons, 2011).

Model description. Change blindness is defined as the surprising failure to notice large changes in a natural scene, photograph, or film when these coincide with a brief visual disruption (Jensen et al., 2011; O'Regan, Rensink, & Clark, 1999). It can be induced by a variety of situations such as saccades (the rapid eye movements that occur as the eye fixation point changes from one location to another), eye blinks, mud-splashes on a windscreen, blank screens, flickers, and film cuts (Caird et al., 2005; O'Regan et al., 1999; Velichkovsky, Dornhoefer, Kopf, Helmert, & Joos, 2002). Under usual circumstances a movement signal would attract attention to the change location making it easy to detect, but a transient disruption overwhelms this signal causing attention loss or redirection, inducing change-blindness (Galpin et al., 2009; Koustanaï et al., 2012; Rensink, 2000). This disruption effectively removes 'bottom up' activity (e.g., a luminosity change) that automatically attracts attention (Richard, Wright, Ee, & Prime, 2002). The observer is forced to use 'top down' methods to scan various elements of the scene, deciding where to focus attention (Richard et al., 2002). Hence, in driving, there is potential for dangerous events occurring in full view to go unnoticed if they occur simultaneously with even innocuous disturbances such as an eye blink or saccade (Caird et al., 2005; O'Regan et al., 1999). The effect is even more pronounced in novel or complex situations with competing attention demands (Becker & Rasmussen, 2008; Jensen et al., 2011; Koustanaï et al., 2008).

Theoretical background. Rensik (2000, 2002, 2005) provided an in-depth analysis of the theoretical concepts and mechanism involved in the change blindness model. He proposed Coherence Theory to explain the nature of attention based on results of his own change detection experiments and the change blindness literature. His main tenants are: (a) attention and short-term memory significantly overlap and may be part of the same process; (b) objects that are attended to are placed in a coherent complex that falls apart as soon as attention is withdrawn; and, (c) a higher-level structure determines what can be done with the coherent complex (Rensink, 2000, 2002, 2005). The content of attention complexes is sparse in detail and specific to the task at hand (Rensink, 2002; Rensink et al., 1997). At least four properties of four to five attentional complexes can be retained at one time (e.g., colour, size, orientation and presence of a gap). Observers are generally good at screening which properties should be entered into attention complexes and which ones should be left out, but can only focus either globally or locally at the one time (Rensink, 2002).

Coherence theory is represented in Figure 2.5. Early processing occurs rapidly at a low level to form structures (proto-objects) with limited spatial and temporal coherence. These dissolve with new stimuli at their location. Focused attention grasps several proto-objects at a mid-level coherence field forming a more stable individuated object. The feedback that occurs between the two levels is broken when attention is released. The object representation then dissolves back into the proto-objects pool and is overwritten (Rensink, 2005; Rensink et al., 1997)

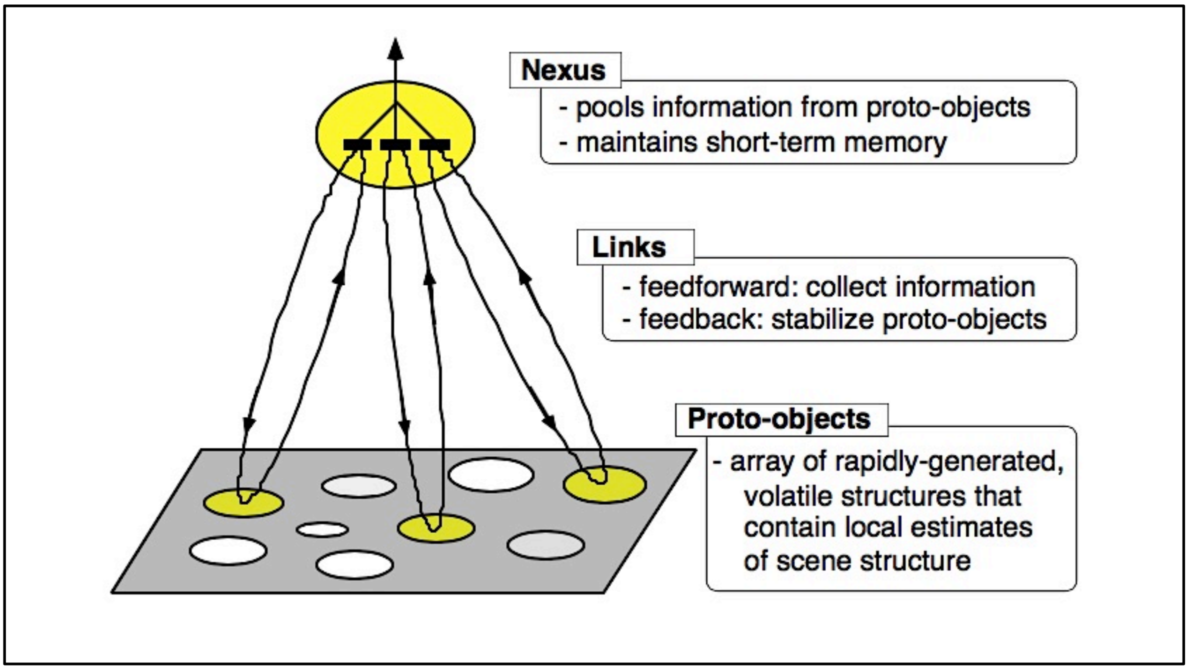


Figure 2.5. Rensink's (2005) pooling of attended information in Coherence Theory (p.78).

There is a large body of evidence to support Rensink's (2000, 2002, 2005) theory, which indicates remarkably little information is stored about objects that are not being directly attended to (Becker & Rasmussen, 2008; Blackmore, Breistaff, Nelson, & Trościanko, 1995; Hayhoe, Bensinger, & Ballard, 1998; Jensen et al., 2011; Koustanaï et al., 2008; O'Regan et al., 1999; Rensink, 2000, 2005; Rensink et al., 1997; Simons & Levin, 1997). Additionally, accurate visual representations only exist as long as people are attending to a particular area or object: meaning that changes occurring in other parts of the scene may go unnoticed because there is no detailed representation of the changed location at that moment (Caird et al., 2005; Pringle et al., 2001; Rensink, 2000, 2002; Rensink et al., 1997).

Rensink (2002) proposed that if formation of coherent structure requires attention, then successful participation in daily activities depends on attention management. The limited amount of information that is stabilized must be used as effectively as possible for the task at hand (Rensink, 2002). This limited capacity impacts speed of information processing, performance accuracy, and reaction time in situations such as driving (Hasher & Zacks, 1998; Hoffman et al., 2006). Adults with age-related changes and medical impairments would be expected to have even greater capacity limitations in these areas (Hoffman et al., 2006), indicating the impact of these factors on performance could be measured via the change blindness model. Thus, within the change blindness model, safe driving would depend on the ability to orient attention in response to external stimuli and internal goals (Hoffman et al., 2006).

Model Mechanisms

I identified key model mechanisms in the change blindness literature that fitted well with the DriveSafe subtest protocol. These must be present to ensure the correct visual perceptual mechanisms are being assessed (Rensink, 2002).

Decoupling motion and change. Motion and change must be separated to decouple the outputs of the change and motion detection systems (Rensink, 2002). Tasks must be designed so that there is no motion cue in the detection of change. Change is defined as an alteration to a well-defined structure over time, maintaining spatiotemporal (space-time) continuity (Rensink, 2002); it is referenced to structure. In contrast, motion is defined as a temporal (time-related) variation in a quality (e.g., luminosity) at a fixed point. Motion is referenced to location (Rensink, 2002). It is important to separate motion and change because motion detection does not need structure and motion detectors do not require attention for their operation (Rensink, 2002). A key aspect of focused visual attention is the ability to create and maintain consistent spatiotemporal representation, requiring more sophisticated processes than for motion detection (Rensink, 2002).

Decoupling detection of change-in-progress and completed-change. The process of seeing a change in progress must be decoupled from the detection of a completed change;

otherwise the detection of the transformation itself is being measured (Rensink, 2002). The detection of completed change involves noticing that a change has occurred at some point in the past, perhaps when vision was briefly occluded, but there is no sense of dynamic change (Rensink, 2002). The viewer must compare a representation in memory with the representation of the structure currently visible. Continuity must be maintained by something other than the visual representation itself (Rensink, 2002).

Decoupling change and difference. Change and difference are separate concepts in change-detection studies even though both rely on similarity, and are referenced to structure without the nature and complexity of the structure being important (Rensink, 2002). Change refers to transformation over time to a single structure. In contrast, difference refers to lack of a similarity in the properties of two structures that are side-by-side, relying on comparisons. There is no sense of the structures being related in time or transforming (Rensink, 2002). Spotting the difference between two structures side-by-side is different to detecting change in two images presented sequentially (Rensink, 2002). The detection of dynamic change involves spatiotemporal continuity in both the external and internal representations, requiring sophisticated memory that maintains continuity and the perception of dynamic transformation (Rensink, 2005). Uncoupling change from difference involves separating the effect of long-term memory from the effects of visual attention (Rensink, 2005).

Methods for assessment of correct perceptual mechanisms. Motion and change outputs to stimuli can be uncoupled by either making the change slow enough so there is no change signal to draw attention (slow fade in or out) or by creating an event that swamps the motion signal at the point of the change, such as a flicker, flash or occlusion (Rensink, 2005). Change and difference can be uncoupled by minimising the impact of memory so that visual attention can be assessed. The impact of memory can be minimized by adding time pressure and by making the viewer respond to a completed change, not dynamic change.

Application to DriveSafe. The review of the change blindness literature enabled me to identify key change blindness model mechanisms critical for the assessment of the attention search ability of drivers, then consider how these mechanisms could underlie DriveSafe. The outcome of this analysis is summarised in Table 2.1.

Table 2.1

Proposed Touchscreen DriveSafe Assessment Mechanisms Based on the Change Blindness Model and Relationship to Driving

DSDA Stimuli	Action	Mechanism Effect	Relationship to Driving
Digital, static, naturalistic driving scenes Objects appropriate to driving context Driving related task	Activates driving-related schema to guide attention Triggers planned and serial search for change	Cause viewer to observe the scene from the perspective of a driver	Systematic scanning of the road environment for potential hazards and change
Countdown screen and bell (Mask)	Cues attention Masks change	Separates motion and change to decouple outputs of the change and motion detection systems (i.e., no motion cue so visual attention can be assessed) Swamps motion signal Forces top down methods to encode scene	Simulates eye blinks, saccades, mudsplashes, and rapid changes on-road, potential causing hazards to be missed
Time limited object display (changed elements)	Minimises the impact of memory	Decouples change and difference so detection of the transformation is not measured Ensures viewer responds to completed change. Viewer must compare representation in memory with representation of visible structures. Separates the effect of long term memory from visual attention so visual attention can be assessed Forces spontaneous and economical search for change	Simulates the amount of time a driver has to scan for information in the driving environment and make turn decisions at intersections
Presentation of stimuli once (one-shot)	Minimises input of memory and eye movements for extended visual search	Allows measurement of accuracy in conditions similar to real world driving	Changes happen only once without the opportunity for extended search for encoding and recall of driving scene elements

I found that motion and change are decoupled in touchscreen DriveSafe by masking object change prior to each new display via a ghosted image of the intersection along with a count down and auditory bell. The mask swamps any motion signal that would otherwise exist if objects changed within a continuously presented scene. I found that change is decoupled from difference by the time pressure. The viewer has 4 seconds to commit the objects to memory before they disappear, ensuring they respond to completed and not dynamic change. The time pressure means the viewer must take a more spontaneous and economical search for change as in real driving (Blackmore et al., 1995; Veirk & Kiesel, 2008).

Change Blindness Test Protocols

All change detection studies have the same basic design despite evolving over many years with varying tasks and stimuli: a viewer is presented with a stimulus (e.g., a picture); an alteration is made to the stimulus (e.g., an element is removed or altered); the response of the viewer is measured (Rensink, 2002). The content of the change is not important except that a radical change in appearance should not be introduced. For example, a car should not be placed in the sky in a real-world driving scene; otherwise performance may be related to the influence of the anomaly rather than the change itself (Rensink, 2005). The simplest change that can be made is item appearance or disappearance. Change can also be made to item properties, such as: colour, brightness, shape, orientation, location and importance (Rensink, 2005). Most studies use static computer displays or images of naturalistic scenes but the methods described below can be broadly applied, including to real-world situations.

One-shot task. In one-shot tasks the image is presented (A), followed by a blank, then the original image is presented again but with a change (A1) (i.e., A, blank, A1) (Jensen et al., 2011). The blank screen between the two images represents a saccade and is used to mask the appearance of the new object (O'Regan et al., 1999). The viewer may only see each image once (Velichkovsky et al., 2002). Performance is usually measured by accuracy in detecting the change, although response time may also be measured (Rensink, 2002). This task minimises the input of long-term memory and eye movements (Rensink, 2002).

Flicker tasks. In the late 1990's, Rensink et al. (1997) adapted the one-shot task into the 'flicker technique' by to enable a richer phenomenological experience of change blindness for the viewer (Jensen et al., 2011). In the flicker task two images are presented in quick succession (e.g., 240ms) with a blank in between (e.g., 80ms) (O'Regan et al., 1999). The viewer sees a continual cycle of the alternating images until the change is noticed or the task times out (i.e., A, blank, A1, blank, A, blank, A1…) (O'Regan et al., 1999; Rensink, 2000). The viewer is usually very surprised to have missed the change for so long once it is identified (Jensen et al., 2011). An example of the flicker task sequence and timing is presented in Figure 2.6. Performance is primarily measured via response time but accuracy may also be measured (Rensink, 2002).

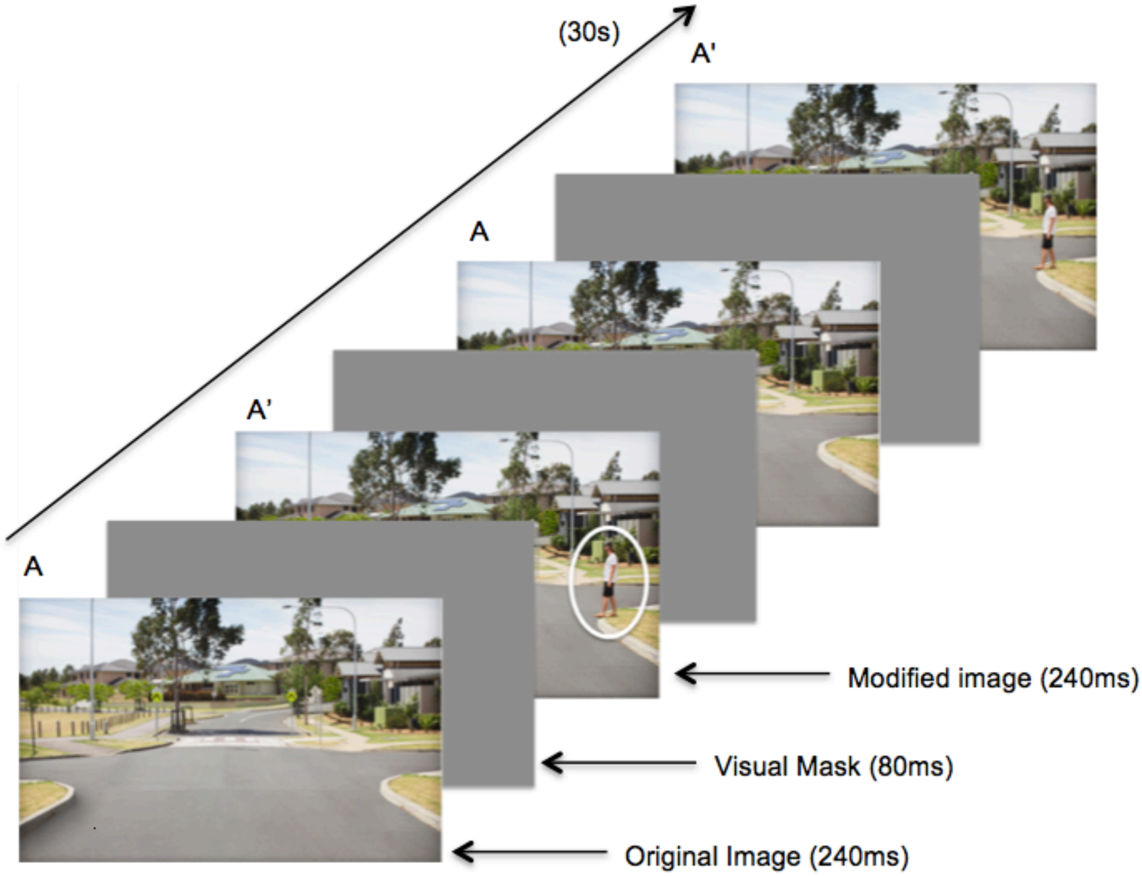


Figure 2.6. Typical flicker task sequence and timing.

Modified flicker method. Caird et al. (2005) modified the flicker method so it could be used to test a driver's attentional capabilities at intersections. This involved adding a time pressure, because drivers only have a few seconds to observe the driving scene, and adding a goal-oriented decision that must take place rapidly. Caird et al. (2005) labelled their technique the modified flicker method (MFM). The authors suggested application of this method as a low-cost way of conducting traffic safety research, such as evaluating driver attention for road infrastructure including traffic lights and signs (Caird et al., 2005).

Mudsplashes. O'Regan et al. (1999) discovered that change blindness could also be induced by scattering a few small, high contrast shapes (similar to mud splashes on a windscreen) over a natural scene instead of a blank (see Figure 2.7 for an example array).

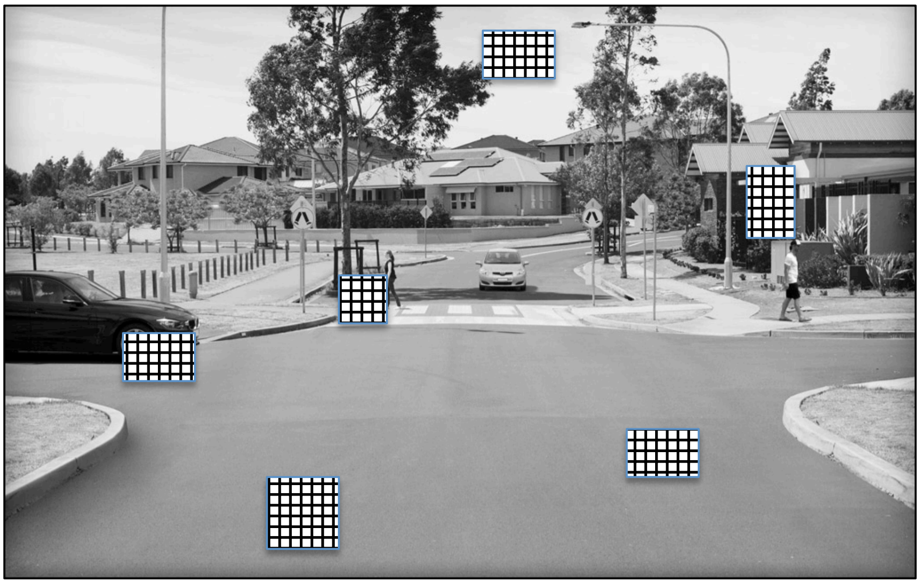


Figure 2.7. Example mudsplashes array: an example changing element would be removal of the rear car.

The small, brief (80ms) disruptions mask even large changes occurring simultaneously, making them very difficult to detect (O'Regan et al., 1999; Velichkovsky et al., 2002).

Theoretical difference of tasks. One-shot tasks and flicker tasks are similar but have important theoretical differences (Jensen et al., 2011). In one-shot tasks, the viewer must encode elements from image A and hold as much information as possible in memory for subsequent comparison with image A1, to increase their chance of finding the change (Jensen et al., 2011). The flicker task allows the viewer to scan small sections of the image to hold in memory for comparison across the blanks until the change is noticed (Jensen et al., 2011). One-shot tasks are more applicable to the real world where change is presented only once. Therefore, this approach is usually applied to test visual memory for scenes. Viewers are generally aware they are looking for a change in both techniques (Jensen et al., 2011). Flicker technique has been criticised for being less naturalistic and causing eye strain and fatigue for the viewer trying to maintain image constancy (Macknik, Fisher, & Bridgeman, 1991). Both one-shot and flicker tasks have been used extensively in driving research to measure how drivers allocate attention and process information (Koustanaï et al., 2012) (see Table 2.2).

Table 2.2

Examples of Driving Related Change Blindness Research

Driving Research Topic	References
Impact of chronological age on change detection	Caird et al., 2005; McCarley et al., 2004; Pringle et al., 2001; Rizzo et al., 2009
Impact of experience change detection	Crundall, 2009; Koustanaï et al., 2012; McCarley et al., 2004; Zhao et al., 2014
Impact of cognitive load and distraction on scanning and change detection	Lee, Lee, & Boyle, 2007; McCarley et al., 2004; Pearson & Schaefer, 2005; Richard, Wright, Ee, & Prime, 2002; Schömig, Metz, & Krüger, 2011; Wallis & Bulthoff, 2000; White & Caird, 2010
Impact of mode of vision occlusion on change detection	Velichkovsky, Dornhoefer, Kopf, Helmert, & Joos, 2002
Impact of training on change detection	Gaspar, Neider, Simons, McArely, & Kramer, 2013
Impact of Alzheimer's disease on change detection	Rizzo et al., 2009
Impact of route familiarity on change detection	Martens & Fox, 2007
Impact of time constraints on decision accuracy at intersections	Caird et al., 2005; White & Caird, 2010
Applicability of the change-detection paradigm to driving	Caird et al., 2005; Galpin et al., 2009; Koustanaï et al., 2012
The study of visual attention mechanisms whilst driving	Bédard et al., 2006; Charlton & Starkey, 2011; Hoffman et al., 2005; Hoffman et al., 2006; Hughes & Cole, 1986; Pringle et al., 2001; Shinoda, Hayhoe, & Shrivastava, 2001
Impact of object features on change detection	Caird et al., 2005; Hoffman et al., 2006; Koustanaï et al., 2012; Pringle et al., 2001

Ecological Validity of the Change Blindness Model

The change blindness model has been criticised for being unrelated to the attentional demands of the real world and inadequate to understand human cognition in complex systems like driving (Smilek, Eastwood, Reynolds, & Kingstone, 2007, 2008). Smilek et al. (2007, 2008) criticised the assumption that breaking down cognitive processes into basic mechanisms and studying them a laboratory would indicate how they operate in everyday life; human behaviour should instead be studied in real life situations for ecological validity (Smilek et al., 2007, 2008).

Change blindness researchers (Beck, Levin, & Angelone, 2007b) countered this criticism with multiple examples of experiments conducted either in real-world settings or with a whole task approach using stimuli with high face validity (i.e., driving simulators, driving videos, and naturalistic driving images). These included change blindness experiments conducted during chess games (Reingold, Charness, Pomplun, & Stampe, 2001), naval combat operations (Divita, Obermayer, & Nugent, 2004), and various other work settings such as hospitals, offices, emergency departments, and traffic control offices (Beck et al., 2007b; Levin, Simons, Angelone, & Chabris, 2002). Change blindness researchers argued their method is valid because real-life driving scenes provide a natural visual environment where certain objects and events needed to be prioritised (Hoffman et al., 2006). For example, driving scenes provide a context for viewers to use a goal-directed and stimulus-driven process to guide eye movements and attention, as in real driving (Hoffman et al., 2006; Hughes & Cole, 1986). Evidence indicates participants clearly view scenes from the perspective of a driver, detecting change to relevant targets much faster regardless of the location and using scanning patterns unique to driving (Caird et al., 2005; Galpin et al., 2009; Koustanaï et al., 2012). Participants consider the change blindness protocol face valid in the few surveys conducted to examine their views regarding relationship to real world driving (Caird et al., 2005; Crundall, 2009).

Whether the vision occlusion methods used in change blindness research are comparable to real-world blinks, saccades, and the sudden object change that occur in driving may be questioned. However, when modes of occlusion are compared in research (e.g., blinks versus artificial blanks and flickers), they have equal effects regardless of the source of the disruption (Fernandez-Duque & Thornton, 2000; Jensen et al., 2011; Rensink, 2002; Velichkovsky et al., 2002). Change blindness has been shown to have strong and generalised effects that are stable across a wide variety of situations (e.g., photographs of real scenes, video, animations, digital images, movie clips, and real-life situations), stimuli (e.g., changed object colour, salience, type, orientation or location) and disruption source (e.g., saccades, blinks, movie cuts and real-world interruptions) (Rensink, 2000, 2002, 2005).

Test Protocols Similar to DriveSafe

I conducted a review of the change blindness literature to identify driving assessments developed from within the change blindness model, because I noticed that most driving related research from this paradigm included experimental protocols similar to original and touchscreen DriveSafe (i.e., digital presentation of photographs of naturalistic driving scenes taken from the perspective of a driver, with participants required to recall information regarding serially presented driving related objects). However, I only identified four studies conducted specifically to develop formal driver assessment tools: DriverScan (Hoffman et al., 2006); Deceleration Detection Flicker Test (Crundall, 2009); Modified Flicker Technique (Caird et al., 2005); and the Hazard Change Detection Test (Wetton et al., 2010). I compare the four test protocols in Table 2.3.

Table 2.3

Comparison of DriverScan, MFM, DDFT, HCDT, and DriveSafe Test Protocols

Test	CB Task	Equipment	Image Pairs	Stimuli	Object changes	Timing	Driving Goal / Measures
Driver-Scan (Hoffman et al., 2006)	FT	PC (43cm) / 24° screen angle. Viewer 76cm from screen.	46	P	- Relevance - Brightness - Location (central vs. peripheral)	**Image:** 280ms **Mask:** 80ms **Time out:** 45s	**Tasks:** Find change as soon as possible (click mouse / report) **Measures:** Response speed / accuracy
DDFT (Crundall, 2009)	FT	PC(25.5cm) / 19.5°angle. Viewer 75cm from screen (chin-rest).	30	P	Trial 1: Vehicles size (front / rear view) Trial 2: Vehicle size (front) and orientation	**Image:** 600ms **Mask:** 200ms **Time out:** M = 35s	**Tasks:** Press keyboard when cars come closer / Identify threatening vs non-threatening hazards (via keyboard) **Measures:** Accuracy / RT
MFM (Caird et al., 2009)	FT	PC (1.5 x 1.3m) Viewer 3m from screen Driver console: brake, accelerator, steering wheel	36	P	1 change per image: - Relevance to turn - Type	**Image:** 250ms **Mask**: 80ms **Time out:** 5s or 8s	**Tasks:** Decided if safe to go in arrow direction (brake / accelerate) / Answer 4 questions about decision **Measures:** Accuracy / Confidence / Information used / Change detection
ACT HCTD (A) / **HCDT** (B) (Anstey et al., 2012; Marrington et al., 2008; Wetton et al., 2010)	FT	A – PC (38cm) / touchscreen Viewer 45cm from screen. B – PC with mouse	A: 64 B: 59	P	1 hazard disappearing in second image	**Image:** A-480ms / B-250ms **Mask:** A-320ms / B- 80ms **Timeout:**32ms	**Task:** Identify the hazard via touching the touchscreen **Measure:** RT
DriveSafe (Kay, Bundy, Clemson, 2009a)	One Shot	9.7inch iPad (iOS 9) Headphones / stylus available iPad angled at 20°.	10	P	- Object type (array of 5) - Number (1-4) - Location - Orientation	**1st:** 4s **2nd:** VI **Mask**: 3s **Time out:** 4s	**Task:** Recall object type, location and direction of movement **Measures:** Recall accuracy

Note: CB=Change Blindness; FT=Flicker Task; P=Photograph of real driving scene from driver perspective; VI= Viewer Initiated; RT=Response Time

DriverScan. DriverScan was developed to measure attentional search ability in older adults, adapted from Rensink et al.'s (1997) flicker task (Hoffman et al., 2006). The authors selected the change blindness model because the viewer must monitor the environment and search for change similarly to driving (Hoffman, McDowd, Atchley, & Dubinsky, 2005). I summarise the test protocol in Table 2.3 above. DriverScan was developed via a pilot study then administered to 155 older drivers for instrument development. Item response theory statistical analysis indicated the items were unidimensional and reliable, that individuals with attentional deficits were the most precisely measured, and that ability to detect change was significantly predicted by the study variables: visual clutter, change brightness, and change relevance (Hoffman et al., 2006).

DriverScan and UFOV ™ were used in a subsequent study to determine if both tests could predict driving simulator performance and self-reported accidents (*N* = 155; aged 63 – 87) (Hoffman et al., 2005). DriveScan could not predict past accidents but could predict simulator driving performance with a sensitivity of 71% at the cost of 35% false positives: performing better than UFOV ™ for false positives (sensitivity 85% at a cost of 48% false positives) (Hoffman et al., 2005). The authors attributed this result to lack of a contextual background in UFOV ™ making the selective attention process more artificial. Neither test could reliably identify at-risk older drivers when compared to self-reported crash data and performance in a driving simulator (Hoffman et al., 2005). A limitation of the study was the validity of these two outcome measures. There is contradictory research regarding validity of driving simulators and their relationship to real-world driving (Dickerson et al., 2014; Kay, 2008) and questionable reliability of self-reported crash data due to over- and under-reporting, and insensitivity of the measure due to the low frequency of the events (Clay et al., 2005; Lew et al., 2005). DriverScan was applied in one further study within a battery of tests used to predict performance in UFOV ™ (Matas, Nettelbeck, & Burns, 2014). UFOV ™ has three subtests: 1) processing speed, 2) divided attention, and 3) selective attention (Ball & Owsley, 1993). The authors concluded that DriveScan was a significant predictor of performance in subtest 1 and 2, and that performance in subtest 2 (divided attention) is best explained by change detection (DriveScan) and processing speed (Matas et al., 2014).

Deceleration detection flicker test (DDFT). Crundall (2009) developed the DDFT as a simple, low-cost method for exploring driving-related research questions with some real-world validity. This was in response to the risk and ethical difficulties associated with on-road driving research, the cost of questionable validity of driving simulators, and the cost and time associated with developing driving videos (Crundall, 2009). I summarised the DDFT test protocol in Table 2.3 (the comparison of test protocols). DDFT was developed over two experiments conducted with experienced and inexperienced drivers including a primary and secondary driving task (e.g., detection of car in front decelerating) (Crundall, 2009). Results indicated a significant main effect of driver experience in accuracy of change detection (93.6% compared to 86.9%). Experienced drivers performed better on the secondary task (87% compared to 78%) (Crundall, 2009). Participants in the study reported they felt the DDFT assessed driving-relevant skill (Crundall, 2009). Crundall (2009) advised further research to determine if the test could become a diagnostic tool.

Modified Flicker Method (MFM). Caird et al. (2005) developed the MFM to test the attentional ability of drivers at intersections. The authors modified the flicker task to make it more related to real-world driving (Caird et al., 2005). Observers were given a driving-related goal (i.e., deciding if safe to proceed at an intersection) to allow driving experience to impact performance and guide attention. A time pressure was added because drivers often have only a few seconds to observe the driving scene before making a decision. It was assumed this would negatively impact decision accuracy (Caird et al., 2005). I summarised the test protocol in the Table 2.3 above.

Caird et al. (2005) conducted a study with MFM to assess the effect of age and time on intersection turn decision accuracy. Participants viewed images for either 5 or 8 seconds, approximating the time required for intersection approach. Participants ($N = 62$) comprised four age groups (i.e., 18-25; 26-64; 65-73; 74+). Participants responded to four questions after taking each item to measure confidence (e.g., "How confident are you in your decision to go or not to go?") (Caird et al., 2005). Findings indicated object contrast and size did not predict decision accuracy for any age group even though visual acuity and contrast sensitivity were worse for the two older groups (65+). The youngest group had greater turn decision accuracy. Older drivers had especially low accuracy for detecting pedestrians. Road sign change detection also declined with age. Older drivers relied heavily on traffic control devices to make decisions, often to the exclusion of other objects such as pedestrians: suggesting older drivers adopt coping strategies to identify the most important objects but may miss unexpected hazards (Caird et al., 2005). Fourteen test items significantly predicted accuracy (Caird et al., 2005).

Caird et al. (2005) concluded that MFM was useful for assessing attention to fixed objects (e.g., traffic lights and signs) and for testing drivers' visual search and decision making under time pressure, allowing measurement of the impact of experience (Caird et al., 2005). The authors proposed the test could be applied to assess working memory and study the looked-but-failed-to-see phenomenon, which can be hard to measure in real traffic (Caird et al., 2005). They advised further research to establish test validity and reliability (Caird et al., 2005).

Australian Capital Territory (ACT) Hazard Change Detection Task (HCDT). Wetton et al. (2010) developed the ACT HCDT as one of two hazard perception measures created to separate the processes required in hazard detection. The ACT HCDT was based on a change detection test developed by Marrington et al. (2008) to evaluate the impact of simulated cataracts on hazard perception. The HCDT was the outcome measure in the 2008 study. As could be anticipated, participants wearing cataract simulating goggles had slower reaction times in the HCDT (Marrington, Horswill, & Wood, 2008). Anstey et al. (2012) subsequently used the Marrington et al. (2008) version of the HCDT as an outcome measure in a study conducted to examine the Multifactorial Model of Driver Safety (described earlier in the thesis). The ACT HCDT test protocol appears in Table 2.3.

Both the Wetton et al. (2010) and Anstey et al. (2012) examined correlations of HCDT scores with scores on other tests (e.g., UFOV ™). Wetton et al. (2010) concluded that the ACT HCDT was not suitable for testing novice drivers because they performed faster than experienced drivers. The test was also not correlated with other measures used in the study. Anstey et al. (2012) found evidence that HCDT was correlated with age, gender (males being faster) and tests of spatial and executive speed. Neither version of the HCDT has been compared with criterion measures such as accident rates, performance in a driving simulator, or performance on-road. Predictive validity has not been examined.

Comparison of test protocols. Similarities between the four test protocols and original and touchscreen DriveSafe include: (a) presentation of photographs of real driving scenes taken from the driver's perspective and modified via editing software to introduce changed elements relevant to the driving task; (b) display of images via digital equipment (computer, data projector, or iPad); and (c) measurement of performance via change detection accuracy and / or response times: with DriveSafe and MFM also adding questions for participants similar to those in DriveAware. All protocols have a driving related task to complete (i.e., scan the driving scene for objects or hazards and recall changed elements) but MFM and DDFT add a separate driving related goal to increase applicability to driving.

A significant difference between original and touchscreen DriveSafe and the other four protocols is timing of object display. MFM and DriveSafe have the most similar object presentation time: 4s for DriveSafe; 5s or 8s for MFM. Caird et al. (2005) found no difference in performance whether the objects were displayed for 5s or 8s. This timeframe appears to reasonably represent the amount of time a driver has to scan the driving environment at intersections and make a turn decision (Caird et al., 2005). The mask is presented for much longer in touchscreen DriveSafe (4s) compared to most change blindness studies (around 80ms). However, findings from change blindness research indicate the length of the mask is not important, as long as it is at least the length of a blink or saccade (i.e., 80ms) (Summerfield, Lepsien, Gitelman, Mesulam, & Nobre, 2006). In fact, viewers are able to retain memory of scenes and detect change after 1 to 2 days, as long as there has been sufficient time to encode the original image (Summerfield et al., 2006).

Original and touchscreen DriveSafe apply a unique change detection task. The other protocols implement a flicker task. Both versions of the DriveSafe subtest present stimuli only once for each item: with the blank first (i.e., blank, A1, A). The mask precedes the change to swamp any motion signal so that only the completed change is seen (see Figure 2.8).

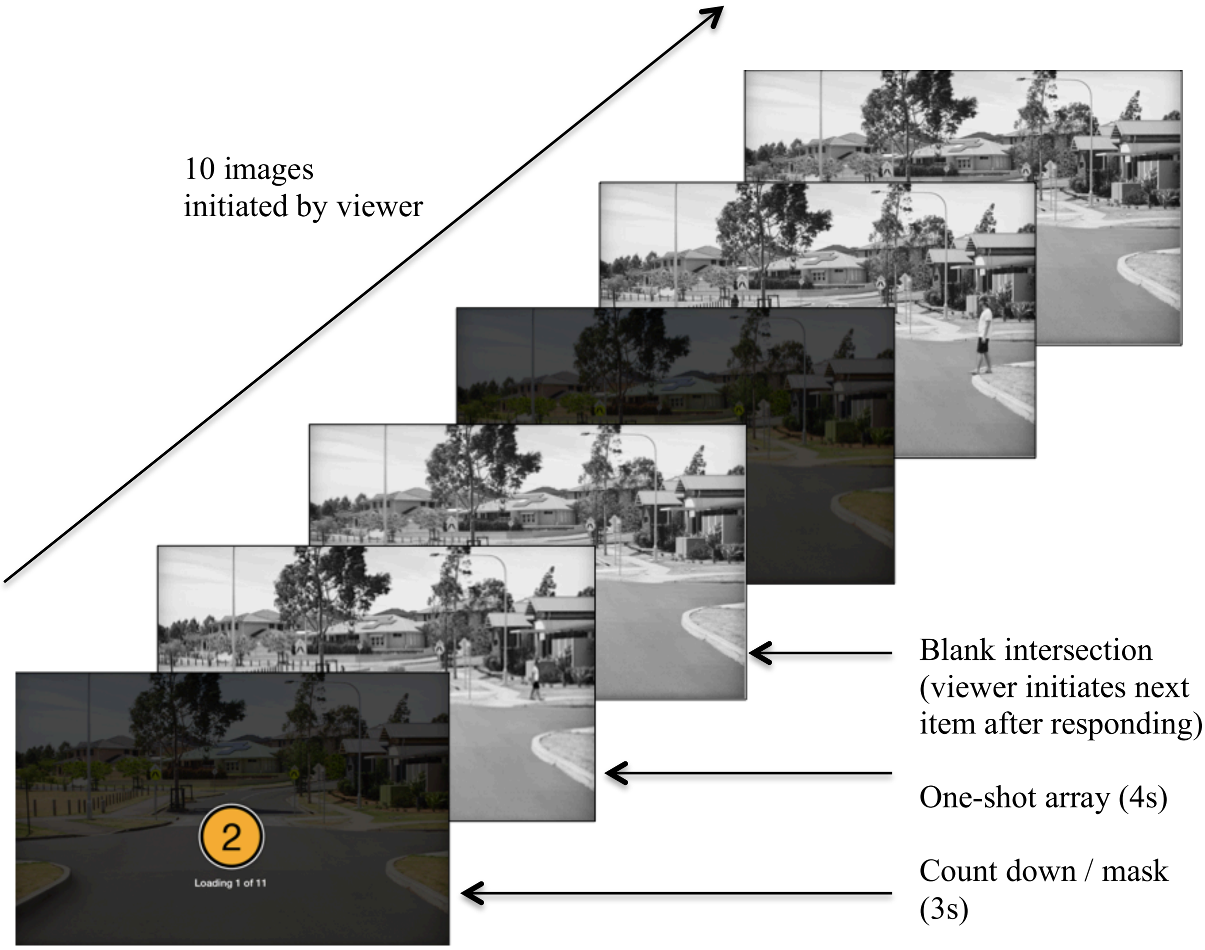


Figure 2.8. Image sequence and timing for touchscreen DriveSafe.

The mask consists of a heavily ghosted image of the intersection with a countdown and bell in touchscreen DriveSafe, rather than the blank screen used in the other test. The touchscreen DriveSafe protocol has similarities to Rizzo et al.'s (2009) study of change detection for older drivers with Alzheimer's disease, where driving-related objects (e.g., cars and pedestrians) slowly faded in and out of the scene without a change to the base image in the item.

Driver-Scan, MFM, DDFT and HCDT are too lengthy to be used as GP fitness-to-drive screens due to the number of items presented (between 30 and 64). However, the tests could be shortened with further research to identify and remove redundant or poorly performing items. Driver-Scan presents images via a data projector and large screen and seats the viewer at a mock vehicle console: a set-up not practical for GPs. The other tests are presented via a personal computer, a set-up that could be managed in a GP office. The main barrier to use of the four tests as GP fitness-to-drive screens is that they are at the beginning stage of development. MFM, DDFT, and Driver-Scan have not been studied further despite being developed around a decade ago. The psychometric properties and predictive validity of data gathered with the tests have not been sufficiently investigated using suitable criterion measures. However, study findings support change blindness as a face valid model for driver assessment that can discriminate between individuals and measure the impact of age and experience (see Appendix B for further information regarding how the model measures these impacts).

Does DriveSafe Fit the Change Blindness Model?

Original and touchscreen DriveSafe meet Rensink's (2002) criteria for change blindness protocol because the viewer is presented with a stimulus (i.e., image of a natural driving scene), an alteration is made to the stimulus (i.e., object appearance and disappearance) and the response of the viewer is measured (via accuracy of ability to recall object type, location and orientation). All changes are appropriate to a driving scene. Additionally, DriveSafe utilises mechanisms identified by change-blindness model researchers (Caird, Edwards, Creaser, & Horrey, 2005; Crundall, 2009; Galpin, Underwood, & Crundall, 2009; Jensen, Yao, Street, & Simons, 2011; Koustanaï, Van Elslande, & Bastien, 2012; O'Regan, Rensink, & Clark, 1999; Rensink, 2002; Rensink, O' Regan, & Clark, 1997) as necessary for assessment of driver visual attention: use of naturalistic driving scenes and stimuli, a driving related task, masking of object change, and a time-limited, one-time display of objects.

One-shot task is the change blindness test protocol most similar to DriveSafe because the change is only made once then the viewer must respond. Consistent with one-shot protocols, accuracy is the most important aspect being measured; response time is also recorded but is not included in scoring. Some researchers (Jensen et al., 2011) have proposed that a one-shot task is better for studying what observers do under naturalistic conditions. It is also advantageous for driving research because the input of memory and eye movements for extended search is reduced (Blackmore et al., 1995; Veirk & Kiesel, 2008). In real driving, the viewer does not have up to 45s to scan each small section of the scene and compare repeated change over time as in the flicker task. The flicker task can still be used successfully but is sometimes modified to incorporate a driving related task to increase application to driving (Caird et al., 2005; Crundall, 2009; Wallis & Bulthoff, 2000).

Original and touchscreen DriveSafe are similar to almost all other driving-related change blindness test protocols in that a static, naturalistic driving scene taken from the perspective of a driver, appears on a screen. Only a few driving related protocols use alternative forms of presentation including video (Wallis & Bulthoff, 2000) and driving simulators (Charlton & Starkey, 2011; Martens & Fox, 2007; Shinoda, Hayhoe, & Shrivastava, 2001). Touchscreen DriveSafe is unique in the use of touchscreen technology and the presentation of objects immediately after the mask, followed by an unchanged image. However, this method is more suited to newer, touchscreen technologies where attention must first be cued to a small screen before the change can be made. The countdown and bell in touchscreen Drive Safe shift the viewer's attention to the centre of the screen so they are optimally primed for the change, otherwise there would be a risk that the change would occur without the viewer looking at the screen. A cueing screen was also used in MFM (Caird et al., 2005).

Conclusion. DSDA is a unique assessment among published psychometric tests of cognitive fitness to drive. However, I found that the DriveSafe subtest protocol is not new and fits well within the change blindness model. According to the model safe driving would depend on the attentional management of limited stabilised information, impacting speed of information processing, performance accuracy, and reaction time (Hasher & Zacks, 1998; Hoffman et al., 2006; Rensink, 2002). The impact of fitness-to-drive factors, such as age related changes, cognitive impairment, reduced vision, and the presence of medical conditions, could further limit capacity, allowing a test such as DriveSafe to discriminate among individuals and predict safe driving. In the context of the change blindness model, DriveSafe could be described as a fitness-to-drive screen that measures individual differences in attentional search ability of drivers.

I was able to apply the change blindness model to consider how the key change blindness model mechanisms could underlie DriveSafe and operate to assess the attentional search ability of drivers. This step enabled me to retain potentially critical test mechanisms in the touchscreen DriveSafe conversion. Trait models have been criticised for relying on correlations without underlying theoretical concepts, clearly defined constructs, and operationalization of mechanisms: resulting in lack of success in identifying predictors of safe driving (Adler & Silverstein, 2008; de Winter & Happee, 2012; Heikoop et al., 2015; Michon, 1985; Ranney, 1994). However, I propose that these criticisms can be addressed by: (a) applying the change blindness model to describe underlying theoretical concepts and mechanisms; (b) a study design that includes a suitable outcome measure (e.g., a standardised on-road assessment); and, (c) statistical methods that allow construct validity and internal reliability to be examined, cut-off scores to be set, and sensitivity, specificity, PPV and NPV to be calculated. I found the change blindness model fit the criteria of a simple, explicit, usable, validated, and predictive model, applicable to driving related research.

CHAPTER 3

Critical Review of Existing Fitness-to-Drive Screening Tests

The conversion of a standardised assessment to digital administration requires a significant investment of time and financial resources. Therefore, I conducted a needs analysis prior to project commencement to justify the research. I reviewed the literature to: (a) identify existing screening tests used for determining fitness to drive; (b) determine if these tests had sufficient accuracy to predict on-road performance; and, (c) evaluate if these tests were feasible for medical practice. Driver screening refers to brief, simple tests that identify drivers who are either clearly without deficits that will impact driving or those who require further in-depth assessment to determine fitness to drive (Bédard & Dickerson, 2014). The aim of the literature review was to determine if GPs already had access to suitable fitness-to-drive screens. First, I examined literature to identify optimum fitness-to-drive test design criteria recommended by researchers, so existing tests could be compared against these criteria. Next, I reviewed literature to identify existing fitness-to-drive tests that were potentially practical and valid for medical practice. I searched five online databases: Medline via OvidSP (1946 to present), Cinahl via Ebsco (1982 to present); Ageline via Ebsco (1978 to present); Scopus; and, PsycTESTS. Search terms included: automobile driving; automobile driver examination; safety; driver evaluation; risk assessment; cognition; psychometrics; neuropsychological tests; general practitioner; older driver; and, fitness-to-drive. I reviewed reference lists of relevant articles and articles citing these publications. The literature search did not have a time or locality limitation.

GP Fitness-to-drive Test Design Criteria Recommended in the Literature

Bédard et al. (2008) described the development a suitable fitness-to-drive screen for GPs as a quest:

> A quick and valid screening tool has been treated by many researchers, including ourselves, as the "holy grail". Such a tool, which could be administered quickly at a licensing bureau or by a health professional, would have considerable value. (p. 336)

A vast amount of research has been dedicated to this endeavour. Whilst it is not possible to predict driving performance with 100% accuracy, researchers have set a number of criteria for a useful driver screen. Researchers (Bédard & Dickerson, 2014; Bédard et al., 2008; Molnar et al., 2006; Weaver & Bédard, 2012) recommend evidence-based cut-off scores where sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) are reported. Tests that rely on a statistically significant relationship between the results and poor driving performance do not have sufficient accuracy to predict driving performance due to a significant overlap in safe and unsafe scores: a limiting characteristic of correlation research (Bédard et al., 2008; Heikoop et al., 2015; Michon, 1985; Ranney, 1994). Having two cut-off scores can address this problem, allowing categorisation of patients into pass, fail, and intermediate categories (Langford, 2008; Laycock, 2011; Molnar et al., 2006). Bédard et al. (2008) reported that a driver screen is only clinically useful if it classifies a large proportion of drivers as either safe or unsafe, with only a small percentage (10 – 20%) falling into the further testing category. Asimakopulos et al. (2012) described an ideal fitness to drive screen as one that is simple to administer, takes 5-10 minutes, and has sensitivity, specificity, PPV and NPV values over 80%. Additionally, researchers (Classen, Velozo, Winter, Bédard, & Wang, 2015; Classen et al., 2010; Hargrave, Nupp, & Erickson, 2012; Kay et al., 2012; Langford et al., 2008; Laycock, 2011; Wheatley & Di Stefano, 2008; Yale et al., 2003) recommend that assessment tools be compared against a suitable and valid criterion measure, preferably a standardised occupational therapy on-road assessment.

Dalchow, Niewoehner, Henderson, and Carr (2010) highlighted the importance of driver screen face validity. GPs require a test that patients will accept as related to real-world driving, so patients feels they have had a fair assessment of driving ability and will be more likely to accept the results, particularly if they fail (Dalchow, Niewoehner, Henderson, & Carr, 2010). Driving tests with the highest rate of acceptance amongst patients are: on-road assessment, visual acuity testing, hazard perception tests, road rules tests, and components of neuropsychological tests with driving-related scenes and activities (e.g., map reading) (Crundall, 2009; Dalchow et al., 2010). Tests traditionally used for driver screening such as Trail Making Test (TMT) Part A and B, Clock Drawing Test, Rapid Pace Walk, and maze tests, have a low rate of acceptance (Dalchow et al., 2010). Driver assessment is stressful and has potentially significant quality-of-life implications. Therefore, patients need to be able to understand why a particular test has been adopted and perceive that it is related to their driving ability (Dalchow et al., 2010).

A significant concern for the high stakes task of driver assessment is that tests be sufficiently predictive to avoid misclassifications, particularly minimising identification of patients as safe when they are not, due to the risk of injury and death for the driver and others (Bédard et al., 2008). Although less critical for community safety, misclassification of safe drivers as unsafe also has potentially significant negative health outcomes. There is much evidence that driving cessation is associated with declining cognitive, physical and social functioning, increased depression, higher rates of admission to long-term care facilities, and increased mortality (Chihuri et al., 2016; Edwards, Lunsman, Perkins, Rebok, & Roth, 2009).

Table 3.1 summarises the GP fitness to drive test design criteria recommended in the literature.

Table 3.1

GP Fitness-to-Drive Test Design Criteria Recommended in the Literature

Criteria for a GP driver screen	Criteria references
Evidence based cut-off scores	Bédard et al., 2008; Molnar et al., 2006
2 cut-off scores (generating a "further testing" category)	Bédard et al., 2008; Langford et al., 2008; Molnar et al., 2006; Laycock, 2011
Sensitivity, specificity, PPV and NPV reported and above 80%	Asimakopulos et al., 2012; Bédard et al., 2008; Langford et al., 2008; Molnar et al., 2006; Weaver & Bédard, 2012
Criterion measure: a standardised on-road assessment	Classen et al., 2015; Classen et al., 2010; Hargrave et al., 2012; Laycock, 2011; Wheatley & Di Stefano, 2008; Yale et al., 2003
Small percentage (e.g., 10-20%) of patients classified "further testing"	Bédard et al., 2008; Weaver & Bédard, 2012
Face valid for the driving task	Crundall, 2009; Dalchow et al., 2010; Weaver & Bédard, 2012
Brief (i.e., 10 minutes or less)	Asimakopulos et al., 2012
User friendly for GPs (i.e., portable, no unique testing consoles, no training, and simple)	Asimakopulos et al., 2012; Dalchow et al., 2010; Fildes, 2008; Molnar et al., 2006

This list provided us with a useful template for evaluating existing fitness-to-drive tests and also informed both the touchscreen DSDA app design and the design of research projects conducted in later phases of the touchscreen DSDA conversion.

Review of currently available driver screens. Many neuropsychological tests commonly used in driving clinics, such as Trail Making Test (TMT) Parts A & B, Motor-Free Visual Perception Test (MVPT), and maze tests, have sensitivity and specificity either not calculated or well below an acceptable level for accurately predicting driving performance (Anstey, Wood, Lord, & Walker, 2005; Bédard et al., 2008; Classen, Wang, Crizzle, Winter, & Lanford, 2013; Kay et al., 2012). Therefore, these tests are not suitable to use alone to screen fitness to drive. Numerous authors have presented self-assessment or family and carer driver assessments used for determining if older drivers are safe to drive (Classen et al., 2016; Classen, Wen, Velozo, Bedard, et al., 2012; Classen, Winter, Velozo, Hannold, & Rogers, 2013; Classen et al., 2010; Levasseur et al., 2014; Medhizadah, Classen, & Johnson, 2018). Few of these are rigorously developed and evaluated (Levasseur et al., 2014) or have sensitivity and specificity reported. Additionally, older driver self-assessment has low accuracy in predicting on-road assessment outcomes (Marottoli & Richardson, 1998).

However, one promising, reliable and accurate driver screening tool is the Fitness-to-Drive Screening Measure (FTDS), a 54-item online screening tool for a proxy rater (e.g., family and carers) to use in determining if older drivers are having difficulty with driving (Classen et al., 2016; Classen et al., 2015; Classen, Wen, Velozo, Bédard, et al., 2012; Classen, Wen, Velozo, Bedard, et al., 2012; Classen, Winter, et al., 2013; Classen et al., 2010). The proxy rates the level of challenge 54 driving skills present to the drive (e.g., driving in the correct lane) on a 4-point scale from very difficult to not difficult. Results place older drivers into one of three categories: accomplished driver, routine driver, and at-risk driver. Two studies (Classen et al., 2015; Classen, Wang, Crizzle, et al., 2013), have examined whether FTDS can predict on-road assessment outcomes. Classen et al. (2015) reported the specificity and sensitivity of 33 FTDS items at two cut points as 98.2% and 19.4% respectively, with 28/200 misclassifications; and, 60.4% and 80.6%, with 73/200 misclassifications. The authors noted a ceiling effect, with items assessing the performance of only the least competent drivers (Classen et al., 2015). Classen et al. (2013) reported good sensitivity (79%) in a study involving 168 community dwelling older drivers and their family/carers but specificity was 59%, with a large number of false positives (Classen, Wang, Crizzle, et al., 2013). FTDS is a free, easily-accessible driver screening tool with favourable face valid and content validity (Classen, Winter, et al., 2013; Classen et al., 2010). It is a helpful tool for starting a conversation regarding cessation of driving and guiding older drivers to resources for further assessment where required (Classen et al., 2015; Classen, Wang, Velozo, et al., 2013; Classen, Wen, Velozo, Bédard, et al., 2012; Classen, Wen, Velozo, Bedard, et al., 2012). However, the assessment takes 20 minutes (Classen et al., 2014), is designed to be administered by a family/carer proxy familiar with the older drivers' driving performance (Classen et al., 2015; Classen, Wen, Velozo, Bédard, et al., 2012; Classen, Winter, et al., 2013), and classifies only 6% of participants in the "at-risk" category (Classen et al., 2014). Therefore, the test is neither practical nor suitable for medical practice, or sufficiently predictive to screen fitness to drive.

In a structured review to identify accurate and practical fitness-to-drive screens for clinicians, Asimakopulos et al. (2012) identified driver screens that are accurate enough to predict driving performance: with sensitivity, specificity, PPV and NPV over 80%. However, among the 53 tools reviewed, only four had sufficient accuracy to predict driving performance: Useful Field of View (UFOV ™), original DSDA, Sensory Motor Cognitive

Tests (SMCTests), and Stroke Driver Screening Assessment (SDSA). However, none of these tools met all of the criteria (Asimakopulos et al., 2012). Asimokopulos et al. (2012) identified SDSA (Lincoln, Radford, & Nouri, 2012) as a promising test. SDSA is a battery of cognitive tests developed in the UK to determine fitness to drive for stroke patients (Lincoln et al., 2012). SDSA takes approximately 30 minutes to administer (Selander, Johansson, Lundberg, & Falkmer, 2010) and has been adapted for Sweden and Norway (Selander et al., 2010). The outcome of an on-road assessment was the criterion measure used in the research (Asimakopulos et al., 2012; Lincoln et al., 2012; Selander et al., 2010). Initial research indicated SDSA was accurate enough to predict driving performance, correctly classifying approximately 80% of stroke patients (Asimakopulos et al., 2012; Lincoln & Fanthome, 1994; Lincoln et al., 2012). However, a validation study conducted in Sweden and Norway indicated fewer than 70% of patients were correctly classified (Lundberg, Caneman, Samuelsson, Hakamies-Blomqvist, & Almqvist, 2003). A further study found only 50% of patient with cognitive impairment and 62% of stroke patients were correctly classified (Selander et al., 2010). Selander et al., (2010) concluded that SDSA cannot predict the outcome of an on-road assessment and should not be used as a stand-alone test (Selander et al., 2010).

Our group conducted a similar review of clinical fitness-to-drive tests (Kay et al., 2012). In agreement with Asimokopulos et al. (2012), we identified SMCTests (Innes et al., 2007) as potentially sufficiently accurate to predict fitness to drive. SMCTests is a computerised group of visuomotor and visuoperceptual tests administered via an integrated car body apparatus (Innes et al., 2007). Result of an on-road assessment was the criterion measure used in the research (Hoggarth, Innes, Dalrymple-Alford, & Jones, 2013; Innes et al., 2007; Innes, Jones, Anderson, Hollobon, & Dalrymple-Alford, 2009). Early research conducted by Innes et al., (2007) indicated high sensitivity and specificity: 97% and 89%

respectively. However, results of a subsequent study with cognitively impaired older drivers (*N* = 279), evidenced lower sensitivity and specificity (73.5% and 70.2% respectively), indicating SMCTests are not sufficiently accurate to predict driving performance (Hoggarth et al., 2013). SMCTests have high face validity (Innes et al., 2009) but they are not practical for GPs because specialised testing equipment is required and the administration time is lengthy: reported as several hours in one study (Heitger et al., 2004).

UFOV ™ (Ball, Beard, Roenker, Miller, & Griggs, 1988; Ball & Owsley, 1993), a 15-minute computerised test of functional vision and visual attention, is a driver screen that has been widely used and extensively studied (Bédard et al., 2008; Classen, Wang, Crizzle, et al., 2013; Dickerson et al., 2014; Kay et al., 2012). Although criterion measures vary between studies, UFOV ™ has been compared with results of an on-road assessment in some research (Austroads, 2004; Bédard et al., 2008; Myers, Ball, Kalina, Roth, & Goode, 2000). UFOV ™ has the potential to be practical for medical practice because it is easy to administer via computer and has a short administration time compared to other fitness-to-drive tests. However, UFOV ™ demonstrates variable sensitivity and specificity estimates between studies and these are not consistently high enough for the test to be used to predict driving performance accurately (Bédard et al., 2008; Classen, Wang, Crizzle, et al., 2013; Dickerson et al., 2014; Kay et al., 2012).

DriveAble (Dobbs, 2005) is another computerised test widely used to determine fitness to drive, primarily in Canada. However, evidence from driver screen review studies indicates the test is not a good predictor of on-road performance, with sensitivity 76%; specificity 90%; PPV 97%; and, NPV 47% (Kay et al., 2012; Korner-Bitensky & Sofer, 2009). Additionally, the author has not published the DriveAble algorithms. A recent large (*N* = 3662) study conducted by the test author indicated a very low error rate for pass and fail predictions for patients categorised as requiring further testing: 1.7% for pass predictions and 5.6% for fail predictions (Dobbs, 2013). Whilst these results are promising, DriveAble is not suitable for medical practice due to the need for expensive administration equipment (i.e., a unique testing console); ongoing subscription for interpretation of results; lengthy test administration time (35 to 50 minutes according to Korner-Bitensky & Sofer, 2009); and because a large percentage of participants are classified as requiring further testing (36.5% to 46%) (Dobbs, 2013; Korner-Bitensky & Sofer, 2009).

I briefly mention the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) and Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) here because these are the cognitive screening tools most frequently used by GPs (Dobbs et al., 1998; Hoggarth, 2013; Wilson & Kirby, 2008): adopted because they are quick, inexpensive, and easy to administer (Dobbs et al., 1998; Hoggarth, 2013; Wilson & Kirby, 2008). The MMSE was developed to estimate the severity of cognitive impairment (Folstein et al., 1975). The MoCa was developed to detect mild cognitive impairment because the authors noted that individuals with mild impairment consistently scored within the normal range on MMSE (Nasreddine et al., 2005). Both tests are scored on a 30-point scale, with a higher score representing better function. MoCA includes tasks that assess a broader range of cognitive domains and is more sensitive to mild cognitive impairment than MMSE (Bowers et al., 2013; Nasreddine et al., 2005; Trzepacz et al., 2015). The validity of MMSE in predicting on-road assessment outcome has been widely researched but many studies only examine statistically significant relationships between MMSE and accident rate or on-road performance, rather than calculating sensitivity and specificity: many demonstrating no correlation (Anstey et al., 2005; Asimakopulos et al., 2012; Bédard et al., 2008; De Raedt & Ponjaert-Kristoffersen, 2001; Dickerson et al., 2014; Hollis, Duncanson, Kapust, Xi, & O'Connor, 2015; Vrkljan et al., 2011).

However, a number of studies have compared the accuracy of MMSE and MoCA in predicting on-road assessment outcome, calculating sensitivity and specificity. Bowers et al. (2013) examined a range of cognitive measures, including MMSE and MoCA, to determine if combining tests would better predictive driving performance. The authors concluded MMSE and MoCA were equivalent predictors (specificity: 0.76 and 0.68; sensitivity 0.67 and 0.80 respectively) but neither test was sufficiently accurate alone (Bowers et al., 2013). The best combination of tests for predicting on-road assessment outcome was UFOV ™ subtest 2, MMSE, visual acuity and contrast sensitivity measures, and removal of Trails A and B from the battery (specificity 0.95; sensitivity 0.80) (Bowers et al., 2013). Whilst this combination was very accurate in identifying safe drivers, it failed to identify 20% of at risk drivers (Bowers et al., 2013). Additionally, GPs require a brief, stand-alone screen of fitness to drive rather than a battery of tests due to practical limitations, such as short patient consultation time. Wood et al. (2013) reported a similar sensitivity for MMSE compared to Bowers et al. (2013) (0.65 compared to 0.67 respectively) but particularly low specificity (0.37) (Wood et al., 2013). Hollis et al. (2015) found that neither test is predictive for drivers without cognitive impairment (Hollis et al., 2015). The MoCA was a stronger predictor for drivers with cognitive impairment but neither test was sufficiently accurate (Hollis et al., 2015).

MoCA was recommended as quick cognitive screen for health professionals to use in determining the need for on-road assessment in one study of patients with neurological impairments (N=135) (Esser et al., 2016). However, whilst specificity was high (94%) results indicated low sensitivity (44%). The authors proposed a lower cut-off score of <12 (sensitivity 100%; specificity 16.7%) and an upper cut-off of >27 (sensitivity 100%; specificity 4.9%). However, these proposed cut-off scores placed 78% of participants in the further testing category, impacting clinical utility of the test (Esser et al., 2016). In summary, researcher evidence consistently shows that MMSE and MoCA do not have sufficient predictive accuracy to be use alone to make decisions about fitness to drive (Asimakopulos et al., 2012; Bédard et al., 2008; Bowers et al., 2013; De Raedt & Ponjaert-Kristoffersen, 2001; Dickerson et al., 2014; Hollis et al., 2015; Kay et al., 2012; Vrkljan et al., 2011; Wood et al., 2013).

Summary. The fitness-to-drive tests identified in the literature with sufficient accuracy to predict on-road assessment outcomes were compared with GP fitness-to-drive test design criteria identified in the literature (summarised earlier in Table 3.1): so feasibility for medical practice could be determine. Table 3.2 contains the results.

Table 3.2

*Comparison of Current Fitness-to-Drive Screens with GP Fitness-to-Drive Screen Design Criteria Identified in the Literature**

Test name and reference	Evidenced based cut-off scores	Two cut-off scores	Sensitivity, specificity, PPV, and NPV above 80%	Criterion measure: Standardised on-road assessment	10-20% classified "further testing"	Face valid for driving	Brief (≤ 10 minutes)	Practical for medical practice?
DriveAble	✓	✓	✗	✓	✗	✓	✗	✗
Dobbs, 2005	Pass, intermediate & fail categories Dobbs et al., 2013	Dobbs et al., 2013	Not above 80% Korner-Bitensky et al., 2009	Standardised on-road test Korner-Bitensky et al., 2009	46% Dobbs et al., 2013 36.5% Korner-Bitensky et al., 2009	Some items have face validity	35-50 minutes Korner-Bitensky & Sofer, 2009	- Special testing console - Results must be sent to DriveAble for interpretation - Too lengthy
Original DriveSafe DriveAware	✓	✓	✓	✓	✗	✓	✗	✗
(DSDA) Kay et. al., 2009a	Hines & Bundy, 2014; Kay et al., 2009a; Kay et. al., 2009b	Kay et al., 2009a; Kay et. al., 2009b	Hines & Bundy, 2014; Kay et al., 2009a; Kay et. al., 2009b	Kay et al., 2009a; Kay et. al., 2009b	50% Kay et al., 2009a	Driving related scenes / task Kay et al., 2009a	20-30 minutes	- Training required - Not available to GPs - Data projector & screen

Note: ✓= Yes; ✗ = No; N/A = Not applicable; *Criteria references listed in Table 3.1

Test name and reference	Evidenced based cut-off scores	Two cut-off scores	Sensitivity, specificity, PPV, and NPV above 80%	Criterion measure: Standardised on-road assessment	10-20% classified "further testing"	Face valid for driving	Brief (≤ 10 minutes)	Practical for medical practice?
Sensory Motor Cognitive Tests (SMCTests) Innes et al., 2007	✗ Evidence based / logistic regression model for predicting on-road performance Hoggarth et al., 2013; Innes et al., 2013	✗	✗ Variable across studies. Below 80% in more recent research Hoggarth et al., 2013	✓ Innes et al., 2007; Innes et al., 2009	N/A	✓ Car body apparatus and driving related tasks	✗ Multiple computerised subtests / reported as several hours in one study Heitger et al., 2004	✗ - Specialised equipment and training - Too lengthy
Stroke Drivers Screening Assessment / Nordic SDSA (SDSA / NorSDSA) Lincoln et al., 2012	✓ Score predicts pass / fail (higher score = better result) Lincoln et al., 2012; Selander et al., 2010	✗ - Predicts on-road pass / fail - Large overlap of scores Lincoln et al., 2012	✗ Contradictory findings / most recent research indicates cannot predict on-road test outcome Selander et al., 2010	✓ Asimakopulos et al., 2012; Selander et al., 2010	✗ Large overlap of scores Lincoln et al., 2012	✗ Group of timed cognitive tests	✗ Approximately 30 minutes Selander et al., 2010	✗ - Too lengthy

Note: ✓= Yes; ✗ = No; N/A = Not applicable

Test name and reference	Evidenced based cut-off scores	Two cut-off scores	Sensitivity, specificity, PPV, and NPV above 80%	Criterion measure: Standardised on-road assessment	10-20% classified "further testing"	Face valid for driving	Brief (≤ 10 minutes)	Practical for medical practice?
Useful Field of View (UFOV) Ball et al., 1988; Ball & Owsley, 1993	✗ No consistent, reliable cut-off scores identified in the research	✗	✗ Variable / Not sufficiently accurate to predict driving performance	✓ For some studies e.g., Austroads, 2004; Bédard et al., 2008; Myers et al., 2000	N/A	✗ Limited face validity	✗ 15 minutes (short enough to be administered by a practice nurse)	✓ - 15-minute administration time - Easy instructions - Computer administration

Note: ✓= Yes; ✗ = No; N/A = Not applicable

These findings indicate no fitness-to-drive screen meets all the criteria recommended by researchers for a practical and valid driver screen, consistent with findings from published fitness-to-drive test reviews (Bédard et al., 2008; Dalchow et al., 2010; Dickerson, 2014b; Fildes, 2008; Molnar et al., 2006). Short, simple tests (e.g., UFOV ™) do not have the discriminative power required to screen drivers (Kay, Bundy, & Clemson, 2008). Longer tests approaching acceptable predictive validity (e.g., DriveAble, SMCTests, and original DSDA) are not practical for medical practice due to lengthy administration times, the need for specialised equipment and training, or lack of face validity (Dalchow et al., 2010; Fildes, 2008; Kay, 2008; Kay, Bundy, & Clemson, 2008; Molnar et al., 2006). One limitation of the Chapter 3 literature review was that it was not a comprehensive review. However, the findings were consistent with those of numerous other researchers (Asimakopulos et al., 2012; Bédard et al., 2008; Dickerson, 2014b; Kay et al., 2012; Molnar et al., 2006; Vrkljan et al., 2011) who have conduced comprehensive reviews and concluded that GPs did not currently have access to one tool that can be used to screen drivers. This outcome supported the dedication of resources to the conversion of original DSDA into a brief, practical, and valid touchscreen fitness-to-drive screen for GPs.

CHAPTER 4

Extended Methodology: Development of Touchscreen DSDA

The conversion of original DSDA into a touchscreen fitness-to-drive screen for GPs and other health professionals involved many steps. We conducted the following processes prior to the touchscreen development: (a) needs analysis to determine if there was a demand for a digital version of DSDA, with findings presented to Pearson to secure project funding (findings discussed later in this chapter); (b) project planning, research design, and submission of ethics applications; (c) identification of optimum touchscreen test design criteria for GPs and older adults (findings discussed later in this chapter); and, (d) a tender process to select a suitable app developer.

We conducted the following processes in the design and programming phase: (a) developed product specifications and prepared elements for usability testing in conjunction with the app developer; (b) conducted usability testing with GPs, occupational therapists, and older adults concurrently with app design and programming to ensure the app was user friendly (including set up of research sites, participant recruitment, and data analysis as per the ethics applications); (c) developed an audio script and conducted studio voice recording of in-app instructions; and, (d) conducted systematic touchscreen DSDA quality assurance checks and product testing throughout all development stages to ensure the app was user-friendly and free from design faults.

We conducted the following steps in the validation research phase: (a) set up of ten research sites: including equipment provision, administrator training, and set up of data collection and storage mechanisms as per the ethics application; (b) data analysis including evaluation of construct validity and internal reliability, setting touchscreen DSDA cut-off scores, and determining sensitivity, specificity, PPV and NPV; and, (c) finalisation of the fitness-to-drive classification process based on the cut-off scores, including ensure this was programmed and scenario tested with the app developer.

Finally, we conducted the following processes prior to release of the iPad app into the Apple Store: (a) designed the touchscreen DSDA in-app results and reporting mechanisms for administrators: including consent forms, report forms, disclaimers, and advisory warnings; (b) wrote the technical manual (see Appendix A); (c) assisted Pearson to develop product technical support and maintenance systems, customer support, and sales staff training; (d) assisted Pearson to develop marketing content (e.g., website and flier product information); (e) developed and provided webinar training for touchscreen DSDA administrators; (f) released touchscreen DSDA to the Apple Store for approval then to the market via the app developer; and, g) identified GPs and medical specialists willing to implement touchscreen DSDA into their medical practices as part of a study planned to evaluate the feasibility of the test for this context: including research site set up and data collection (discussed in Chapter 8).

At the beginning stage of the project we had a concept of how a touchscreen version of DSDA might look. However, we needed to confirm our design assumptions with the intended administrators (general practitioners) and the group most likely to take the test and most likely to struggle with the technology (older adults). We assumed GPs required a brief test but were unsure how GPs would define brief (e.g., 3 or 10 minutes). We did not know if GPs would prefer to administer the test themselves, in conjunction with a practice nurse, or via referral to an external service (e.g., an occupational therapist). We also did not know if

GPs had access to an iPad or an android tablet and which mode of administration they would prefer. Additionally, we were unsure what mode of digital administration would be optimum for older adults and how the app should be best designed to suit their needs. Therefore, we consulted GPs directly and I conducted a review of the literature as described below.

Fitness-to-drive Screen Needs & Preferences Identified by GPs

I needed to understand GPs' needs and preferences regarding practical test design prior to commencing app design, so the final product would be functional and useful for them. GPs' views and practices regarding assessment of patient fitness to drive have been explored via a large number of surveys investigating: (a) GP familiarity with current legislation and the need for further education (Hoggarth, 2013; Jang et al., 2007; Kahvedzic, McFadden, Cummins, Carr, & O'Neill, 2015; Pfaffli, Thali, & Eggert, 2012; Wilson & Kirby, 2008); (b) opinions and attitudes towards the driver-assessor role (Jang et al., 2007; Jones, Rouse-Watson, Beveridge, Sims, & Schattner, 2012; Kahvedzic et al., 2015; Marshall et al., 2012; Omer, Dolan, Dimitrov, Langan, & McCarthy, 2014; Sims et al., 2012); (c) current clinical practice regarding fitness-to-drive assessment (Braekhus & Engedal, 2009; Hoggarth, 2013; Jang et al., 2007; Marshall et al., 2012; Wilson & Kirby, 2008); and, (d) managing the consequences of the fitness-to-drive assessment (Jones et al., 2012; Nouri, 1988). However, there is a paucity of research exploring GPs' needs and preferences regarding the practical aspects of fitness to drive test design (e.g., how long it should take and how it should be administered).

In the absence of published research to guide test design, Pearson surveyed a representative sample of 200 Australian GPs via a market research company. Results are reported in Appendix C (Brown, Cheal, Cooper, & Joshua, 2013). GPs reported they needed a driver screen, but it must be brief (*M* and *Mdn* = 10 minutes); valid, simple to administer in an office setting; and have the option of practice nurse administration (Brown et al., 2013). Most GPs (84%) had access to a practice nurse whom they considered could spend slightly longer time administering the screen (*M* = 14m; *Mdn* = 15m). Most GPs (84%) preferred in-house administration, either via the GP, a practice nurse, or a combination. Only 10% preferred referral to another professional to complete the screen and 6% considered the test would not be feasible. Most GPs (78%) reported iPad access; only 7% preferring android tablets (Brown et al., 2013). These findings contributed to the decision to: (a) select iPad rather than android tablet as the administration mode; (b) allow a 10-minute administration time; and, (c) design the test to be partially self-administered with capacity for a practice nurse to setup and supervise the self-administered components, to reduce testing time. The next step was to ensure touchscreen DSDA was designed to be user-friendly for those who will take the test.

Suitability & Design of Touchscreen Technology for Older Adults

Because of a world ageing population and the higher rates of cognitive impairments among older adults (Global Burden of Disease Study 2013 Collaborators, 2015), the group most likely to take touchscreen DSDA are older adults. Therefore, age-related changes and the requirements of older adults needed to be taken into consideration at the commencement of the design process (Rogers & Mitzner, 2016). Consequently, I conducted a review of the literature to determine the suitability of touchscreen administration for older adults and to determine which design criteria should be applied to ensure the test was user-friendly for older adults who may be less familiar with digital technology. I searched three online databases: Medline via OvidSP (1946 to present), Cinahl via Ebsco (1982 to present); and, Ageline via Ebsco (1978 to present). Search terms included: older adults; touchscreen; digital tablet; iPad; mobile application; hand held computers; and, health information technology. I reviewed reference lists of relevant articles and articles citing these publications. The literature search did not have a time or locality limitation. I also searched for statistical reports and surveys available on the Internet via research companies such as Nielsen and Price Waterhouse Coopers, due to the rapid changes occurring in digital technology development and adoption rates in the community.

Survey results (Nielsen & IAB Australia, 2015; Statista, 2017) indicated rapid uptake of touchscreen technology broadly across age groups. Fifteen million Australians own a smartphone and eleven million own a tablet, with demand increasing rapidly each year (Nielsen & IAB Australia, 2015); September 2015 was the first month in which time spent on a tablet exceeded time spent on a personal computer (Nielsen & IAB Australia, 2015). Convenience, portability, user-friendly operating systems, and larger screens are driving the high uptake of mobile touchscreen devices (Statista, 2017). I sought to determine if touchscreen technology was suitable for application in health care, particularly among older adults.

The National Health Service (NHS) in the United Kingdom noted a growing gap between the proliferation of digital devices for everyday tasks and the continued delivery of health care via traditional modes, such as face-to-face consultation and paper administration of forms and tests (Clionsky & Clionsky, 2014). The NHS estimates significant cost-benefit (£5,059 million over 10 years) in transitioning health services online for access via digital devices (e.g., for assessment of minor ailments, primary care pre-assessment, pre-operative screening, post-surgical and secondary care remote follow-up, and administrative tasks such as appointment bookings and test result notifications) (National Health Service, 2011; Price Waterhouse Coopers & Department of Health, 2013).

The benefits of digital modes of administration are clear for clinicians, who readily adopt digital tablet apps for healthcare delivery due to the time and accuracy advantages (Daly, Xu, & Levy, 2015; Downing Peck, 2011; Glaser, Jain, & Kortum, 2013; Godwin, Tan, Bockhold, Ma, & Tran, 2015; Howell, Hood, & Jayne, 2015; Turney & Reynard, 2014; Wiarda, McMinn, Peterson, & Gregor, 2014). More than 165,000 medical apps are available to clinicians, with thousands added each year (McCarthy, 2015). Patients miss significantly fewer items when they self-administer assessments via healthcare apps compared to paper administration (Matthew et al., 2007; Ryan, Corry, Attewell, & Smithson, 2002; Wilson et al., 2002). Additionally, equivalence studies indicate that touchscreen and paper administration of standardised assessments are equivalent measures (Clionsky & Clionsky, 2014; Lavi, Malki, & Kornowski, 2014; Matthew et al., 2007; Ryan et al., 2002; Wilson et al., 2002). Further, there is evidence that capacity to form a therapeutic patient relationship is not reduced by touchscreen test administration compared to traditional methods (Clough & Casey, 2015; Eonta et al., 2011; Wiarda et al., 2014). The ubiquity of touchscreen personal devices and cost-benefit for clinicians supported the selection of touchscreen for DSDA administration. However, I sought to determine if patients, particularly older adults, were equally positive about assessments being presented to them via digital tablet and if they could successfully use them.

Older adult perspectives regarding touchscreen technology use. Stereotypes suggest older adults are less interested in technology, have lower uptake, and struggle with the format (Kunemund & Tanschus, 2014). Some surveys indicate that older adults are less likely to engage in health services via technology or perceive the need for digital devices (Gitlow, 2014; Gordon & Hornbrook, 2016; Kontos, Blake, Chou, & Prestin, 2014). Nonetheless, Kunemund and Tanschus (2014) advised against treating older adults as a homogeneous group when it comes to understanding adoption of technology due to the complex interplay of factors such as health status, living situation, available supports, socio-economic status, and needs on uptake. Acceptance also changes with the specific technology: whether it is seen as socially desirable, easy to use, and useful (Kunemund & Tanschus, 2014).

There is anecdotal evidence that older adults readily adopt touchscreen technology, finding smart phones and tablets more user friendly that computers. However, there are relatively few studies that examine the use of iPad technology with older adults (Delello & McWhorter, 2016). Moussa et al. (2017) conducted a literature review to examine the implementation of mobile health technology among older adults aged ≥65 diagnosed with a mental illness or cognitive impairment. Only 7 out of the 1941 studies reviewed that met the inclusion criteria (i.e., that participants must be diagnosed with a mental illness), examined mobile health technologies (Moussa et al., 2017); 3 addressed touchscreen technology, and only one (Onoda et al., 2013) examined self-administration of an iPad application. Onoda et al. (2013) found that iPad could be successfully used to screen a large ($N = 222$) older population (M_{age} 70.7) for dementia. However, usability was not evaluated from the perspective of the participants.

Studies evaluating delivery of healthcare via digital tablets in hospital and community settings indicate technology is not a barrier for older adults, particularly in the case of iPad (Cook et al., 2014; Dixon, Bunker, & Chan, 2007; Turney & Reynard, 2014). For example, results of a feasibility study conducted to evaluate delivery of education to cardiac surgery in-patients ($N = 149$; $M_{age} = 68$ years; Range 52-90 years), via iPad, indicated 84% of education modules presented over 5 days were consumed and 90% of patients reported understanding 90% of the content (Cook et al., 2014). Increasing age was not associated with decreasing iPad use and the older patients quickly learnt to use the technology, despite being initially unfamiliar with it. The authors perceived a cost-benefit in reducing length-of-stay and re-admission due to providing engaging, simple, and timely education regarding daily healthcare, without reliance on busy hospital staff to deliver education verbally (Cook et al., 2014). Perhaps surprisingly, iPads have also been used successfully with older adults with dementia, aiding participation in group well-being activities (Leng, Yeo, George, & Barr, 2014), aiding reminiscence and recall, and increasing communication and socialisation (Upton, Upton, Jones, Jutlla, & Brooker, 2011; Yack & Camic, 2017).

The following three studies examined older adults experiences using digital tablet devices in community and residential care settings. Tsai, Shillair, Cotten, Winstead, & Yost. (2015) conducted 21 semi-structured interviews with older (69-91), community based residents who owned a digital tablet, to examine technology self-efficacy. Participants reported the decision to adopt a tablet device was driven by recommendation from family or seeing others with the technology. All participants reported finding digital tablets quick and easy to learn to use and user-friendly; all were happy with their device. Participants described feeling confident and "current" using their tablets; 90% (n = 19) felt more connected to family and the world since adopting the device (Tsai, Shillair, Cotten, Winstead, & Yost, 2015). There was a clear difference in perceived self-efficacy using a tablet compared to a computer, with computers described as being frustrating, complex and intimidating. Residents valued their iPads due to their size, portability and convenience (Tsai et al., 2015). These findings are consistent to those of Watkins and Xie (2015) who explored older adult (N = 22; age ≥60) perceptions of iPad use for increasing fruit and vegetable consumption via relevant applications; participants were unfamiliar with the technology. Five focus group sessions were conducted incorporating instruction in iPad use. The participants reported similar influencers regarding choice to adopt (family, social contacts and advertising) and found iPad user friendly, positive and engaging (Watkins & Xie, 2015).

Østensen et al. (2017) provided 15 residents in aged care facilities (M_{age} 78.3) with an internet-connected iPad as part of an innovative care programme to promote thriving and participation. All had little prior exposure to digital tablets (Østensen, Gjevjon, Øderud, & Moen, 2017). Participants attended a weekly, 12-month iPad training programme in how to use various iPad applications (e.g., Pintrest and Facebook) and functions (e.g., email). Training was supported by community care nurses and provided by adolescent youth volunteers. Semi-structured interviews were conducted pre-training, after 6 months, and after 12 months (Østensen et al., 2017). All older adults reported finding iPads easy to use and feeling more socially connected after training. For example, many participants re-established contact with family and friends they could no longer visit. A group of participants formed a social group based on a mutual interest in music, using iPad applications such as Spotify. One resident wanted to try online shopping to increase her sense of independence (Østensen et al., 2017). Participants described a sense of mastery, satisfaction and pride in learning to use an iPad, and found them an interesting and positive addition to their lives. Community nurses reported the residents seemed happier, more social, and left their rooms more often. The authors proposed a model of care using technology to prevent isolation and functional decline based on the success of the intervention (Østensen et al., 2017).

Whilst both these studies provide evidence that older adults find iPads intuitive and simple to learn to use, sample sizes were small and results cannot be generalised. However, the views of participants were consistent, and findings supported the use of touchscreen tablet devices with older adults. No studies were found indicating iPad was unsuitable for the healthcare context or standardized assessment delivery for older adults. Regardless, the unremitting pace of technological development and almost universal adoption of digital devices within the community suggest use of touchscreen devices in healthcare will only increase, particularly as cohorts who are native users of technology age. The findings from this literature review gave me confidence that iPad administration of assessments is advantageous for both clinicians and older and cognitively impaired adults. Next, I searched the literature to identify suitable touchscreen design criteria for older adults.

DSDA iPad app design for seniors. Many older adults experience age-related changes that could impact interaction with an iPad (e.g., reduced vision, hearing loss, reduced reaction time, reduced coordination, and cognitive changes) (Loureiro & Rodrigues, 2014). We wanted to design an interface that would take these limitations into consideration so DSDA examined the desired construct (awareness of the driving environment) and not individual differences in ability to enter responses via iPad. Motor and cognitive performance in touch based functions (e.g., tap and drag) are critical to a person's ability to successfully interact with an iPad (Findlater, Froehlich, Fattal, Wobbrock, & Dastyar, 2013). Many studies have examined the performance of older adults with touchscreen functions to determine what works best. A study comparing iPad to computer mouse performance of 20 older (M_{age} = 74) and 20 younger (M_{age} = 28) adults found a lower error rate and faster performance on iPad for both groups (decreased by 35% for older adults; 16% for younger) (Findlater et al., 2013). Consistent with other research (Cockburn, Ahlström, & Gutwin, 2012; Kobayashi et al., 2011), drag was the slowest iPad function for both groups, perhaps because the object is obscured by the finger or due to the additional dexterity and pressure required to "hold" and slide the target against resistance. However, participants considered all touchscreen functions relatively easy (Findlater et al., 2013). In a study comparing finger, stylus and mouse inputs, fingers were slowest for dragging (Cockburn et al., 2012). Fingers were fastest for tapping but less accurate for smaller targets, probably due to the target being obscured by the finger and the finger being a larger, cruder pointing device. However, overall error rate was very low for all three modes (Cockburn et al., 2012).

Useful guidelines arose from this body of literature that we applied when designing the touchscreen interface for older adults. The most relevant for this project included: larger sizes for targets (≥ 8mm); larger, simple fonts; highly contrasting colours, targets, and backgrounds; adjustable audio volume; caution in design of 'drag' tasks: testing these with seniors first; simple, meaningful icons; avoidance of scrolling or distracting and irrelevant elements; and reduction of the gap between intended and actual touch location due to parallax (i.e., where object position appears to differ when viewed from different angles) (Kobayashi et al., 2011; Loureiro & Rodrigues, 2014; Schneider, Wilkes, Grandt, & Schlick, 2008).

In summary, comparison of keyboard, mouse, and touchscreen input methods consistently show that older adults perform best on touchscreen devices (Cockburn et al., 2012; Findlater et al., 2013; Kobayashi et al., 2011; Schneider et al., 2008) and that iPad reduces the performance gap between older and younger adults compared to traditional desktop setups (Findlater et al., 2013; Schneider et al., 2008). These finding supported the choice of iPad as the optimum mode of digital DSDA administration for older adults. However, in addition to the literature review we required a clearer understanding of how older adults may enter responses via touchscreen in practice, to identify potential barriers to effective performance. Therefore, we took a user-centred design (UCD) approach to the touchscreen DSDA development: to ensure test design was informed by the needs and preferences of the intended user group and that we avoided potentially costly and time-consuming user-interface problems during the clinical research phase. A UCD approach included usability testing elements of touchscreen DSDA with older adults in the field, concurrently with the design, programming, and evaluation stages. Usability can be defined as the extent to which a product or service can be used with effectiveness, efficiency and satisfaction by the target user to achieve specified goals in the specified context of use (Hussain & Kutar, 2009; International Organization for Standardization, 2010; Nayebi, Desharnais, & Abran, 2012) and has been described as a cornerstone of best practice when designing medical devices (Hegde, 2013). I presented the results in Chapters 5 and 6.

CHAPTER 5

Journal Article 1

Usability testing of touchscreen DriveSafe DriveAware with older adults: A cognitive fitness-to-drive screen

The following manuscript describes a study in which we examined the usability of touchscreen DSDA with older adults concurrently with the app design, programming, and evaluation. This allowed us to test and validate our design assumptions with the group most likely to take touchscreen DSDA and most likely to have difficulty with the technology. We also tested the automatic data collection system we had designed to ensure touchscreen DSDA reflected the data collection decisions that would have been made by a trained administrator. The present study was essential to the development of touchscreen DSDA because we could not have predicted what difficulties older adults may have without usability testing. I formatted this chapter in accord with the style of the journal to which the manuscript was submitted. The manuscript was published by Cogent Medicine on December 5, 2018 (Cheal, Bundy, Patomella, & Scanlan, 2018).

Statement of Contribution of Author

On behalf of the co-authors of "Usability Testing of Touchscreen DriveSafe DriveAware with Older Adults: A Cognitive Fitness-to-Drive Screen", we confirm that Beth Cheal has made the following contributions:

- Conceptualisation and design of the research
- Collection of data
- Analysis and interpretation of findings
- Writing the paper and critical appraisal of content

Anita Bundy Signed: ____________________ Date: 04.09.2018

Beth Cheal Signed: ____________________ Date: 04.09.2018

Usability testing of touchscreen DriveSafe DriveAware with older adults:

A cognitive fitness-to-drive screen

Beth Cheal, MAppSc(OT) [a,c], Anita Bundy, ScD[a,d], Ann-Helen Patomella, PhD[b], Justin Newton Scanlan, PhD[a]

[a] *University of Sydney, Sydney, Australia;* [b]*Karolinska Institutet, Stockholm, Sweden;* [c]*Pearson, Sydney, Australia;* [d]*Colorado State University, Fort Collins, US.*

Abstract

Background. DriveSafe DriveAware is a cognitive fitness-to-drive screen that can accurately predict on-road performance. However, administration is restricted to trained assessors. General practitioners are ultimately responsible for determining fitness to drive in many countries but lack suitable tools. We converted DriveSafe DriveAware to touchscreen to provide general practitioners and other health professionals with a practical fitness-to-drive screen. This necessitated the development of an automatic data collection system. We took a user-centred design approach to test usability of the system with older adults, the group most likely to take the test.

Method. Middle aged and older adult volunteers were asked to try an iPad application to assist in the development of a fitness-to-drive screen. Seventeen males and 18 females (mean age 70 years) participated in four trials, each participant tested only once. We tested all text and function changes until all older adults could successfully self-administer the screen.

Results. Older adults found basic touchscreen functions easy to perform, even when unfamiliar with the technology.

Conclusion. Usability testing allowed us to develop a user-friendly touchscreen data collection system and ensured design errors were not missed. Psychometric evaluation of data gathered with touchscreen DriveSafe DriveAware is required prior to use in clinical practice.

Keywords. Older adults; automobile driving; cognitive assessment screening instrument; touchscreen

Background

People in many cultures consider driving as one of their most valued daily living activities (Al-Hassani & Alotaibi, 2014; Dickerson et al., 2012; Fricke & Unsworth, 2001). Gaining a license is considered a rite of passage for young people and older adults want to drive for as long as possible, often with no plan for cessation (Coxon & Keay, 2015; Kostyniuk & Shope, 2003). However, driving is complex and therefore easily disrupted by illness, injury, or age-related changes. Chronic medical conditions, particularly among older drivers, are associated with increased crash risk and driving errors (Barco et al., 2015; Charlton et al., 2010; Dobbs et al., 1998; Kay, Bundy, & Clemson, 2008; Marshall, 2008; Marshall & Man-Son-Hing, 2011; Marshall et al., 2007; Molnar et al., 2006; Papa et al., 2014). Despite this, researchers recommend fitness to drive is determined on an individual basis, with a focus on functional status, rather than diagnosis (Charlton et al., 2010; Marshall, 2008; Marshall & Man-Son-Hing, 2011).

General practitioners are the professionals ultimately responsible for determining medical fitness to drive in most countries and are in an ideal position to screen drivers because: 1) patients usually present to them in the first instance; 2) they are required to fill out license authority medical forms (Dobbs et al., 1998; Sims et al., 2012); and, 3) there is mandatory reporting of medically 'at risk' drivers in jurisdictions of many countries including the US, Canada, and Australia (Austroads & National Transport Commission, 2016; Jang et al., 2007). Surveys show general practitioners believe they should be responsible for making determinations about fitness to drive but lack valid and reliable driver screens that are practical for use in medical practice (Dobbs et al., 1998; Fildes, 2008; Jang et al., 2007; Marshall et al., 2012; Molnar et al., 2006; Sims et al., 2012; Wilson & Kirby, 2008; Woolnough et al., 2013; Yale et al., 2003).

The desktop (original) version of DriveSafe DriveAware is a cognitive fitness-to-drive test showing promise as a driver-screening instrument. Data gathered with original DriveSafe DriveAware are face valid, reliable, sufficiently predictive, test-retest reliable, and trichotomise patients via two evidence-based cut-off scores based on the likelihood of passing an on-road assessment (i.e., "Likely to Pass", "Requires Further Testing", and "Likely to Fail") (Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2008; Kay et al., 2009a, 2009b; Kay et al., 2012; O'Donnell et al., 2018). However, original DriveSafe DriveAware is not practical for medical practice and requires a trained administrator. Therefore, we further developed the test so it would be suitable for administration by general practitioners. Prior to development of the new screen, we surveyed a representative sample of 200 Australian general practitioners to identify their preferences regarding a driver screen (Brown et al., 2013). General practitioners reported they needed a brief (mean and median 10 minutes), valid and simple test. Thus, we designed DriveSafe DriveAware to be largely self-administered via iPad, with capacity for a practice nurse to setup and supervise the self-administered components.

Because the majority of drivers likely to take touchscreen DriveSafe DriveAware will be older adults (typically classified as age ≥65), we wanted to be sure that older adults could use tablet technology and feel comfortable with it. Whilst research findings suggest that older adults find digital tablets easy to learn to operate, intuitive and user-friendly (Delello & McWhorter, 2016; Moussa et al., 2017; Onoda et al., 2013; Østensen et al., 2017; Tsai et al., 2015), older adults experience age-related changes that could impact interaction with a digital tablet (e.g., reduced vision, hearing loss, reduced reaction time, reduced coordination, and cognitive changes). We wanted to design an interface that would consider these limitations so that touchscreen DriveSafe DriveAware examined the desired construct (i.e., awareness of the driving environment and one's own driving performance) (Kay et al., 2009a) and not individual differences in ability to enter responses via a touchscreen. First, we reviewed optimum touchscreen design guidelines for older adults. The most relevant for this project included: large targets (minimum .31" or 8 mm); large, simple fonts; high-contrast colours; contrasting targets and backgrounds; caution in the design of drag tasks including testing with seniors first; avoidance of scrolling; simple and meaningful icons; and, avoidance of distracting or irrelevant elements (Kobayashi et al., 2011; Loureiro & Rodrigues, 2014; Schneider et al., 2008).

We took a user-centred design approach to avoid potentially costly, time-consuming user-interface problems in the clinical research phase. Hegde (2013), described usability testing as the cornerstone of best practice when designing medical devices. Usability is defined as the extent to which a product or service can be used with efficiency, effectiveness, and satisfaction by the target users to achieve specified goals in the specified context of use (International Organization for Standardization, 2010). The user-centred design philosophy places end users at the centre of the design process (Dorrington, Wilkinson, Tasker, & Walters, 2016; McCurdie et al., 2012). Elements of the design are refined via an iterative process (Hegde, 2013; McCurdie et al., 2012; Rogers & Mitzner, 2016).

We sought to make the touchscreen version of DriveSafe DriveAware as similar as possible to the original version in order to retain test validity. However, we were transitioning from a test where a trained administrator collected, interpreted and scored variable data via participant verbal responses, to a test where variable data were collected via participant touchscreen responses and scored automatically. This necessitated the development of an automatic variable data collection and scoring system that would reflect the decisions that would otherwise have been made by a trained assessor. In the present study we addressed the research question "Does the touchscreen data collection system we designed collect variable data in a way that is user-friendly for older adults who may be unfamiliar with the technology?" We sought to answer this question by testing the usability of the touchscreen DriveSafe DriveAware data collection system with older adults concurrently with touchscreen DriveSafe DriveAware software design and programming.

Method

The University of Sydney Human Research Ethics Committee provided approval for the study. We conducted four rounds of usability testing on four days over one month. Results from each round informed the next stage of design and programming. We aimed to test approximately 10 participants per round. Testing with larger numbers was not considered beneficial because a repeated pattern of errors emerged after trials with 7 to 10 individuals. These errors needed to be addressed in programming before further feedback was useful. Each participant was tested only once.

Setting. We conducted rounds 1 and 4 at a large aged care residential facility. We conducted rounds 2 and 3 at a community centre. Round 2 occurred within the context of a social group for older adults. DSDA was administered to Round 3 participants individually. Both centres were located in Sydney, Australia.

Participants. We placed an advertisement in community meeting areas at both centres, asking for older adult volunteers to assist in the development of a fitness-to-drive screen for general practitioners. Potential volunteers were informed that their information would be anonymous; we would provide no advice regarding their driving. Volunteers advised centre staff if they wished to participate. A total of 35 adults volunteered: 17 males and 18 females aged between 41 and 89 (mean age 70 years); 83% (29 participants) were over 65. We collected no identifying data. No one withdrew or was excluded after agreeing to participate. Participant characteristics including iPad use are listed in Table 1.

Table 1

Summary of participant characteristics

Round	*n*	Gender	M_{age} (Range)	Previous iPad use	
1	13	Male = 6 Female = 7	78 (66 – 89)	Never: Rare: Frequent:	5 3 5
2	9	Male = 4 Female = 5	74 (66 – 77)	Never: Rare: Frequent:	4 2 3
3	6	Male = 2 Female = 4	51 (41 – 59)	Never: Rare: Frequent:	1 1 4
4	7	Male = 5 Female = 2	77 (69 – 85)	Never: Rare: Frequent:	6 1 0
Total	35	Male = 17 Female = 18	70 (41 – 89)	Never: Rare: Frequent:	16 7 12

Note. Rare = <1 x month; Occasional = ≥1 x fortnight (no responses); Frequent = ≥1 x week

Rounds 1 and 4 participants lived in supported care units. Eleven were ambulant; 2 mobilised via wheelchair. Three reported a past stroke; 3 reported hearing impairment; 1 reported significant vision impairment. Round-2 participants were all retired, ambulant, generally well, driving, and living independently in the community. Round-3 participants were younger and more active than the other groups. All were in paid employment and driving. Inclusion of a middle-aged group allowed comparison with a different generational cohort. The educational status of the sample was: post graduate degree (*n* = 1), university degree (*n* = 7), college certificate (*n* = 6), completion of high school (*n* = 5), completion of middle high school (*n* = 9) and completion of primary school (*n* = 1), not reported (*n* = 6).

Instruments. DriveSafe measures awareness of the driving environment (Kay, Bundy, & Clemson, 2008; Kay et al., 2009a). Touchscreen DriveSafe consists of 10 images of a 4-way intersection (see example image in Figure 1). Each image includes between two and four potential hazards (i.e., people or vehicles). These hazards appear for 4 seconds then disappear, leaving only the blank intersection. Participants are asked to recall the hazards that were displayed, touching the blank intersection to identify hazard type, location, and direction of movement.

Figure 1. DriveSafe image example.

DriveAware measures awareness of one's own driving abilities (Kay et al., 2009a, 2009b). Touchscreen DriveAware (Kay et al., 2009b) consists of two self-administered questions and five questions administered by a general practitioners or suitably qualified health professional (see example question in Figure 2). The DriveAware items yield a discrepancy score based on the difference between the patient's self-ratings and the clinician's ratings, or performance in DriveSafe.

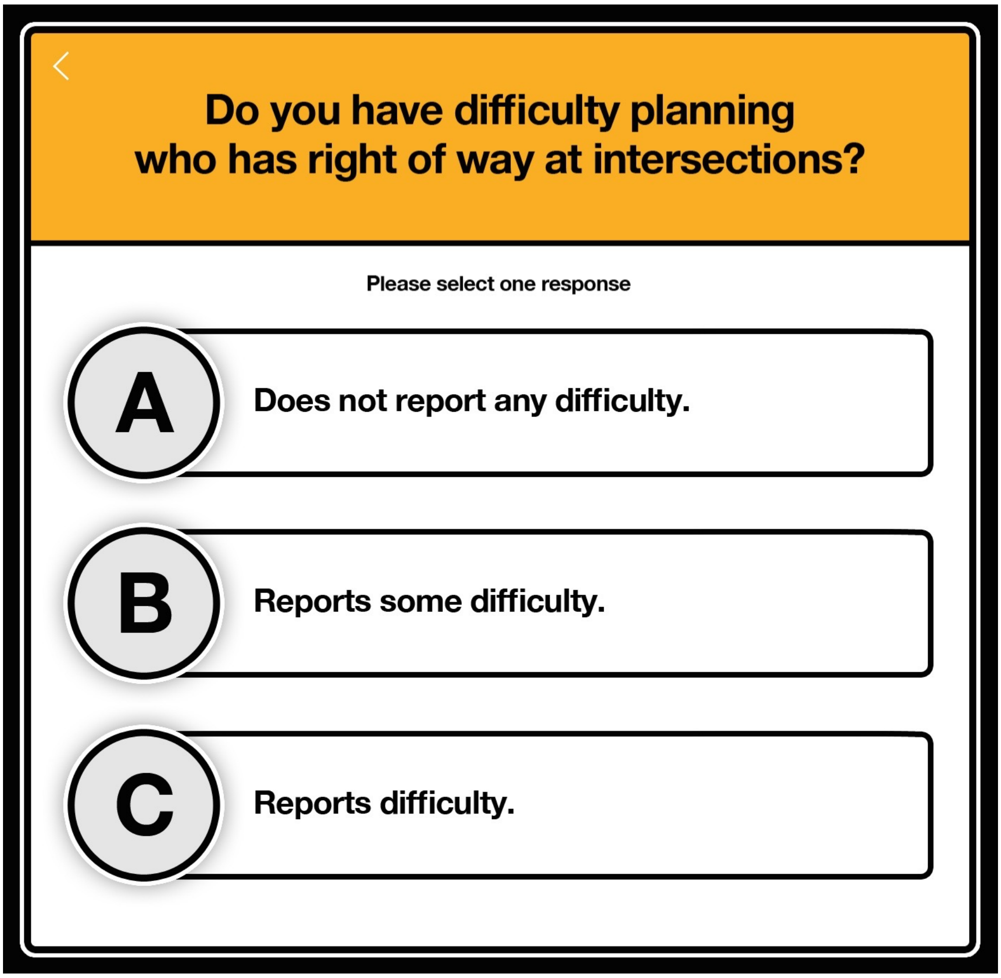


Figure 2. DriveAware image example.

Touchscreen DriveSafe DriveAware was presented to participants via a 3rd generation iPad (operating system 9), with a 9.7-inch, 2048 x 1536 mm (264 pixels per inch), multitouch, "retina" display. Headphones and a stylus were available to use depending on each participant's preference.

Procedure

We developed a pilot version of touchscreen DriveSafe DriveAware based on screen blueprints, which provided a visual guide of the skeletal framework and arrangement of elements in the iPad application. We performed several rounds of in-house testing and quality checks until we had created a satisfactory version for the present study.

The first author trialled the pilot version at the two centres. At the time of the trial, each participant sat on a chair with the iPad on a table in front. Volume and brightness were adjusted to full. Participants adjusted the position of the iPad to suit their focal length and chose whether to use a stylus or headphones. The examiner recorded participant actions and comments during testing with attention to apparent ability to understand test requirements, operate functions (e.g., tap, drag, undo, and buttons), and evidence of any anxiety or frustration. The examiner also recorded any technical difficulties related to programming.

The first author conducted a brief interview with each participant post testing including questions regarding test difficulty, ability to understand and follow written and audio instructions, and ability to operate the device.

After each round of testing, the first author discussed any difficulties encountered and potential solutions with the iPad application developers. Agreed programming and design changes were made and quality checks performed, followed by re-testing with participants. Testing continued until touchscreen DriveSafe DriveAware was fully programmed and participants could independently self-administer the test.

Analysis. The analysis focused on functional outcomes: whether participants could understand the test requirements, successfully perform the associated actions via the touchscreen (e.g., tap a target or adjust an arrow direction), and complete the required task in a timely manner and without errors. We assessed these outcomes against the project goals and objectives presented in Table 2.

Table 2.

Touchscreen DriveSafe DriveAware development goals and objectives.

Subtest / Area	Goals	Objectives
Test set up	Standardise administration without administrator training	– Optimize vision – Maximise hearing – Cue attention to the screen
Drive Safe	Enable user to successfully enter the following responses for each hazard observed: • Location • Type • Direction of movement	– Ensure timely and smooth progression of items – Ensure user can tap target location – Ensure user can tap object type from an array – Ensure user can input object direction of movement – Ensure user can undo unwanted responses
Drive Aware	Enable user to comprehend each question and enter one response from three options	– Enable user to read and hear each question and response options – Design user-friendly buttons – Ensure smooth and timely progression of items

Results

Following is a summary of the main challenges encountered and the solutions implemented prior to testing in subsequent rounds.

Test Setup. Usability testing provided important insight regarding optimum test set-up. For example, four Round-1 participants forgot to wear their reading glasses and all failed to attend to the screen at the commencement of each item. Some had difficulty entering responses via touch, largely due to incorrect finger angle when the iPad was flat on the table. We addressed these difficulties by adding a written and audio prompt to put reading glasses on if worn; adding a countdown (i.e., "3, 2, 1") and bell to cue timely attention; and placing the iPad on a stand angled to 20°. Result from Round-2 and -3 testing indicated these measures had resolved the testing difficulties; we identified no further test set-up challenges.

Three participants reported hearing loss in Round 4. One had a significant loss, did not wear hearing aids, and could not hear the instructions with full volume. However, all three participants reported hearing the instructions clearly once wearing headphones. One participant had a significant hand tremor that impaired touch ability. A stylus resolved this problem. Round-2 participants were asked to try both stylus and finger inputs and could use either equally successfully. Round-3 participants did not have any difficulty with test set-up.

DriveSafe. The primary goal for the DriveSafe evaluation was to determine user-friendly input methods.

Object location input. Round-1 participants (*n* = 9) triggered unwanted object location responses by resting a hand on the screen or by incorrect touch. Also, we encountered technical difficulties for 'tap' functions because some objects were situated too close together. We resolved all difficulties via use of an iPad stand and optional stylus, a programmed 0.5-second delay between taps to allow menus to open, and adjustment of object proximity. We encountered no further difficulties in subsequent rounds.

Object type input. Round-1 participants had difficulty with object-type menus overlapping each other or the screen edge, preventing option selection. We overcame these difficulties by programming a 1cm exclusion zone around the perimeter of the screen and by moving close objects further apart. Round-2 participants failed to notice or use the icon depicting two people (see Figure 3a). Therefore, we enlarged it and moved it to a more prominent location on the menu (from left to right). We added forced selection of that icon to a practice item and inserted an error message to provide clarification (i.e., "For 2 people walking together, use the 2-person icon"). We encountered no further difficulties in subsequent rounds.

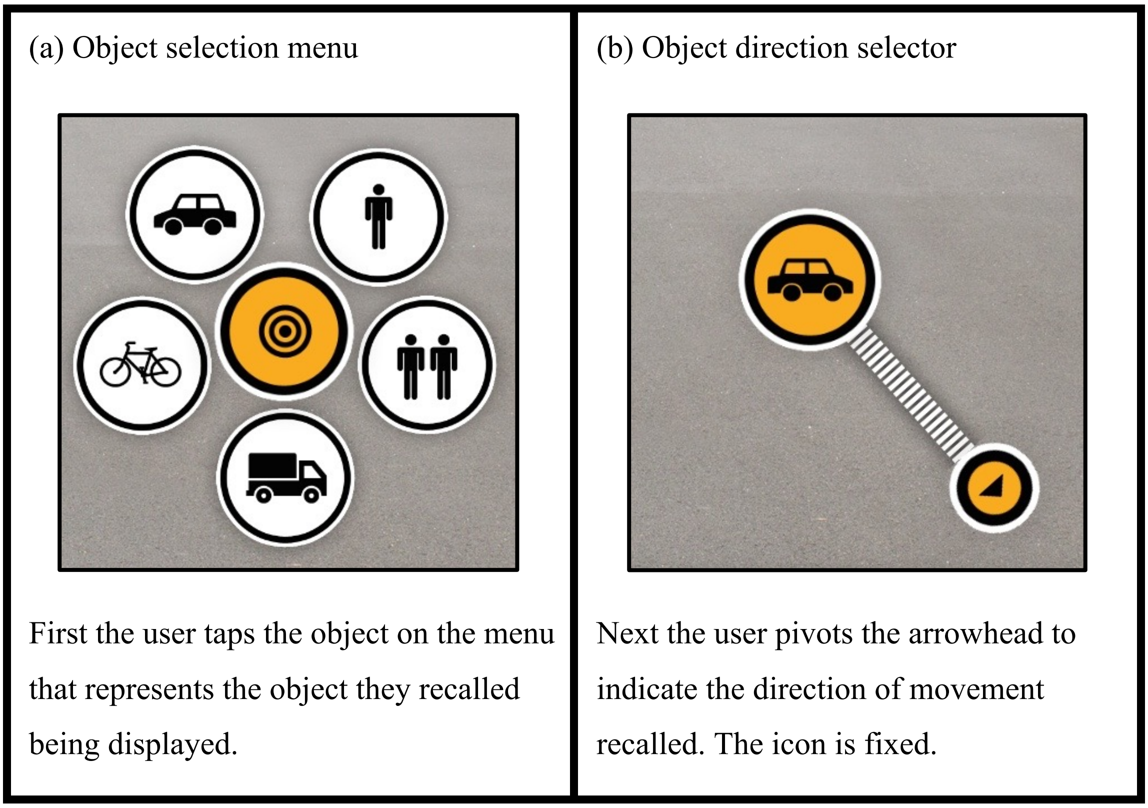


Figure 3. DriveSafe object selection menu and object direction selector.

Direction of movement input. We trialled two direction input methods in Round 1: an 8-way arrow icon (i.e., tap the arrow representing the object's direction) and a drag icon (i.e., drag the object in the desired direction). More people preferred "drag" to "arrows" (*n* = 10:2). One participant found both options difficult. Participants reported drag provided a more accurate reflection of their intended direction stating, "The arrows are not really spot on" and "I wish there was one in the middle". Because the drag function was the most successful, we discontinued arrow testing. Five participants in Round 2 had difficulty with the drag motion, either not dragging the icon far enough or trying to drag it too far. We resolved this in subsequent rounds by fixing the icon at the tap location and snapping out a ghosted arrow once the participant selected the object type. The ghosted arrow became solid once touched (see figure 3b). We also extended the drag radius and required a pivot action, which older adults found easier than drag alone.

We wanted to make the touchscreen interface user-friendly without prompting responses. For example, in the original version of DriveSafe, patients could 'forget' to provide a verbal response for the 'direction' category; this contributed to scoring and served to discriminate among individuals. We initially tried "no cue" for this category but each of the 13 participants forgot to enter at least one response. Lack of a visual cue was therefore not useful for discriminating ability among participants of varying ability levels. We addressed the problem of eliciting a response, whilst minimising cueing, by inserting the ghosted arrow. This prompted participants to enter a direction response but, if they failed to respond, they could still proceed.

Item progression. In Round-1, we identified a significant problem in self-administration of item progression via the "next" button. All participants over-focused on tapping this button, missing the brief display of objects. We corrected this by a) delaying the appearance of the “next” button until a first response had been entered; b) adding a written and audio instructions once objects had disappeared (i.e., “Tap where you saw the object”); and c) inserting a message at the conclusion of each item (i.e., “Are you sure you have completed this screen?”). Some participants tried to enter responses before objects had disappeared. This was largely resolved by the above programming changes. Inserting an error message was not viable, as this would have distracted participants from committing the hazards to memory.

Designing a user-friendly “undo” process proved challenging in all rounds. Middle-aged (Round-3 participants) and older participants had contrasting experiences and difficulties. We added undo instructions into the demonstration image and forced undo practice because Round-1 and -2 older adults did not understand how to use the undo button. However, Round-3 participants strongly disliked being forced to practice (e.g., stating, “I don’t think I should have to press undo. I didn’t make a mistake” and “No. I’m not undoing because I did it right”). Making practice optional in conjunction with additional instructions worked well for all adults. Round-3 participants did not have difficulty with any other challenges on usability testing.

Understanding of test requirements. In Round 1, only 1 of the 13 participants understood the test requirements. Participants continually pressed the next button, missing many items. We resolved this in Round 2 with forced practice of incorrect demonstration items and two levels of in-application instructions. To avoid disrupting test flow for participants who quickly understood test requirements, only participants having difficulties received second level instructions. The wording of some test instructions confused participants in Round 2. For example, one DriveSafe instruction stated, “You will now see 10 images of an intersection”. Including numerals confused the participants (e.g., “That's a problem. I thought I needed to look for 10 items"), so we removed them. Field-testing allowed identification of particular words or colours that were problematic. For example, participants failed to understand the word "marker". Thus, we substituted the word "arrow". Participants reported a red message background made them feel like they had done something wrong. We resolved this by changing the background colour to blue.

Round 4 participants had the least iPad exposure: only one participant reported owning an iPad but rarely using it. The others had never used one. One 79-year-old male with significant hearing loss, full vision loss in one eye, and glaucoma in the other eye, presented a particular challenge (he was still driving). He reported he had never seen an iPad and did not understand the touchscreen concept. Despite this he completed the test successfully and with ease, wearing headphones. The only difficulty he had was working out how to drag the direction icon. The second level of in-application instructions addressed the difficulty and he required no administrator assistance.

We learned about the need to create an administrator-assisted option for test administration because two participants struggled with the practice items (neither had used an iPad before). A brief verbal prompt addressed their difficulties but the first participant became frustrated after five item repetitions. We developed an administrator-assist procedure to meet needs we observed during the study. For example, we concluded that an administrator should intervene after four unsuccessful practice attempts.

We chose not to develop a solution for every observed problem. For example, one participant took twice as long as the others to complete the test and demonstrated behaviours not observed among other participants (e.g., placing a car in the foreground of every image although there was no object at this location). These difficulties may have related to test design or to the participant's cognitive deficits. Because we did not observe a similar problem in any other participant, we did not develop a programming solution.

DriveAware. Participants self-administered DriveAware with ease through all rounds of testing. Therefore, we made no adjustments. The original design worked well in practice.

Discussion

Consistent with findings from other research, older adults in this study found basic touchscreen operations easy to perform without training, even when unfamiliar with the technology (Findlater et al., 2013; Leng et al., 2014; Østensen et al., 2017; Tsai et al., 2015; Watkins & Xie, 2015). All participants quickly understood how to interact with the iPad, although 23 of 29 participants aged over 65 had never or rarely used one. Participants' overall response to the touchscreen test was positive; spontaneous comments from 4 participants indicated they found the test face-valid (e.g., "I think it is fair in the fact that it makes you look at your surrounds, which a lot of older people don't, and just basic road rules I guess" and "That seemed quite good. If my doctor made me do this to check my driving, that would be fair").

DriveSafe was the more difficult subtest to convert to iPad administration because it was the more complex aspect of DriveSafe DriveAware. The original version relied on an administrator to interpret patient verbal responses. It was fundamental to consider motor and cognitive performance required for touch-based functions such as "tap" and "drag" in the DriveSafe DriveAware conversion (Findlater et al., 2013). Consistent with other research (Cockburn et al., 2012; Findlater et al., 2013; Kobayashi et al., 2011), participants found tap intuitive and drag more difficult. However, participants were not precise when entering responses via either action. This may be due to parallax (where object position appears to differ when viewed from different angles), because the target is obscured by the finger, or due to the additional dexterity and pressure required to "hold" and slide the target against resistance (Cockburn et al., 2012; Findlater et al., 2013; Kobayashi et al., 2011; Loureiro & Rodrigues, 2014). The arrow design that avoided the need for a precise dragging action worked better than the other options trialled. We considered this lack of precision when determining scoring in a subsequent study, allowing generous zones to be scored as correct. Consistent with design guidelines for older adults (Loureiro & Rodrigues, 2014), we found timing of interface elements critical to smooth progression of the test. For example, participants encountered difficulties in Round 1 because the "next" button appeared before it was needed, resulting in over-focus on this button and misapplication.

Cueing timely attention to the small iPad screen was fundamental in the conversion because participants determined when to progress to the next item. In touchscreen DriveSafe DriveAware, a loud bell and countdown timer successfully cued attention via auditory and visual prompts. We did not standardise distance to the screen based on touchscreen design guidelines for older adults which recommended seniors be free to adjust the iPad distance for comfortable viewing (Loureiro & Rodrigues, 2014). This worked well in practice. Participants did not have difficulty observing the smaller hazards, which varied in size from .04" (10 mm) to .24" (62 mm).

We tested all text and function changes across age groups as any change significantly impacted performance. The middle-aged cohort reported different challenges to the older age groups (i.e., with forced undo practice). Round-3 participants did not have any difficulty with test set up, following instructions, or with touchscreen operations. This is considered related to their younger age, familiarity with digital technology compared to the older adults, and because they were trialling the third iteration of DSDA. Inclusion of participants with challenges common to older age, such as hearing and vision loss, tremors, and potential cognitive changes, avoided user-interface problems that might have occurred if we had considered only the needs of able-bodied users. The performance of participants with these challenges informed test set-up and administration procedures and identified the need for an examiner-assisted version of administration.

Limitations & Future Implications. A limitation of this study was the small sample size, which may have resulted in low probability errors not being detected (Hegde, 2013). However, we considered a range of other sources to mitigate this: literature review, task analysis, prototyping, interviews, expert reviews, and continuous quality checking. A further study was conducted to develop an automatic data scoring system to reflect the decisions that would otherwise have been made by an expert-rater (Cheal, Bundy, Patomella, Scanlan, & Wilson, 2018). Additionally, a further study was conducted to examine the psychometric properties and predictive validity of data gathered with touchscreen DriveSafe DriveAware before it could be used in clinical practice (Cheal & Kuang, 2015).

Conclusion. A user-centred design process allowed us to develop a user-friendly touchscreen data collection system for older adults, the group most likely to take touchscreen DriveSafe DriveAware and most likely to struggle with the technology. The approach taken allowed us to test and validate our design assumptions and ensured design errors were not missed. We believe that conducting usability testing concurrently with iPad application design, programming and evaluation resulted in significant cost, time, and efficiency benefits.

Acknowledgements

We acknowledge the contributions of the iPad application developer Very Livingstone.

Declaration of interest

This project was supported by Pearson, who publishes DriveSafe DriveAware. The first author was employed by Pearson to project manage the conversion to touchscreen. The second author is a DriveSafe DriveAware author and receives royalties from sale of DriveSafe DriveAware.

References

Al-Hassani, S. B., & Alotaibi, N. M. (2014). The impact of driving cessation on older Kuwaiti adults: Implications to occupational therapy. *Occupational Therapy in Health Care, 28*(3), 264-276. doi:10.3109/07380577.2014.917779

Austroads, & National Transport Commission. (2016). *Assessing Fitness to Drive for Commercial and Private Vehicle Drivers*. Sydney, Australia: Austroads.

Barco, P. P., Baum, C. M., Ott, B. R., Ice, S., Johnson, A., Wallendorf, M., & Carr, D. B. (2015). Driving errors in persons with dementia. *Journal of American Geriatric Society, 63*(7), 1373-1380. doi:10.1111/jgs.13508

Brown, F., Cheal, B., Cooper, M., & Joshua, N. (2013). *DriveSafe DriveAware Digital Development: General Practitioner Survey Findings*. Sydney, Australia: Pearson Clinical Assessment Australia & New Zealand.

Charlton, J., Sjaanie, K., Odell, M., Devlin, A., Langford, J., O'Hare, M., . . . Scully, M. (2010). *Influence of Chronic Illness on Crash Involvement of Motor Vehilce Drivers* (2nd ed.). Melbourne, Australia: Monash University Accident Research Centre.

Cheal, B., & Kuang, H. (2015). *DriveSafe DriveAware for Touch Screen Administration Manual.* Sydney, Australia: Pearson Australia Group Pty Ltd.

Cheal, B., Bundy, A., Patomella, A., Scanlan, J. N., & Wilson, C. (2018). Converting the DriveSafe subtest of DriveSafe DriveAware for touchscreen administration. Australian Occupational Therapy Journal, (in press). doi:10.111/1440-1630.12558

Cockburn, A., Ahlström, D., & Gutwin, C. (2012). Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies, 70*(3), 218-233. doi:10.1016/j.ijhcs.2011.11.002

Coxon, K., & Keay, L. (2015). Behind the wheel: Community consultation informs adaptation of safe-transport program for older drivers. *BMC Research Notes, 8*, 764. doi:10.1186/s13104-015-1745-0

Delello, J. A., & McWhorter, R. R. (2016). Reducing the Digital Divide. *Journal of Applied Gerontology, 36*(1), 3-28. doi:10.1177/0733464815589985

Dickerson, A. E., Reistetter, T., & Gaudy, J. R. (2012). The perception of meaningfulness and performance of instrumental activities of daily living from the perspectives of the medically at-risk older adults and their caregivers. *Journal of Applied Gerontology, 32*(6), 749-764. doi:10.1177/0733464811432455

Dobbs, A. R., Heller, R. B., & Schopflocher, D. (1998). A comparitive approach to identify unsafe older driver. *Accident Analysis & Prevention, 30*(3), 363-370.

Dorrington, P., Wilkinson, C., Tasker, L., & Walters, A. (2016). User-centered design method for the design of assistive switch devices to improve user experience, accessibility, and independence. *Journal of Usability Studies, 11*(2), 66-82.

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

Findlater, L., Froehlich, J. E., Fattal, K., Wobbrock, J. O., & Dastyar, T. (2013). *Age-Related Differences in Performance with Touchscreens Compared to Traditional Mouse*

Input. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI) 2013, Paris, France.

Fricke, J., & Unsworth, C. (2001). Time use and importance of instrumental activities of daily living. *Australian Occupational Therapy Journal, 48*(3), 118-131.

Hegde, V. (2013). *Role of human factors / usability engineering in medical device design.* Paper presented at the Proceedings Annual Reliability and Maintainability Symposium (RMAS), Orlando, FL.

Hines, A., & Bundy, A. (2014). Predicting driving ability using DriveSafe and DriveAware in people with cognitive impairments: A replication study. *Australian Occupational Therapy Journal, 61*(4), 224-229. doi:10.1111/1440-1630.12112

International Organization for Standardization. (2010). ISO 9241-210: 2010 Ergonomics of Human System Interaction - Part 210: Human-centered Design for Interactive Systems. Genève, Switzerland: International Organization for Standardization.

Jang, R. W., Man-Son-Hing, M., Molnar, F. J., Hogan, D. B., Marshall, S. C., Auger, J., . . . Naglie, G. (2007). Family physicians' attitudes and practices regarding assessments of medical fitness to drive in older persons. *Journal of General Internal Medicine, 22*(4), 531-543. doi:10.1007/s11606-006-0043-x

Kay, L., Bundy, A., & Clemson, L. (2008). Predicting fitness to drive using the Visual Recognition Slide Test (USyd). *American Journal of Occupational Therapy, 62*, 187-197.

Kay, L., Bundy, A., & Clemson, L. (2009a). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation, 90*(9), 1514-1522. doi:10.1016/j.apmr.2009.03.011

Kay, L., Bundy, A., & Clemson, L. (2009b). Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation, 31*(13), 1074-1082. doi:10.1080/09638280802509553

Kay, L., Bundy, A., Clemson, L., Cheal, B., & Glendenning, T. (2012). Contribution of off-road tests to predicting on-road performance: A critical review of tests. *Australian Occupational Therapy Journal, 59*(1), 89-97. doi:10.1111/j.1440-1630.2011.00989.x

Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., & Ifukube, T. (2011). *Elderly User Evaluation of Mobile Touchscreen Interactions.* Paper presented at the INTERACT 2011, Lisbon, Portugal.

Kostyniuk, L. P., & Shope, J. T. (2003). Driving and alternatives: Older drivers in Michigan. *Journal of Safety Research, 34*(4), 407-414. doi:10.1016/j.jsr.2003.09.001

Leng, F. Y., Yeo, D., George, S., & Barr, C. (2014). Comparison of iPad applications with traditional activities using person-centred care approach: Impact on well-being for persons with dementia. *Dementia, 13*(2), 265-273. doi:10.1177/1471301213494514

Loureiro, B., & Rodrigues, R. (2014). *Design guidelines and design recommendations of multi-touch interfaces for elders*. Paper presented at the ACHI 2014: The Seventh International Conference on Advances in Computer-Human Interactions, Barcelona, Spain.

Marshall, S. C. (2008). The role of reduced fitness to drive due to medical impairments in explaining crashes involving older drivers. *Traffic Injury Prevention, 9*(4), 291-298. doi:10.1080/15389580801895244

Marshall, S. C., Demmings, E. M., Woolnough, A., Salim, D., & Man-Son-Hing, M. (2012). Determining fitness to drive in older persons: A survey of medical and surgical specialists. *Canadian Geriatrics Journal, 15*(4), 101-119. doi:10.5770/cgj.15.30

Marshall, S. C., & Man-Son-Hing, M. (2011). Multiple chronic medical conditions and associated driving risk: A systematic review. *Traffic Injury Prevention, 12*(2), 142-148. doi:10.1080/15389588.2010.551225

Marshall, S. C., Molnar, F., Man-Son-Hing, M., Blair, R., Brosseau, L., Finestone, H. M., . . . Wilson, K. G. (2007). Predictors of driving ability following stroke: A systematic review. *Topics in Stroke Rehabilitation, 14*(1), 98-114. doi:10.1310/tsr1401-98

McCurdie, T., Taneva, S., Casselman, M., Yeung, M., McDaniel, C., Ho, W., & Cafazzo, J. (2012). mHealth consumer apps the case for user-centered design. *Horizons, Fall 2012*, 49-54.

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 American Geriatric Society, 54*(12), 1809-1824. doi:10.1111/j.1532-5415.2006.00967.x

Moussa, Y., Mahdanian, A. A., Yu, C., Segal, M., Looper, K. J., Vahia, I. V., & Rej, S. (2017). Mobile health technology in late-life mental illness: A focused literature review. *American Journal of Geriatric Psychiatry, 25*(8), 865-872. doi:10.1016/j.jagp.2017.04.003

O'Donnell, J. M., Morgan, M. K., & Manuguerra, M. (2018). Patient functional outcomes and quality of life after microsurgical clipping for unruptured intracranial aneurysm. A prospective cohort study. *Journal of Neurosurgery, Published online March 2, 2018*. doi:10.3171/2017.8.JNS171576

Onoda, K., Hamano, T., Nabika, Y., Aoyama, A., Takayoshi, H., Nakagawa, T., . . . Yamaguchi, S. (2013). Validation of a new mass screening tool for cognitive impairment: Cognitive Assessment for Dementia, iPad version. *Clinical Intervention in Aging, 8*, 353-360. doi:10.2147/CIA.S42342

Østensen, E., Gjevjon, E. R., Øderud, T., & Moen, A. (2017). Introducing Technology for Thriving in Residential Long-Term Care. *Journal of Nursing Scholarship, 49*(1), 44-53. doi:10.1111/jnu.12268

Papa, M., Boccardi, V., Prestano, R., Angellotti, E., Desiderio, M., Marano, L., . . . Paolisso, G. (2014). Comorbidities and crash involvement among younger and older drivers. *PLoS One, 9*(4), e94564. doi:10.1371/journal.pone.0094564

Rogers, W. A., & Mitzner, T. L. (2016). Envisioning the future for older adults: Autonomy, health, well-being, and social connectedness with technology support. *Futures*. doi:10.1016/j.futures.2016.07.002

Schneider, N., Wilkes, J., Grandt, M., & Schlick, C. (2008). *Investigation of input devices for the age-differentiated design of human-computer interaction.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting - 2008, New York.

Sims, J., Rouse-Watson, P., Schattner, A., Beveridge, A., & Jones, K. (2012). To drive or not to drive: Assessment dilemmas for GPs. *International Journal of Family Medicine, 2012*, 1-6. doi:10.1155/2012/417512

Tsai, H., Shillair, R., Cotten, S. R., Winstead, V., & Yost, E. (2015). Getting Grandma Online: Are Tablets the Answer for Increasing Digital Inclusion for Older Adults. *Educational Gerontology, 41*(10), 695–709. doi:10.1080/03601277.2015.1048165]

Watkins, I., & Xie, B. (2015). Older adults' perceptions of using iPads for improving fruit and vegetable intake: An exploratory study. *Care Management Journals, 16*(1), 2-13. doi:10.1891/1521-0987.16.1.2

Wilson, L. R., & Kirby, N. H. (2008). Individual differences in South Australian general practitioners' knowledge, procedures and opinions of the assessment of older drivers.

Australasian Journal on Ageing, 27(3), 121-125. doi:10.1111/j.1741-6612.2008.00304.x

Woolnough, A., Salim, D., Marshall, S. C., Weegar, K., Porter, M. M., Rapoport, M. J., . . . Vrkljan, B. (2013). Determining the validity of the AMA guide: A historical cohort analysis of the assessment of driving related skills and crash rate among older drivers. *Accident Analysis & Prevention, 61*, 311-316. doi:10.1016/j.aap.2013.03.020

Yale, S., Hansotia, P., Knapp, D., & Ehrfurth, J. (2003). Neurologic conditions: Assessing medical fitness to drive. *Clinical Medicine & Research, 1*(3), 177-188.

CHAPTER 6

Journal Article 2

Converting the DriveSafe subtest of DriveSafe DriveAware for Touchscreen Administration

The following study was conducted to determine if the automatic data collection and scoring system we designed represented the decision that would otherwise have been made by a trained administrator, and if the scoring parameters we set discriminated between at-risk and comparison drivers. The present manuscript was accepted for publication by the *Australian Journal of Occupational Therapy* on 23.10.2018.

Statement of Contribution of Author

On behalf of the co-authors of "Converting the DriveSafe subtest of DriveSafe DriveAware for touchscreen administration", we confirm that Beth Cheal has made the following contributions:

1. Conceptualisation and design of the research
2. Collection of data
3. Analysis and interpretation of findings
4. Writing the paper and critical appraisal of content

Anita Bundy Signed: ____________________ Date: 04.09.2018

Beth Cheal Signed: ____________________ Date: 04.09.2018

Converting the DriveSafe subtest of DriveSafe DriveAware for Touchscreen Administration

Beth Cheal, MAppSc(OT)[a,c], Anita Bundy, ScD[a,d], Ann-Helen Patomella, PhD[b], Justin Newton Scanlan, PhD[a], Christopher Wilson, BHealthSc, BA(Psych) (Hons)[c,e]

[a]*University of Sydney, Sydney, Australia;* [b]*Karolinska Institutet, Stockholm, Sweden;* [c]*Pearson, Sydney, Australia;* [d]*Colorado State University, Fort Collins, US;* [e]*The University of Melbourne, Melbourne, Australia.*

Abstract

Background. DriveSafe measures awareness of the driving environment. It is one subtest of DriveSafe DriveAware, a cognitive fitness-to-drive screening instrument. We converted DriveSafe to a touchscreen format for ease of administration; this necessitated development of an automatic variable data collection and scoring system to reflect the decisions that would otherwise have been made by an expert rater. We applied a structured process to determine what constituted "correct" scores. We then examined the resulting scoring parameters to determine if these discriminated at-risk drivers from a comparison sample.

Methods. Thirty older drivers referred for a fitness-to-drive assessment, identified as 'at-risk', and 30 comparison drivers took touchscreen DriveSafe. Following presentation of images containing between two and four objects / hazards for 4 seconds, participants indicated their recall of object / hazard characteristics (type, location, and direction of movement) by touching the screen. We analysed responses via descriptive statistics to compare spread, accuracy and consistency; and via a Fisher's exact test to determine whether the set scoring parameters could discriminate between at-risk and comparison drivers.

Results. Fisher's exact test results indicated 24 of 28 location zones and 18 of 28 direction ranges discriminated at-risk drivers from the comparison group (p<0.05). Frequency of missed or incorrectly identified hazards was much higher for the at-risk group for all variables. At-risk drivers missed or misidentified significantly (p<0.00) more object types (34%). directions (47%) and locations (36%) than the comparison group. The comparison group missed ≤4% for each variable. At-risk drivers entered 31 additional responses for objects / hazards not displayed; the comparison group entered no additional responses.

Conclusion. The automatic variable data collection and scoring system reflected decisions that would have been made by an expert rater. This systematic process provided automated scoring decisions that enabled us to discriminate at-risk drivers from a comparison group. Psychometric evaluation of data gathered with touchscreen DriveSafe is required prior to use in clinical practice.

Keywords

Cognitive Assessment Screening Instrument; Automobile Driving; Computer, Handheld

Background

Occupational therapists require valid and reliable measures to inform clinical practice, service access, funding decisions, policy, and research in healthcare. Standardised measures are increasingly converted to touchscreen due to time and accuracy advantages for clinicians

(Daly, Xu, & Levy, 2015; Downing Peck, 2011; Glaser, Jain, & Kortum, 2013; Godwin, Tan, Bockhold, Ma, & Tran, 2015; Howell et al., 2015; Turney & Reynard, 2014; Wiarda, McMinn, Peterson, & Gregor, 2014). However, there is a risk that ability of the test to capture the underlying construct may be lost through the conversion process, affecting validity and reliability. Significant time and financial resources are typically dedicated to digital test conversion. Therefore, it is critical the process is conducted in a way that will maintain ability of the test to record and measure variation in the underlying construct.

DriveSafe DriveAware (DSDA) (Kay et al., 2009a) is a clinical screen of cognitive fitness to drive that has been used by driver-trained occupational therapists in Australia for many years. The DSDA total score classifies participants into one of three categories based on likelihood of passing an on-road assessment: “Likely to pass”, “Requires further testing”, and “Likely to fail.” Administration of the original version of DSDA is limited to driver-trained occupational therapists because training is required to interpret participants’ verbal responses. Occupational therapists who are not specialists in driving are frequently asked to address driving with their clients: advising medical teams if further assessment is required; advising clients and family regarding community mobility options; and providing support in cessation of driving (Dickerson, 2014a, 2014b; Korner-Bitensky et al., 2007b). Occupational therapists need a suitable cognitive fitness-to-drive screen so they can give reliable advice in this high-stakes area and clients need not undergo extensive cognitive testing unsuitable for predicting driving performance (Dickerson, 2014b; Vrkljan et al., 2011). We converted original DSDA into a first-level screen that could be administered via touchscreen by occupational therapists and other health professionals without training to determine the need for on-road assessment.

DriveSafe Subtest. The focus of the present paper is on touchscreen conversion of the DSDA subtest, DriveSafe, because this was the most complex of the two DSDA subtests to convert. DriveSafe measures the construct of awareness of the driving environment (Kay, Bundy, & Clemson, 2008; Kay et al., 2009a) via visual search and assumes that attention is critical to safe driving. A driver scanning the driving environment must rapidly and accurately process simultaneous information critical for safe decision-making (Hoffman et al., 2006; Rizzo et al., 2009). DriveSafe utilises mechanisms identified by researchers (Caird et al., 2005; Crundall, 2009; Galpin et al., 2009; Jensen et al., 2011; Koustanaï et al., 2012; O'Regan et al., 1999; Rensink, 2002; Rensink et al., 1997) as necessary for assessment of driver visual attention: use of naturalistic driving scenes and stimuli, a driving related task, masking of object change, and a time-limited, one-time display of objects.

We sought to make touchscreen DriveSafe as similar as possible to original DriveSafe in order to be sure we retained accurate representation of the construct being measured. However, we were transitioning from a test where scores were assigned based on verbal response interpreted by an expert rater, to a test where touchscreen responses are scored automatically in-app. This necessitated development of an automatic variable data collection and scoring system that would reflect the decisions that would otherwise have been made by an expert rater. We applied a structured process (Aparasu, 2011; Summers, 1970) for the measurement of abstract constructs such as behaviour to guide this development, operationalised via the following interlinked steps: Step 1) identification of acceptable variables that reflect the underlying construct; Step 2) data collection including development of data collection rules; and Step 3) assignment of values to data with respect to a specified scale. The final step, evaluation of the validity and reliability of data gathered with touchscreen DriveSafe, was completed in a subsequent study.

Previous researchers (Kay, Bundy, & Clemson, 2008; Kay et al., 2009a) had identified the original DriveSafe variables (Step 1) and provided evidence of construct validity, internal reliability (Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2008; Kay et al., 2009a) and test-retest reliability (O'Donnell et al., 2018). Original DriveSafe items require recalling characteristics of objects presented for 4 seconds in each of 10 driving images. The characteristics represent four variables: a) object type (e.g., car); b) object location (e.g., road); c) object side of screen (e.g., right); and d) object direction of movement (e.g., going left) (Kay et al., 2009a). The present study concerns Steps 2 and 3: how to collect data with respect to the four variables via touchscreen, including development of data collection rules; and how to assign values based on these rules with respect to a scale specified in original DriveSafe. This scale awarded 1 point for each piece of information for each object, correctly recalled (Kay et al., 2009a).

We addressed the research question "How generous should scoring parameters be when awarding points for the touchscreen DriveSafe object location and direction of movement variables?" (a sub-process for achieving Step 3). We sought to answer this question by examining touchscreen responses entered by at-risk older drivers and a comparison group. Our hypothesis was that at-risk drivers would be more likely than the comparison group to incorrectly identify object variables in the driving environment.

Method

The University of Sydney Human Research Ethics Committee provided approval for this study. In addition, the St Vincent's Hospital Sydney Human Research Ethics Committee provided ethics approval for the hospital sites.

Setting

We employed a prospective design over a 2-month period. We administered touchscreen DriveSafe to two groups: a) older adults referred to eight driving clinics in Australia and New Zealand; and b) adult volunteers recruited from a community centre and office in Sydney, Australia.

Participants

Group 1 comprised 30 older adults (*n*=8 female) referred for a driving assessment to determine the impact of a medical condition on fitness to drive; 16 participated in New Zealand and 14 in Australia. Diagnoses included dementia (*n*=12); stroke (*n*=6); cardiovascular conditions (*n*=4); Parkinson's disease (*n*=3); traumatic brain injury (*n*=2); and, other neurological conditions (*n*=3). Participants were aged 57-89 years (M_{age}=78; *SD*=9.2), with a mean Mini-Mental State Examination 2: Standard Version (MMSE2:SV) (Folstein, Folstein, White, & Messer, 2010) score of 24/30 (range=17-30; *SD*=2.8): a 30-point standardised assessment of cognition interchangeable with Mini-Mental State Examination (MMSE) (Folstein et al., 1975; Folstein et al., 2010).

We invited consecutive participants who met the inclusion criteria: a valid driver's license; vision within license authority guidelines; completion of at least 1 year of high school; English as a first language; age >55 years; and drivers identified as at-risk. 'At-risk' was operationalised as missing four or more objects in touchscreen DriveSafe (completed in the clinical assessment of fitness to drive prior to on-road testing) and/or on-road assessment results indicating failure to meet license authority criteria for safe driving. (Group 1 participants completed a standardised on-road assessment as part of a separate study. Results were used to identify at-risk drivers for this study). Six participants received a pass or conditional pass in the on-road assessment but were included because they missed a high number of objects (*M*=7/28) and details (*M*=17/56) in DriveSafe. We excluded potential participants with a psychiatric illness, because previous experience suggests that people diagnosed with a psychiatric illness can do well in DriveSafe but perform poorly on-road due to disturbances of mood, behaviour, or perception. We considered these to be potential confounders. We also excluded participants with a developmental delay or aphasia. No participants declined to participate.

Group 2 comprised 30 adult volunteers from a community centre. We recruited participants via advertisement in community meeting areas asking for volunteers to try a new iPad application to assist in the development of a fitness-to-drive screen. We informed potential volunteers that any information provided would be anonymous and we would give no advice regarding their driving. We collected no identifying information. Once volunteers contacted the first author or centre staff, we made an appointment for them to complete touchscreen DriveSafe. Group 2 participant age ranges in years were 20-29 (*n*=7); 30-39 (*n*=7); 40-49 (*n*=12); 50-59 (*n*=1); and 60-65 (*n*=3). Most (*n*=27) were employed full-time: clerical/administration (*n*=12); health professions (*n*=10); education (*n*=4); and electrician (*n*=1); 3 were retired. All were licensed drivers. Participants were not aged match because the focus of the study was test development. Therefore, we wanted one group of drivers who represented those who were most likely to take the test (older drivers with impairments who may struggle with the technology) and one group to represented young drivers without impairment who were likely not at-risk. We tested every participant only once. No one withdrew from the study after agreeing to participate. We removed data from 2 comparison group participants: 1 due to a high number of missing or random responses and 1 due to duplication of results attributed to a programming error identified following completion of data collection. To include these responses would have significantly impacted data validity. As data were de-identified from the point of collection, we could not identify which participants' data had been removed.

Instrument – Touchscreen DriveSafe

Touchscreen DriveSafe is self-administered via iPad with written and audio instructions and consists of 10 images of a 4-way intersection. Each image includes two to four potential hazards (a total of 28). These hazards are presented for 4 seconds then disappear, leaving only the blank intersection. Examinees are asked to recall the hazards previously displayed, touching the intersection to identify (a) object location, (b) object type, and (c) direction of object movement. Object location is indicated by touching the intersection. Object type is indicated by touching an object from a menu of five objects (i.e., pedestrian, two people walking together, bike, car, or truck). Object direction is indicated by adjusting an arrow to represent the direction the object was travelling (objects are static). Touchscreen responses are visible to the participant and can be adjusted until the participant initiates the next item. Each piece of correctly recalled information yields 1 point. The maximum score is 84: 28 for each scoring category (object type, location and direction).

Procedure

At-risk older adults self-administered touchscreen DriveSafe. The occupational therapist provided assistance, if required, without prompting responses. Group 2 participants self-administered touchscreen DriveSafe unassisted. All participants took touchscreen DriveSafe using an iPad (iOS 9) with a 9.7-inch, 2048 x 1536 mm, multi-touch "retina" display. Headphones and a stylus were available to use.

Analysis

Data were collected via touchscreen then analysed as described below.

Object Location. We collected touch point data to answer the question how to collect data with respect to the two original DriveSafe object location variables (i.e., object side of screen and location). We collapsed these two categories because a single touch point comprises both variables. Two sets of data were required to plot values for a single touch point: x-axis and y-axis locations in pixels (Px). We created a Microsoft Excel x-y scatter chart for each item separately for each group, allowing visual inspection of responses. We assigned each touch point a shape and colour for clearer visualisation. The symbols were consistent across object type (e.g., bike = purple circle).

We created zones around each object to determine how to assign values for object location, so that responses placed within the zone could be scored 1 and missing responses or responses outside the zone could be scored 0. Zones were unique for each object due to variation in object type, size, and location. We examined spread, consistency, and accuracy of responses between groups via descriptive statistics (*M*, *SD*, and range) for each axis to inform decisions regarding how generous zones should be. Because we combined two variables (object side and location), the total score for the location variable was 28 (1 point for each correctly recalled object location); this compared to 56 in original DriveSafe. Once correct zones were generated, we examined results of a Fisher's exact test for each zone to determine whether the set parameters discriminated at-risk from comparison drivers to test our hypothesis that at-risk drivers would be more likely than the comparison group to incorrectly identify object location ($p<0.05$).

Object Type. We assigned numbers to each object type (e.g., car = 4) so we could code and match responses. We scored objects matching the displayed object 1 and missing or non-matching responses 0 (Step 3). We examined accuracy and consistency of total scores between groups via descriptive statistics: mean (*M*), range, and standard deviation (*SD*).

Object Direction. We determined correct direction ranges for each object so responses that fell within the range could be scored 1 and missing responses or responses outside the range could be scored 0. Ranges were unique to each object due to the wide variation in object type, size, location and trajectory (objects were static). We compared spread, consistence, and accuracy of responses between groups for objects via descriptive statistics (*M*, *SD* and range) to inform decisions regarding how generous correct direction ranges should be. Once ranges were set, we examined results of a Fisher's exact test to determine whether each range discriminated at-risk from comparison drivers to test our hypothesis that at-risk drivers would be more likely than the comparison group to incorrectly identify object direction ($p<0.05$).

Results

Following are the results for the object location, type, and direction variables as determined in the process of establishing scoring parameters. We divided the results into three sections reflecting the main scoring categories: object location, object type, and object direction.

Object Location

At-risk drivers entered a response (regardless if correct) on average for 78% of responses (range=13-35/28; *SD*=6.1); comparison drivers entered a response (regardless if correct) for 96% of responses (range=21-28/28; *SD*=2.0). (Note that additional responses were recorded for all variables).

Data collection rules (Step 2). We developed the following data collection rules to determine how to collect data with respect to the object location variable. We scored location as an independent variable similarly to original DriveSafe because we judged that participants who placed an incorrect object at a correct location should receive more credit than someone who did not respond. The participants who responded had at least identified that something was present at the correct location that a driver should avoid. Therefore, we scored location data based on whether responses fell within a specified correct zone, regardless of object type data entered. We allowed a maximum of three responses additional to the number of displayed objects per image. We recorded and reported additional responses to provide therapists with qualitative information.

Assignment of values: correct location zone (Step 3). We needed to generate correct location zones to assign values based on the data collection rules developed: so we could award 1 point for each correct location response within the zone and 0 points for missing responses or responses outside the zone. Analysis of the scatter charts indicated object surface area should be included within zones because participant responses covered this area although not representing the correct location. For example, the truck surface area in Figure 1 covers the footpath, park, and house (see Figure 1). Participants from both groups placed location points over these areas once the truck had disappeared. We assumed that participants did not believe the truck was driving on the footpath or grass but rather the target surface area was important to their perception of locality under time pressure. Therefore, the truck surface area was included in the zone (see Figure 1B). Aside from object surface area, we excluded any locations unrelated to the driving context (e.g., trees or sky).

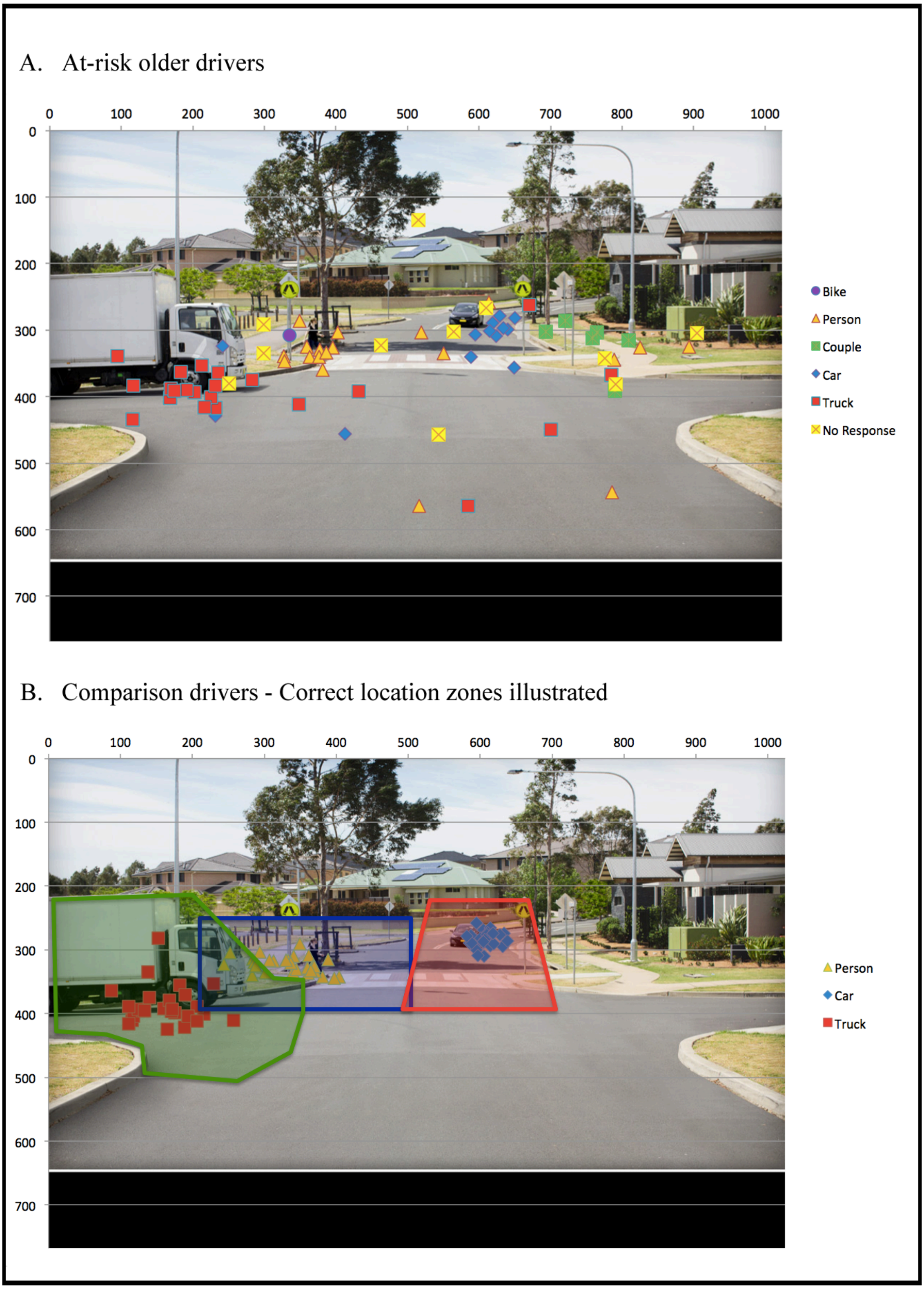


Figure 1. Comparison of Item 2 location responses between groups

Note. Yellow box = no response for object type and direction

We compared location range and *SD* data between groups for both axes to determine the ideal width and height of correct zones. Final zone parameters also reflected our practical experience with original DriveSafe. For example, the red zone around the car in Figure 1B includes the road immediately in front of the car to reflect the previously broad original DriveSafe category, where the entire street would have been accepted as a correct response. Older adult (Group 1) responses demonstrated a substantially greater spread and variation of data across both axes for all objects. However, both groups demonstrated a greater spread across the horizontal when compared to the vertical for all objects (see example in Figures 1 and 2).

Figure 2. Comparison of Item 8 location responses between groups

Note. Yellow box = no response for object type and direction

Examination of the range, *SD*, and scatter charts, indicated the comparison group was more accurate in target location for all objects. In contrast to the at-risk group, comparison driver responses were within 200 Px vertically and horizontally for 25 of 28 objects. Zones were established based largely on the comparison group's correct responses but with more generous parameters to accommodate the expected reduced accuracy of older adult responses.

Determining if zones discriminated between drivers. Results of the Fisher's exact test indicated scoring parameters discriminated at-risk from comparison drivers for 24 of 28 location zones ($p<0.05$). This means that at-risk drivers were more likely to incorrectly identify object location compared to the comparison group for 86% of the objects. There were no between group differences for the following objects: pedestrians on the left and right in item 1 ($p=0.37$ and 0.39 respectively), and the car and person in item 8 ($p=0.06$ and 0.08 respectively) (item 8 hazards are pictured in Figure 2).

The total mean score for at-risk drivers for location was 18/28 (range=6-24; *SD*=4.6) compared to 27.5/28 for Group 2 (range=25-28; *SD*=0.7) (see Table 1). At-risk drivers missed 19% of locations presented (*M*=5.2 locations per person; range=0-15; *SD*=4.0) and misidentified 17% (*M*=4.8; range=1-11; *SD*=2.4). Thus, at-risk drivers missed or misidentified 36% of locations (*M*=10.0; range=3-22; *SD*=4.6). At-risk drivers entered 13 locations not displayed (*M*=0.4; *SD*=0.9). The comparison group missed 2% of locations (*M*=0.5; *SD*=0.7), misidentified 2% (*M*=<0.1; *SD*=0.2), and entered no additional responses. Results of a Kolmogorov–Smirnov test of normality indicated the data were skewed. Therefore, a Mann-Whitney U test was performed with a Bonferroni correction to determine differences between groups. Results indicated a significant difference in ability to recall object locations between groups ($p<0.00$). These results indicated at-risk drivers found recall of object location more difficult than the comparison group for all objects and most of the zones we generated were able to discriminate between drivers in the at-risk category.

Table 1

Difference in recall of object details between at-risk and comparison drivers for the three DriveSafe variables (object location, type and direction).

	Missed %	Misidentified %	Missed & Incorrect %	Additional Responses	Total Mean Score /28
Object Location					
At-risk	19	17	36	13	18.0
Group 2	2	2	4	0	27.5
Object Type					
At-risk	22	12	34	10	18.4
Group 2	2	2	4	0	26.9
Object Direction					
At-risk	30	17	47	8	15.0
Group 2	2	1	3	0	27.2

Object Type

Assignment of values to the object type variable was straightforward because the data were nominal and the scoring method was the same as for original DriveSafe (i.e. participant response were matched with the displayed object then scored).

Data collection rules (Step 2). We coded then matched objects, scoring object type independently of object location, similarly to original DriveSafe. This was because our practical experience with original DriveSafe indicated at-risk drivers found object type easiest to recall and direction of movement the most difficult. We wanted to evaluate performance in the categories separately. We allowed participants to initiate the next item without indicating object type because the opportunity to 'forget' to respond contributed to scoring and served to discriminate among individuals.

Assignment of values (Step 3). We assigned values based on the data collection rules developed by awarding 1 point for each selected object matching the displayed object and 0 points for mismatched objects or missing responses. We awarded points up to the maximum number of objects displayed. Each participant received a total score out of 28, representing the total number of correctly recalled object types. We examined responses across groups to ensure data collection and assignment of values reflected the specified scale and allowed the expected variation in performance between groups to be captured.

Consistent with expectations, at-risk older drivers missed or misidentified significantly more objects than the comparison group. Results of a Kolmogorov–Smirnov test of normality indicated the data were skewed. Therefore, a Mann-Whitney U test was performed with a Bonferroni correction. Results indicated a significant difference ($p<0.00$). At-risk drivers missed 22% of object types (M=6.2 object types per person; SD=3.9) and misidentified 12% (M=3.4; SD=1.7) (see Table 1). Thus, at-risk drivers missed or misidentified 34% of object types (M=9.6; range=2-18; SD=4.2) compared to 4% (M=1.1; range=0-4; SD=1.3) for Group 2. At-risk drivers entered 10 additional responses for objects not displayed (M=0.3; SD=0.9); the comparison group entered no additional responses. The at-risk drivers' total mean score for the object type category was 18.4/28 (range=10-26; SD=4.0) compared to 26.9/28 (range=24-28; SD=1.3) for Group 2.

Only 3 of the 30 at-risk drivers correctly identified all 3 bikes; 10 correctly identified both trucks. At-risk drivers missed or misidentified almost half (48%) of all bikes (M=1.4/3; SD=0.8) and trucks (47%) (M=0.9/2; SD=0.8), 36% of single pedestrians (M=4.7/13; range=0-12; SD=3.1), 25% of cars (M=2.0/8; range=0-7; SD=1.2), and 21% of pedestrian couples (*M*=0.6/3; *SD*=1.0/3). Twenty-one of 28 comparison drivers correctly identified both trucks; 7 missed one and 1 missed both (*M*=0.3/2; *SD*=0.5). Five comparison drivers missed or misidentified one of the 3 bikes (*M*=0.2/3; *SD*=0.4). Comparison drivers missed 1% of cars (*M*=0.1/8; range=0 - 1; *SD*=0.3), 1% of single pedestrians (*M*=0.2/13; range=0-2; *SD*=0.5), and 1% of pedestrian couples (*M*=<0.1/3; *SD*=0.2). Trucks represented 39% of missed or misidentified objects for the comparison group although only two were displayed; bikes represented 22%, although only three were displayed. In comparison, pedestrians made up 55% of objects missed or misidentified by at-risk older drivers (57% of objects displayed), cars 20% (29% of objects displayed), and bikes and trucks 25%. At-risk drivers missed (i.e., failed to enter a response for any variable) 18% of objects displayed (*M*=5/28; range=0-16; *SD*=4.2); comparison drivers missed 2% (*M*=0.5/28; range=0-4; *SD*=<0.1). At-risk drivers entered 10 additional responses for objects not displayed; comparison drivers entered no additional responses. These results indicate at-risk drivers found recall of object type more challenging than the comparison group for all hazards.

Object Direction

At-risk drivers entered a response (regardless if correct) on average for 62% of direction responses (range=9-31/28; *SD*=5.44). Average response entry was much higher for the comparison group (*M*=97%; range=22-28/28; *SD*=1.46).

Data collection rules (Step 2). We set the following data collection rules to answer the question how to collect data with respect to the object direction variable. We scored object direction independently of object type and location because we considered that participants who noticed something travelling in the correct direction but failed to identify the object type or specific location, demonstrated greater awareness than someone who did not enter a response. Similar to the object type variable, we made it possible to forget to enter a response and progress to the next item.

Assignment of values: correct direction ranges (Step 3). To determine how to assign values with respect to the original DriveSafe scale, we determined correct degree ranges for each object so we could score directions that fell within the range 1 and missing responses or directions outside the range 0. We scored directions correct up to the maximum number of directions displayed. The total maximum score for the direction variable was 28 of 28.

We determined how generous correct direction ranges should be (scoring parameters) by examining range, mean, and *SD* across groups. Correct ranges varied for each object depending on trajectory, location, size, and shape. To enable meaningful analysis of participants' perceptions of direction related to specific objects when generating ranges, we combined object type and direction variables for this aspect of the study. We also removed incorrect responses. For example, we removed 6 at-risk driver responses ranging from 37° to 141° from Item 8 for the person on the left walking left (correct answer 270°) (see Figure 2). To retain these responses would have resulted in incorrect scoring and large correct direction ranges that did not accurately reflect the perception of unimpaired participants, reducing the possibility that the variables would discriminate between individuals based on their awareness of the driving environment. We established ranges largely on the comparison group's correct responses but with more generous parameters to accommodate the expected reduced accuracy of older adults.

Determining if set ranges discriminate between drivers. Once we had set scoring parameters, we scored the object direction variable independently of the object type and location variables because we wanted to evaluate whether the direction ranges could discriminate between individuals in the at-risk group. Results of the Fisher's exact test indicated scoring parameters discriminated at-risk from comparison drivers for 18 of 28 direction ranges ($p<0.05$). This means that at-risk drivers were more likely to incorrectly classify direction for 64% of the objects when compared to the comparison group. There were no between-group differences in correctly identifying the direction of 10 objects. This included the pedestrian on the left in DS item 1, for which all participants correctly classified the direction. Fisher's exact test results appear in Table 2 for the remaining 27 objects. (The response count is different between groups because data for 2 participants in Group 2 were removed).

Table 2

Fisher's exact test results for object direction ranges displaying differences between performance of at-risk and comparison drivers where p>0.05 = no difference.

DS Item	Object	*p*	At-risk Group Response count /30	Comparison Group Response Count /28
DS 1	Car	0.04	20	27
	Person R	<0.00	9	27
DS 2	Car	<0.00	7	27
	Truck	0.01	12	28
	Person	<0.00	10	28
DS 3	Car	0.01	23	28
	Couple	0.54	20	28
DS 4	**Couple**	**0.29**	**19**	**28**
	Car	<0.00	18	28

DS 5	**Person**	**0.05**	**20**	**28**
	Truck	**0.14**	**14**	**22**
	Car	0.01	18	25
DS 6	**Bike**	**0.24**	**9**	**28**
	Person R	0.01	13	28
	Person L	<0.00	13	27
	Car	<0.00	23	25
DS 7	**Person**	**0.07**	**20**	**28**
	Couple	0.01	31	28
DS 8	**Bike**	**0.14**	**9**	**28**
	Car	**0.11**	**14**	**28**
	Person R	<0.00	9	28
	Person L	<0.00	13	25
DS 9	Person	0.01	13	28
	Person R	0.02	16	28
	Car	0.01	23	28
DS 10	**Bike**	**0.16**	**19**	**28**
	Person	<0.00	17	28

Note. DS = DriveSafe; R = Right ; L = Left; Bold = no differences (p>0.05) between performance of the at-risk drivers and the comparison group

We did not adjust direction ranges for the 10 objects that did not discriminate between participant groups at this stage because the average response entry was lower for the at-risk group for all objects, indicating older adults found recall of object direction more challenging. Additionally, at-risk drivers missed 30% of directions (*M*=8.4 object directions per person; *SD*=4.1) and misidentified 17% (*M*=4.7; *SD*=2.8) (see Table 1). Thus, at-risk drivers missed or misidentified 47% of directions (*M*=13.0; range=5-21; *SD*=4.6) compared to 3% (*M*=0.8; range=0-4; *SD*=1.0) for the comparison group. At-risk drivers entered 8 additional responses (*M*=0.3; *SD*=0.7); the comparison group entered no additional responses. The at-risk driver's mean total score for the object direction category was 15.0/28 (range=7-23; *SD*=4.7) compared to 27.2/28 (range=24-28; *SD*=1.0) for Group 2. Results of a Kolmogorov-Smirnov test of normality indicated the data were skewed. Therefore, a Mann-Whitney U test was performed with a Bonferroni correction. Results indicated a significant difference (p<0.00). These findings indicated the correct ranges we generated for the direction variable were able to discriminate between drivers in the at-risk category for most objects.

Discussion

The concern of the present study was the development of an automatic, touchscreen variable data collection and scoring system for a standardised assessment that reflected the decisions that would otherwise have been made by an expert rater. This had not been done previously for DriveSafe. Therefore, we applied a structured process (Aparasu, 2011; Summers, 1970) to guide the project, and examined touchscreen responses entered by at-risk older drivers and a comparison group to answer the question how generous scoring parameters should be when awarding points for the location and direction variables. Results indicated the data collection and scoring system we established for each of the three DriveSafe variables accurately recorded data, and that the location and direction variable scoring parameters we set discriminated between at-risk and comparison drivers for most objects. Fisher's exact test confirmed our hypothesis that at-risk drivers would be more likely than comparison drivers to incorrectly identify object variables in the driving environment for 24 of 28 location zones and 18 of 28 direction ranges. Average response entry was lower and frequency of missed or incorrect responses much higher for at-risk drivers for most objects. At-risk drivers also entered 31 additional responses for objects not displayed whereas comparison drivers entered no additional responses. These results indicated at-risk drivers found recalling object characteristics more challenging than the comparison group for all hazards displayed. This suggests that the automated scoring rules we developed reflected the scoring decisions that would have been made by an expert rater when scoring the correctness of responses for object type, location and direction of movement in original DriveSafe. While reporting of this component of the test conversion process is often overlooked, it is an essential step prior to the psychometric testing of the touchscreen version of the DriveSafe. Psychometric testing is the subject of a future study.

Findings from this study have important implications for researchers planning to convert existing standardised assessments for digital administration. We found that success relied on a number of key factors. An understanding of the existing knowledge base regarding the construct was essential so critical mechanisms necessary for assessment of driver visual attention were not lost when making decisions about how to capture data (e.g., masking of object change in DriveSafe). We attempted to replicate the original test as closely as possible to retain the test's psychometric properties and outcomes: including retaining the variables measured and the previously determined scale of measurement where possible. The validity, reliability and predictive validity of touchscreen DSDA were examined in a further study to confirm the test had retained these psychometric properties (Cheal & Kuang, 2015). We found that the characteristics of the data capture method (i.e., touchscreen) affected data collection and scoring decisions. For example, we had to collapse two variable categories in the transition from verbal to touchscreen response. Examining responses from older, at-risk and comparison drivers was essential for informing decisions regarding scoring parameters because we could not predict how participants would respond and there was no right answer regarding how generous scoring parameters should be. We did not want to set parameters so generous it was not possible to discriminate between individuals. Conversely, we did not want to set parameters so narrow participants without deficits could not pass the test.

Because the focus of the study was on determining data capture and scoring methods, we did not consider factors such as object type, size, trajectory, location, and luminosity in the study. The impact of these factors on observation or misidentification of objects is therefore outside the scope of the study. However, it is of interest that trucks and bikes made up 61% of missed and misidentified objects for comparison drivers, even though there were only two trucks and 3 bikes in the test, whereas pedestrians (16 individuals or couples) made up 55% of missed and misidentified objects for at-risk drivers. Comparison drivers demonstrated high accuracy in recalling all variables compared to at-risk drivers, who missed and misidentified a high number of object features broadly across object types. Consistent with our previous experience with original DriveSafe, at-risk drivers found object direction of movement the most difficult variable to accurately recall.

These results raise questions regarding how drivers allocate attention at intersections: whether healthy drivers prioritise areas of the driving scene important for safe decision making along with pedestrians and cars, whereas older drivers may allocate attention more broadly and miss important details. Researchers (Caird et al., 2005; Hoffman et al., 2006; Koustanaï et al., 2008; Rizzo et al., 2009) propose that the common occurrence of drivers failing to perceive hazards in clear view, particularly bikes, may be due to the challenge of managing limited attention resources in complex traffic situations. Drives focused on systematically scanning the driving environment for imminent threats, may fail to perceive less frequent or serious ones: particularly objects such as bikes and pedestrians that have atypical properties and presentations (e.g., speed and movement patterns) (Koustanaï et al., 2008). Given the age-related increased crash risk at intersections and the high number of crashes attributed to missing or delayed hazard perception for all drivers (Baldock et al., 2016; Barco et al., 2015; Caird et al., 2005; Fildes, 2008; Preusser et al., 1998; Rakotonirainy et al., 2012), further research regarding the impact of factors such as object type and location on hazard detection may be beneficial in road safety research and prediction of driving impairment (Hoffman et al., 2006; Koustanaï et al., 2012).

Further research is required to test the usability of touchscreen DriveSafe with older adults, the group most like to take touchscreen DriveSafe and most likely to struggle with the technology, before a final version was ready for psychometric testing (conducted in a subsequent phase of development). Additionally a further study was required to examine the psychometric properties and predictive validity of data gathered with touchscreen DriveSafe before it could be used in clinical practice (results published elsewhere) (Cheal & Kuang, 2015).

Limitations

A convenience sample was selected for the older group, and the comparison group self-selected to participate. Therefore, the groups may not be representative of the wider population. The two groups were not aged matched and it was assumed that group 2 participants were not at-risk driving. However, this was because the purpose of the study was test development: we sought a group of at-risk older drivers to compare with younger, healthy drivers so we could compare responses and set scoring parameters. Sample sizes were also small. However, there was reasonable agreement on location and direction data: the variables of interest for establishing scoring parameters. Therefore, we did not collect additional data.

Conclusion

Occupational therapists and other health professionals require access to sound assessment tools for fitness-to-drive measurement so they can give reliable advice in this high-stakes area: identifying clients at the extremities (i.e., clearly safe or unsafe to drive) and those requiring further assessment. Conversion of original DriveSafe to touchscreen was the first step in a development process conducted to provide occupational therapists with a first-level screen of cognitive fitness to drive. The structured process described in this study and usability testing of touchscreen DriveSafe with at-risk and comparison drivers provided us with information critical to the variable data collection and scoring decisions made and enabled us to retain construct measurement in the digital test conversion.

Key Points for Occupational Therapy

- This study applied a structured process for conversion of a standardised, expert-rated assessment to an automatically-scored touchscreen assessment.
- Scoring rules supported effective discrimination between at-risk older drivers and the comparison group.
- The automated scoring procedures for touchscreen DriveSafe appear to accurately capture participant performance.

Conflicts of Interest

This project was supported by Pearson, who publishes DSDA. The first author was employed by Pearson to project manage the conversion to touch screen. The second author is a DSDA author and receives royalties from sale of the tests. The fifth author is employed by Pearson as a project director.

References

Aparasu, R. R. (2011). *Research Methods for Pharmaceutical Practice and Policy* (1st ed.). London, UK: Pharmaceutical Press.

Baldock, M. R., Thompson, J. P., Dutschke, J., Kloeden, C. N., Lindsay, V. L., & Woolley, J. E. (2016). *Older Road Users: Emerging Trends Research Report AP-R530-16*. Sydney, Australia: Austroads Ltd.

Barco, P. P., Baum, C. M., Ott, B. R., Ice, S., Johnson, A., Wallendorf, M., & Carr, D. B. (2015). Driving errors in persons with dementia. *Journal of American Geriatric Society, 63*(7), 1373-1380. doi:10.1111/jgs.13508

Caird, J. K., Edwards, J. D., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors, 47*(2), 235-249.

Cheal, B., & Kuang, H. (2015). *DriveSafe DriveAware for Touch Screen Administration Manual.* Sydney, Australia: Pearson Australia Group Pty Ltd.

Crundall, D. (2009). The deceleration detection flicker test: A measure of experience? *Ergonomics, 52*(6), 674-684. doi:10.1080/00140130802528337

Daly, B. P., Nicholls, E. G., Patrick, K. E., Brinckman, D. D., & Schultheis, M. T. (2014). Driving behaviors in adults with autism spectrum disorders. *Journal of Autism and Devlopmental Disorders, 44*(12), 3119-3128. doi:10.1007/s10803-014-2166-y

Dickerson, A. E. (2014a). Driving with dementia: Evaluation, referral, and resources. *Occupational Therapy in Health Care, 28*(1), 62-76. doi:10.3109/07380577.2013.867091

Dickerson, A. E. (2014b). Screening and assessment tools for determining fitness to drive: A review of the literature for the pathways project. *Occupational Therapy In Health Care, 28*(2), 82-121. doi:10.3109/07380577.2014.904535

Downing Peck, A. (2011). App-solutely fabulous. Hundreds of new apps for iPad and tablets make MHealth a reality and a lifestyle choice. *Medical Economics, November 25, 2011*, 11-14.

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

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*, 189-198.

Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A. (2010). *Mini-Mental Stated Examination (MMSE-2) User's Manual* (2nd ed.). Lutz, Florida: Psychological Assessment Resources.

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

Glaser, D., Jain, S., & Kortum, P. (2013). Benefits of a physician-facing tablet presentation of patient symptom data: Comparing paper and electronic formats. *BMC Medical Informatics & Decision Making, 13*(99), 1-8.

Godwin, Z., Tan, J., Bockhold, J., Ma, J., & Tran, N. K. (2015). Development and evaluation of a novel smart device-based application for burn assessment and management. *Burns, 41*(4), 754-760. doi:10.1016/j.burns.2014.10.006

Hines, A., & Bundy, A. (2014). Predicting driving ability using DriveSafe and DriveAware in people with cognitive impairments: A replication study. *Australian Occupational Therapy Journal, 61*(4), 224-229. doi:10.1111/1440-1630.12112

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

Howell, M., Hood, A. J., & Jayne, D. G. (2015). Use of a patient completed iPad questionnaire to improve pre-operative assessment. *Journal of Clinical Monitoring and Computing*. doi:10.1007/s10877-015-9818-0

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(5), 529-546. doi:10.1002/wcs.130

Kay, L., Bundy, A., & Clemson, L. (2008). Predicting fitness to drive using the Visual Recognition Slide Test (USyd). *American Journal of Occupational Therapy, 62*, 187-197.

Kay, L., Bundy, A., & Clemson, L. (2009a). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation, 90*(9), 1514-1522. doi:10.1016/j.apmr.2009.03.011

Korner-Bitensky, N., Toal-Sullivan, D., & von Zweck, C. (2007). Driving and older adults: Towards a national occupational therapy strategy for screening. *Occupational Therapy Now, 9*(4), 3-5.

Koustanaï, A., Van Elslande, P., & Bastien, C. (2012). Use of change blindness to measure different abilities to detect relevant changes in natural driving scenes. *Transportation*

Research Part F: Traffic Psychology and Behaviour, 15(3), 233-242. doi:10.1016/j.trf.2011.12.012

Koustanaï, A., Boloix, E., Van Elslande, P., & Bastien, C. (2008). Statistical analysis of "looked-but-failed-to-see" accidents: Highlighting the involvement of two distinct mechanisms. Accident Analysis & Prevention, 40(2), 461-469. doi:10.1016/j.aap.2007.08.001

O'Donnell, J. M., Morgan, M. K., & Manuguerra, M. (2018). Patient functional outcomes and quality of life after microsurgical clipping for unruptured intracranial aneurysm. A prospective cohort study. *Journal of Neurosurgery, Published online March 2, 2018*. doi:10.3171/2017.8.JNS171576

O'Regan, J. K., Rensink, R. A., & Clark, J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398*, 34.

Preusser, D. F., Williams, A. F., Ferguson, S. A., Ulmer, R. G., & Weinstein, H. B. (1998). Fatal crash risk for older drivers at intersections. *Accident Analysis & Prevention, 30*(2), 151-159.

Rakotonirainy, A., Steinhardt, D., Delhomme, P., Darvell, M., & Schramm, A. (2012). Older drivers' crashes in Queensland, Australia. *Accident Analysis & Prevention, 48*, 423-429. doi:10.1016/j.aap.2012.02.016

Rensink, R. A. (2002). Change detection. *Annual Review Psychology, 53*, 245-277.

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.

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., & Vecera, S. P. (2009). Change blindness, aging, and cognition. *Journal of Clinical and Experimental Neuropsychology, 31*(2), 245-256. doi:10.1080/13803390802279668

Summers, G. F. (1970). Introduction. In G. F. Summers (Ed.), *Attitude Measurement* (pp. 1-21). Chicago, IL: Rand McNally.

Turney, B. W., & Reynard, J. M. (2014). Obtaining patient feedback in an outpatient lithotripsy service is facilitated by use of a touch-screen tablet (iPad) survey. *Urolithiasis, 42*(4), 317-321. doi:10.1007/s00240-014-0662-3

Vrkljan, B. H., McGrath, C., & Letts, L. (2011). Assessment tools for evaluating fitness to drive: A critical appraisal of evidence. *Canadian Journal of Occupational Therapy, 78*(2), 80-96. doi:10.2182/cjot.2011.78.2.3

Wiarda, N. R., McMinn, M. R., Peterson, M. A., & Gregor, J. A. (2014). Use of technology for note taking and therapeutic alliance. *Psychotherapy, 51*(3), 443-446. doi:10.1037/a0035075

CHAPTER 7

Journal Article 3

Predicting Fitness to Drive for Medically At-risk Drivers Using Touchscreen DriveSafe DriveAware

This following study was conducted to determine if the touchscreen version of DSDA had acceptable evidence for internal reliability and validity. The predictive validity of touchscreen DSDA was examined to determine if the test was sufficiently accurate to be used as a first-level screen of cognitive fitness to drive. The manuscript has been formatted according to the requirements of the journal to which it has been submitted. The manuscript is currently under review.

Results from this study have been presented at the following conferences:

Cheal, B., & Bundy, A. (2015, October). *Determining fitness to drive for older and cognitively impaired drivers - DriveSafe DriveAware a touchscreen test for medical practice*. Paper presented at The Australasian College of Road Safety Journal and Conference, Gold Coast, Australia.

Cheal, B., & Bundy, A. (2015, September). *DriveSafe DriveAware - A valid cognitive fitness to drive touchscreen test for medical practice?* Paper presented at the GP-15 The RACGP for General Practice, Melbourne, Australia.

Statement of Contribution of Authors

On behalf of the co-authors of "Predicting fitness to drive for medically at-risk drivers using touchscreen DriveSafe DriveAware", we confirm that Beth Cheal has made the following contributions:

- Conceptualisation and design of the research
- Collection of data
- Analysis and interpretation of findings
- Writing the paper and critical appraisal of content

Anita Bundy Signed: ______________________ Date: 04.09.2018

Beth Cheal Signed: ______________________ Date: 04.09.2018

Predicting Fitness to Drive for Medically At-risk Drivers Using Touchscreen DriveSafe DriveAware

Beth Cheal, MAppSc(OT)[a,c], Anita Bundy, ScD[a,d], Ann-Helen Patomella, PhD[b], Haijiang Kuang, PhD[c], Justin Newton Scanlan, PhD[a]

[a]*University of Sydney, Australia,* [b]*Karolinska Institutet, Sweden,* [c]*Pearson, Australia,* [d]*Colorado State University*

Abstract

Objective: To examine the psychometric properties and predictive validity of touchscreen DriveSafe DriveAware (DSDA), a screening test of cognitive fitness to drive for general practitioners and other health professionals.

Design: A prospective study comparing a screening test with a criterion standard.

Setting: Ten community and hospital-based driver assessment and rehabilitation clinics in Australia and New Zealand.

Participants: Older and cognitively impaired drivers (N=134) aged 18 to 91 years (M_{age}=68) referred for an assessment to determine the impact of a medical condition on ability to meet license authority standards. Inclusion criteria: a valid driver's license, vision within license-authority guidelines, completion of at least one year of high school, and English as a first language. Six patients declined to participate (n=4 female; M_{age}=69 years). Of the 142 patients who agreed, 1 withdrew after failing the on-road assessment and 7 did not meet inclusion or study criteria.

Interventions: Not applicable.

Main Outcome Measure: A standardized occupational therapy on-road assessment.

Results: Rasch analysis provided evidence for construct validity and internal reliability of the touchscreen DSDA. Optimal upper and lower cutoff scores were set to trichotomize drivers into three categories: likely to pass an on-road assessment; likely to fail an on-road assessment; and further testing required. Specificity of touchscreen DSDA was 86% and sensitivity 91%. The positive predictive value was 83%; negative predictive value, 92%. Overall accuracy of classification was 88%.

Conclusion: Evidence supports the clinical utility of the touchscreen version of DSDA for predicting with substantial accuracy which patients require on-road assessment.

Keywords: Automobile Driving; Safety; Cognition; Psychometrics; General Practitioners.

Drivers must monitor the visual scene to rapidly and accurately identify simultaneous information critical for safe decision-making (e.g., road signs and pedestrians) (Hoffman et al., 2006; Rizzo et al., 2009). Since failure to detect such information often results in accident, researchers propose that examining drivers' allocation of attention could be beneficial for predicting driving impairment (Caird et al., 2005; Hoffman et al., 2006; Koustanaï et al., 2012). For many years, researchers have examined clinical tests to identify one that can accurately predict driving performance off-road.

Researchers (Asimakopulos et al., 2012; Bédard et al., 2008; Molnar et al., 2006) have advised that a suitable fitness-to-drive screen must have: evidence-based cutoff scores with sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) above 80%; two cutoff scores for trichotomizing drivers into "likely to pass", "likely to fail" and "requires further testing" categories to reduce the problem of overlapping safe and unsafe scores (Laycock, 2011; Molnar et al., 2006); and, a small percentage of patients (10-20%) classified "further testing" for clinical utility (Bédard et al., 2008). Additionally, the screen must be sufficiently predictive to avoid misclassifications (Bédard et al., 2008) particularly minimizing categorizing unsafe drivers as safe, due to the risk of accident.

In most countries, general practitioners (GPs) are ultimately responsible for determining fitness to drive and are in an ideal position to screen drivers as a first point of contact. Surveys show GPs believe they should have this responsibility but lack valid and reliable tests, practical for medical practice (Classen et al., 2016; Dobbs et al., 1998; Fildes, 2008; Jang et al., 2007; Marshall et al., 2012; Molnar et al., 2006; Sims et al., 2012; Woolnough et al., 2013; Yale et al., 2003). Importantly, GPs require a test that patients will accept as related to real-world driving so they feel they have been fairly assessed and are more likely to accept the results (Dalchow et al., 2010). Many existing screening tools are population-specific (Devos et al., 2007; Hoggarth et al., 2013; Lundqvist, Gerdle, & Ronnberg, 2000; Marshall et al., 2007; Matas et al., 2014; Nouri & Lincoln, 1992; Piersma et al., 2016; Wood et al., 2013) (e.g., Stroke Driver Screening Assessment) (Nouri & Lincoln, 1992). However, fitness to drive is an overall concept associated with driving privilege (Drachman, 1988) and GPs require a population-free (i.e., generic) screening test that they can administer to any patient for whom driving is a concern. Considering these design criteria (Asimakopulos et al., 2012; Bédard et al., 2008; Dalchow et al., 2010; Molnar et al., 2006), we adapted an existing psychometrically-sound test (DriveSafe DriveAware; [DSDA]) (Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2003, 2008; Kay et al., 2009a, 2009b; O'Donnell et al., 2018) as a brief (10-minute), user-friendly screen (Cheal & Kuang, 2015).

The purpose of this study was to examine the psychometric properties and predictive validity of data gathered with touchscreen DSDA. An on-road assessment was the criterion measure: Performance Analysis of Driving Ability (P-Drive) (Patomella et al., 2010;

Patomella, Tham, & Kottorp, 2006), a standardized assessment of driving ability, coupled with therapist judgment determined the outcome.

Method

The University of Sydney Human Research Ethics Committee provided approval for this study. St Vincent's Hospital Sydney Human Research Ethics Committee provided ethics approval for the hospital sites.

Setting

This prospective study took place in ten driving clinics in Australia and New Zealand. Sixty percent of patients (n=81) were assessed in Australia. Driver-trained occupational therapists (n=16) conducted the assessments.

Participants

The sample consisted of 134 patients age 18+ (M_{age}=68 years; SD=16.9) referred to determine the impact of a medical condition on fitness to drive. The sample was divided into two groups. Group 1 included 34 participants aged 18 to 59 diagnosed with a neurological condition that had the potential to impact cognitive capacity (e.g., stroke, brain injury or Parkinson's disease). Participants who had primarily physical deficits were excluded from this group; to include them would have led to an overestimation of the accuracy of the prediction, as the purpose of DSDA is to determine cognitive fitness to drive. Group 2 included 100 participants aged 60 and over with any diagnosis, including general medical (n=3) and physical deficits (n=7). People with any diagnosis were included in the older group due to the increasing incidence of cognitive impairment with age and general community concern regarding cognitive fitness of older drivers.

Consecutive patients who met these inclusion criteria were invited: a valid driver's license, vision within license-authority guidelines, completion of at least 1 year of high school, and English as a first language. Patients were excluded if they had a developmental delay, psychiatric illness, or aphasia. Patients with physical deficits were only included if aged 65+ (range = 67-84; M_{age} = 76), to avoid overestimating the accuracy of prediction.

Six patients (*n*=4 female; M_{age}=69 years) declined to participate. Of the 142 patients who agreed, 1 withdrew after failing the on-road assessment; 5 did not meet the inclusion criteria; and the on-road assessment was incomplete for 2. The final sample (94 males, 40 females) is described in Table 1.

Table 1.

Patient Diagnoses

Diagnoses	*N*	(%)	MMSE-2:SV [a] *M*	*SD*
Traumatic brain injury	15	11	28.66	1.72
Stroke / TIA	42	31	27.14	2.62
Dementia / memory loss	38	29	22.84	4.56
Parkinson's disease	7	5	28.00	1.63
Other neurological	13	10	27.09	2.47
Physical deficits	7	5	28.00	1.15
General medical	3	2	26.00	1.73
Unsafe driving report / no formal diagnosis	9	7	25.38	5.26
Total	134	100	26.64	

[a] Mini-Mental State Examination - 2: Standard Version (MMSE-2: SV) is a 30-point standardized assessment of cognitive status, interchangeable with Mini-Mental State Examination (MMSE) (Folstein et al., 1975; Folstein et al., 2010) | Note: TIA = transient ischemic attack.

Instruments

We utilized two standardized assessments: touchscreen DSDA (Cheal & Kuang, 2015) and Performance Analysis of Driving Ability (P-Drive) (Patomella et al., 2010; Patomella et al., 2006).

Touchscreen DSDA. Touchscreen DSDA, a clinical screen of cognitive fitness to drive, comprises two subtests: DriveSafe and DriveAware. Scores on DriveSafe separate drivers into preliminary "safe", "unsafe", and "further testing categories". DriveAware scores are then applied to generate a final trichotomy based on the likelihood of passing an on-road assessment: "Likely to Pass", "Requires Further Testing", and "Likely to Fail." Previous research provides evidence of construct validity, internal reliability (Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2008; Kay et al., 2009a, 2009b) and test-retest reliability (O'Donnell et al., 2018) for the original version (original DSDA).

Touchscreen DriveSafe. DriveSafe measures awareness of the driving environment (Kay et al., 2009a, 2009b) via visual search and assumes attention is critical to safe driving. DriveSafe utilizes mechanisms identified by researchers (Caird et al., 2005; Galpin et al., 2009; Jensen et al., 2011; O'Regan et al., 1999; Rensink, 2002; Rensink et al., 1997) as necessary for assessment of driver visual attention: a driving-related task, use of naturalistic driving scenes and stimuli, masking of object change, and a time-limited, one-time display of objects. Touchscreen DriveSafe is self-administered via iPad (Cheal & Kuang, 2015) with written and audio instructions. The test consists of 10 images of a 4-way intersection (Figure 1 contains one example image). Each image includes two to four potential hazards (i.e., people or vehicles located in a particular place and moving a particular direction). Hazards are presented for 4 seconds then disappear leaving only the intersection. Participants are asked to recall information about the hazards by touching the blank intersection to identify: (i) each object from an array, (ii) its location, and (iii) direction of movement. A total of 28 objects is presented. Each piece of correctly identified information yields 1 point. The maximum score is 84: 28 for each object category (i.e., object type, location and direction).

Figure 1. DriveSafe Item Example

Previous experience with original DriveSafe indicates at-risk drivers find object direction of movement the most difficult to recall and object type the easiest. The item hierarchy presented in the original DriveSafe research confirms at-risk drivers find object type easiest to recall (Kay et al., 2009a) (object location and direction were combined variables in this study). Our hypothesis was that this structure would be represented in the touchscreen version item hierarchy.

Touchscreen DriveAware

DriveAware measures awareness of one's own driving abilities (Kay et al., 2009b). Awareness is operationalized as a lack of discrepancy between the participant's responses and an agreed standard (Kay et al., 2009b). Awareness is critical for safe driving because it enables the driver to self-monitor driving performance (Kay et al., 2009b). Seven DriveAware questions are administered: Two self-administered by the patient and five administered by a clinician (Figure 2 provides a sample item). One of the 7 questions contributes to rating on the agreed standard. Therefore, DriveAware consists of 6 items. Each item yields a discrepancy score based on the difference between the patient's self-rating, the clinician's ratings, or performance in the subtests. The total maximum score is 17.

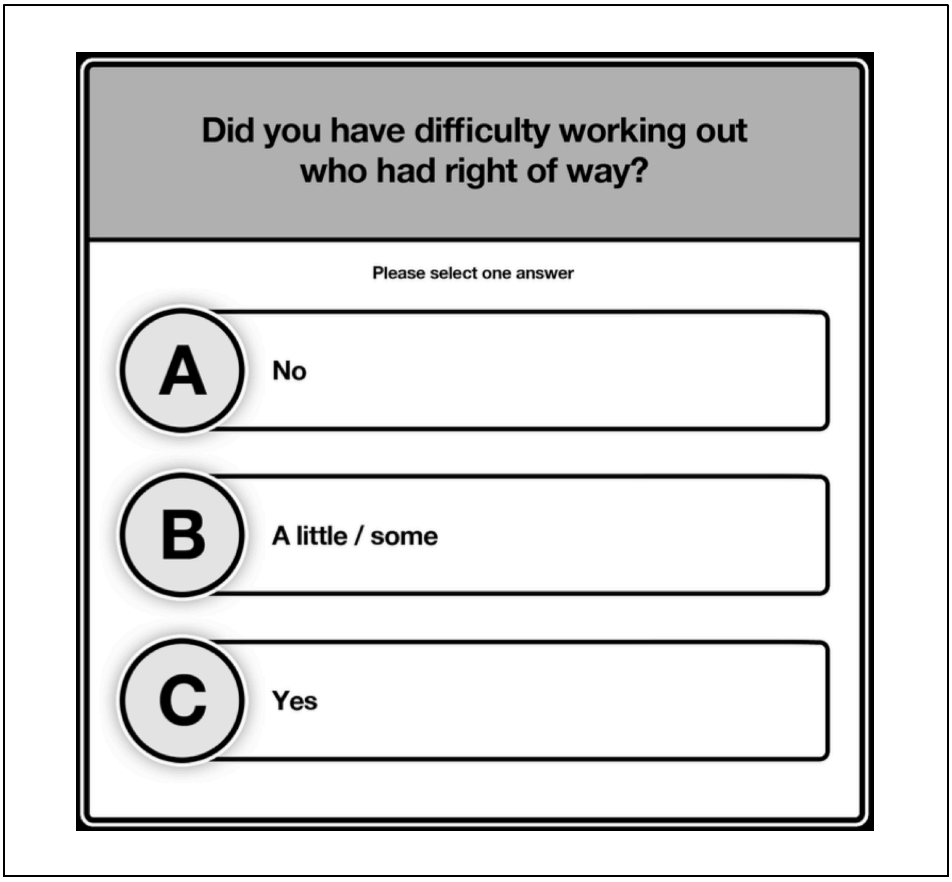


Figure 2. DriveAware Item Sample

DriveAware was developed as a supplement to DriveSafe to increase predictive accuracy, after original DriveSafe research identified awareness as critical for safe driving (Kay, Bundy, & Clemson, 2008). The original DriveAware item hierarchy reflected a progression of awareness from awareness of deficits in performance on cognitive tests, to awareness of deficits in daily life, to awareness of deficits that may impact driving performance (Kay et al., 2009a, 2009b). Our hypothesis was that this hierarchy would be similarly represented for touchscreen DriveAware.

Performance Analysis of Driving Ability (P-Drive)

P-Drive is a standardized observational assessment of driving ability: 25 items comprise four categories: Manoeuvres, Orients, Follows Regulations, and Attends and Acts (Patomella et al., 2010; Patomella et al., 2006). Items are scored on a 4-point scale (4 = competent). A score of 1 on any item is interpreted as unsafe driving. Examiners score the worst behaviour observed for each item. For example for the item 'Obeying stop signs and traffic lights', if the driver obeys three of four stop signs, the failure to stop is scored. P-Drive allows for route variation across centres because scoring is based on observed behaviour. P-Drive was adjusted by the third author for this study, due to the differences in road rules in Australia and New Zealand compared to Sweden. Summed scores yield a raw score out of 100. Cut-off scores of both 85 (Selander, Lee, Johansson, & Falkmer, 2011) and 81 (Patomella & Bundy, 2015) have been proposed. A cut-off score of 81 in the 2015 study yielded a specificity of 0.92 (i.e., 92% of drivers anticipated to pass did pass) and sensitivity of 0.93 (i.e., 93% of drivers anticipated to fail did fail). The positive predictive value was 0.95 and the negative predictive value was 0.90. However, no cut-off scores have been set for the Australian / New Zealand context; Whilst the cut-off scores informed the decision made, final determinations were based on P-Drive score coupled with clinical judgement using the following pass/fail criteria applied in previous DSDA research: Pass (i.e., safe and legal driving without intervention); Conditional pass (i.e., safe and legal driving with restrictions); Intervention (i.e., lessons required); Fail (i.e., failure to meet safe and legal driving criteria/ driving instructor intervention). Data for the pass and conditional pass groups were collapsed for analysis because both groups had achieved the criteria for safe driving. The combination of an on-road assessment tool with sound psychometric properties and a gestalt decision are supported in the literature as the optimum for standardising the on-road assessment and capturing all aspects of driving performance (Di Stefano & Macdonald, 2003; Justiss, Mann, Stav, & Velozo, 2006; Kay, Bundy, Clemson, et al., 2008; Shechtman, Awadzi, Classen, Landford, & Yongsung, 2010).

Procedure

The first three authors trained driver-trained occupational therapists (examiners) via two 1-hour interactive webinars. This included training in route design using compulsory route inclusions recommended by Di Stefano & Macdonald (2010) and Patomella & Bundy, (2015), and testing examiners for consistency in scoring P-Drive. Examiners designed a set-driving route in consultation with the driving instructors, who rated the level of challenge for each route so comparison could be made across centres. The first author reviewed all data to ensure participants met study criteria and assessments were conducted according to protocols.

Examiners administered touchscreen DSDA and MMSE-2:SV off-road, and conducted vision and physical function screenings. Participants completed a 60-minute on-road assessment in a dual-controlled vehicle; patients and examiners were blinded to DSDA results. A driving instructor seated in front provided route instructions and monitored safety. The examiner, seated in the rear, recorded driving performance and scored P-Drive once the on-road assessment concluded.

Statistical Analysis

We analysed DriveSafe and DriveAware data separately to evaluate item quality, remove problem items, and generate cut-off scores to trichotomise drivers based on the outcome of the on-road assessment. DriveAware cut-off scores were added to DriveSafe categories to refine classifications and generate the final trichotomy. We retained raw scores because they yielded the same results as the Rasch scores and are easier for clinicians to understand.

We used Rasch analysis (Winsteps Version 3.72.2) (Linacre, 2014) partial credit model to examine evidence of construct validity and internal reliability (Bond & Fox, 2007). Rasch analysis is suitable for a population with mixed diagnoses because it tests the assumption that easy items are easy for all participants and the most competent participants will perform best on all items, regardless of factors such as diagnosis (Bond & Fox, 2007; Kay, Bundy, & Clemson, 2008). Winsteps converts raw scores into interval scores (measures) expressed as log-odds probability units (logits) and generates a single hierarchy describing item difficulty and participant competence (Bond & Fox, 2007). Our analyses followed an iterative process with results from each phase informing the next. Iterative analyses allowed us to create the most parsimonious item set and to ensure that each item contributed to the unidimensional construct represented by the subtest. We examined the following sources of evidence for both subtests throughout the process.

At each step, we examined point-measure correlations to ensure each was positive (i.e., part of the construct) (Bond & Fox, 2007). Goodness-of-fit statistics revealed how well data conformed to the Rasch model assumptions: (i) more able participants have greater likelihood of passing difficult items; and, (ii) easy items are easy for all people (Bond & Fox, 2007). For each item and person, Winsteps generated two pairs of goodness-of-fit statistics expressed as mean square (*MnSq*) and standardized values (*ZStd*) (Bond & Fox, 2007). We considered removal of items where both *MnSq* and Z*Std* values were outside acceptable ranges (Bond & Fox, 2007; Tennant & Pallant, 2006).

We examined the extent to which the overall spread of items matched participant ability, with attention to areas along the hierarchy where the construct was overrepresented (i.e., more than one item at the same difficulty level), suggesting redundancy. We also identified gaps along the hierarchy indicating insufficient spread of item difficulty, reducing instrument precision.

We examined the person reliability index, a Cronbach's α equivalent, the person separation index, and number of strata (H= [4G+1]/3; where G = the separation index) for evidence of internal reliability. Strata indicate how reliably the test separates participants into statistically distinguishable groups (Fisher, 2007). We sought a person reliability index > 0.80 (Fisher, 2007) and an H value > 2.0.

At each iteration, a principal component analysis (PCA) of the residuals further contributed to examination of construct validity, providing evidence for the strength of any additional constructs represented in the subtests. We checked the empirical variance closely matched the Rasch-predicted model variance in the first factor and the percentage of explained variance was much greater than the percentage of unexplained variance (Linacre, 2003, 2014); an unexplained variance from the first factors < 3 eigenvalue units provides additional evidence that the test is unidimensional (Linacre, 2014). We also completed a uniform differential item function analysis (DIF) to check that females and males did not differ systematically on items (i.e., t values < 1.96) because research (Ackerman et al., 2011; D'Ambrosio, Donorfio, Coughlin, Mohyde, & Meyer, 2008; Molnar & Eby, 2008; Morgan, Winder, Classen, McCarthy, & Awadzi, 2009) indicates that women consistently express less confidence in their driving performance than men.

We examined evidence of the predictive validity of DSDA data by establishing optimal lower and upper cut score for each subtest using ROC curves. Lower cut scores maximized identification of unsafe drivers while minimizing misidentification of safe drivers (i.e., Sensitivity and PPV). Upper cut scores maximized identification of safe drivers while minimizing misclassification of unsafe drivers (i.e., Specificity and NPV).

Results

Of the 132 participants, 49% passed the driving assessment; 40% failed; and 11% required further testing or intervention.

DriveSafe

Point measure correlation coefficients were all positive and ranged from .44 to .75 (*m* = .63). All items had goodness-of-fit statistics within acceptable ranges (see Table 2). However, we removed one item in the first analysis because examiners reported technical difficulties while administering it and removal did not change the results.

Table 2.

Goodness-of-Fit Statistics, Item Measures (IM) and Standard Error (SE) for DriveSafe Items

	DriveSafe Infit		DriveSafe Outfit		IM	SE
Items	*MnSq*	*ZStd*	*MnSq*	*ZStd*		
Location - Image 1	1.06	0.48	0.88	-0.31	-2.02	0.14
Location - Image 2	0.91	-0.64	0.66	-1.27	-2.02	0.15
Location - Image 3	0.88	-0.65	1.36	0.8	-3.32	0.23
Location - Image 4	0.84	-0.86	0.58	-0.76	-3.38	0.23
Location - Image 5	0.91	-0.61	1.54	1.71	-2.47	0.16
Location - Image 6	1.34	2.17	1.37	2.02	-2.01	0.14
Location - Image 7	0.88	-0.58	0.74	-0.33	-2.98	0.22
Location - Image 8	0.97	-0.14	1.14	0.78	-1.95	0.13
Location - Image 9	1.20	1.46	1.42	1.58	-2.62	0.16
Location - Image 10	0.90	-0.5	0.89	-0.04	-2.94	0.22

	DriveSafe Infit		DriveSafe Outfit		IM	SE
Items Continued	*MnSq*	*ZStd*	*MnSq*	*ZStd*		
Object - Image 1	1.08	0.63	0.88	-0.44	-1.79	0.14
Object - Image 2	1.02	0.17	0.98	-0.01	-1.61	0.14
Object - Image 3	1.08	0.55	1.02	0.21	-3.03	0.22
Object - Image 4	0.63	-1.99	0.50	-0.91	-3.18	0.23
Object - Image 5	0.97	-0.18	0.85	-0.73	-1.93	0.15
Object - Image 6	1.17	1.28	1.19	1.41	-1.04	0.13
Object - Image 7	0.76	-1.37	0.54	-0.88	-3.36	0.24
Object - Image 8	1.07	0.56	1.09	0.69	-1.18	0.13
Object - Image 9	0.86	-0.89	0.67	-0.91	-2.44	0.16
Object - Image 10	0.90	-0.61	1.04	0.23	-2.76	0.21
Direction - Image 1	0.92	-0.57	0.75	-1.15	-1.53	0.14
Direction - Image 2	1.00	0.03	0.77	-1.09	-1.24	0.13
Direction - Image 3	1.01	0.12	0.81	-0.47	-1.91	0.17
Direction - Image 4	0.82	-1.35	0.84	-0.49	-1.98	0.18
Direction - Image 5	0.97	-0.2	0.97	-0.14	-1.01	0.14
Direction - Image 6	1.38	2.72	1.35	2.53	-0.61	0.12
Direction - Image 7	0.88	-0.79	0.68	-0.71	-2.07	0.18
Direction - Image 8	1.12	0.9	1.11	0.77	-0.65	0.11
Direction - Image 9	1.06	0.52	1.24	1.42	-0.94	0.13
Direction - Image 10	1.02	0.21	0.79	-0.59	-1.73	0.17

The range of item difficulty was comparable to the range of participant ability in the map of items and drivers, except for the most competent drivers. There were few gaps along the hierarchy (see Figure 3).

```
                    Drivers     Items
Score     More Competent   |   Hardest

  4                  P  P  P  P  P  +
                                    |
               P  P  Q  P  P  P  P  |
                                   T|
                                    |
                                    |
  3                                 +
                           P  P  P  |
                                    |
                                    |
                  P  P  P  Q  P  P  |
                                    |
  2                  P  P  P  P  P  +
                                   S|
                  P  P  P  P  P  P  |
                     P  P  F  P  P  |
                                    |
                     P  Q  P  P  P  |
  1                     P  F  P  P  +
                                 P  |
                                 P  |
                              Q  Q  |
                           P  P  F  |
               F  F  P  P  Q  Q  P  |
  0                              F M+
      P  P  P  F  F  P  P  Q  P  P  |
                  F  F  Q  F  P  P  |
                           F  F  F  |T
               P  P  F  P  P  Q  P  |  dir-Im6   dir-Im8
                        F  Q  Q  Q  |
 -1                           F  P  +  dir-Im9   dir-Im5   ob-Im6
                     P  F  F  F  P  |S dir-Im2   ob-Im8
                     F  F  F  P  F  |
                           P  P  F  |  dir-Im1
                     F  F  Q  F  F  |  dir-Im10  ob-Im2
                        F  P  F  P S|  dir-Im3   ob-Im1
 -2                     F  F  F  F  +M dir-Im4   dir-Im7   loc-Im1   loc-Im2   loc-Im6   loc-Im8   ob_s5
                           F  F  F  |
                           F  F  F  |
                                 P  |  loc-Im5   ob-Im9
                              F  F  |  loc-Im9
                                    |S ob-Im10
 -3                              Q  +  loc-Im10  loc-Im7   ob-Im3
                              F  F  |  ob-Im4
                                    |  loc-Im3   loc-Im4   ob-Im7
                                   T|
                                    |T
                                    |
 -4                                 +
                                 F  |
                                    |
                                    |
                                    |
                                    |
 -5                                 +
          Less Competent   |   Easiest
```

Note: F = failed on-road; P = passed on-road; Q = required further testing; dir = direction of movement; loc = location of object; ob = object; Im = Image number; M = mean; S = 1 standard deviation; T = 2 standard deviations

Figure 3. Map of Drivers and Items for DriveSafe Subtest

The test separated participants into >5 strata (separation index=3.76; H=5.34), providing evidence that DriveSafe is sensitive enough to distinguish multiple levels of awareness of the driving environment. A high reliability index (.93) provided evidence for the replicability of person placement along the hierarchy.

The PCA yielded an empirical variance (55.1%) closely matching the modeled variance (56.4%). The percentage of unexplained variance by the first contrast was 4.6% (Eigenvalue = 3.1), which is slightly more than desired (i.e., < 3) but much less than the variance explained by items (14.8%; Eigenvalue = 9.9) and persons (40.4%; Eigenvalue = 27.0). A uniform DIF analysis revealed no significant difference between males and females on any item.

DriveAware

We excluded data from 2 participants from analysis of DriveAware data due to missing data for those two participants. The iPad apparently failed in these two instances as no data were recorded.

Point measure correlations were positive, ranging from 0.64 to 0.82 (Mean = 0.70), indicating all items were part of the construct. All items except Item 5 had fit statistics within the acceptable ranges (see Table 3).

Table 3.

Goodness-of-Fit Statistics, Item Measures (IM) and Standard Error (SE) for DriveAware Items

	Infit		Outfit		IM	SE
Item	*MnSq*	*ZStd*	*MnSq*	*ZStd*		
1. Ability to recall DriveSafe location	0.95	-0.31	0.95	-0.19	-2.22	0.15
2. Ability to recall DriveSafe direction	0.75	-2.11	0.73	-1.94	-1.52	0.14
3. Concerns regarding driving performance	1.10	0.85	1.10	0.8	1.03	0.15
4. Surprised by hazards whilst driving	0.90	-0.72	0.88	-0.69	-0.40	0.13
5. Awareness of why asked to do DSDA	1.67	3.80	2.36	3.35	-0.58	0.14
6. Awareness of performance on DSDA	0.64	-3.20	0.62	-3.24	-0.99	0.14

Note: DSDA = DriveSafe DriveAware

As expected, the hierarchy of items was consistent with previous iterations of DriveAware (Kay et al., 2009a, 2009b), with awareness of reduced performance in the test being easier than awareness of the reason for being tested; awareness of reduced driving performance was the most difficult (see Figure 4). The map revealed gaps where there were people but no items.

```
                                         Drivers  | Questions
Score                              Intact Awareness | Most Aware
  5                                            F F +
                                                   |
                                                   |
                                                   |
                                                   |
                                                   |
  4                                                +
                                         P P P P F |
                                                   |
                                                   |
                                                  T|
                                                   |
  3                                                +
                                                   |
                                                   |
                                                   |
                                     F F P P P P P |
                                                   |
  2                                                +
                                                   |
                                                  S|
                       P P P P F P F P P P P P P P |
                                                   |
                                                   |T
  1         F P P P P P P P P P F P F F P P P P    + 3.Concern about driving
                                                   |
                                                   |
        F P P F P P P F P F P P F P P P P P P      |
                                                   |
                                                   |S
  0                            P P P P P P P P F P P M+
                                                   |
                                 F F F F F F F F P F | 4.Surprised by hazards
                                                   | 5.Why asked to do DSDA
                                                   |
                                   F F F F F F P P F |M
 -1                                                + 6.Performance on DSDA
                                             F F F |
                                                   |
                                 F P P F F F F F F F | 2.DS direction recall
                                                  S|
                                       P F F F P P |S
 -2                                                +
                                                   | 1.DS location recall
                                       F F F F F F |
                                                   |
                                                   |
                                       F F F F F P |T
 -3                                                +
                                                   |
                                               P F T|
                                                   |
                                                   |
 -4                                            F F +
                              Absent Awareness | Least Aware
```

Note: F = failed; P = passed; M = mean; S = 1 standard deviation; T = 2 standard deviations; DS = DriveSafe; DSDA = DriveSafe DriveAware

Figure 4. Map of drivers and items for DriveAware subtest

The person reliability index (0.80) was acceptable. A separation index of 1.98 revealed almost 3 strata (H = 2.95). The PCA yielded a modeled variance (59.3%) that closely matched the empirical variance (59.4%). The percentage of unexplained variance by the first contrast was 11.4% (Eigenvalue = 1.6), which is within the acceptable range (i.e., < 3) and less than the percentage of variance explained by item (16.1%; Eigenvalue = 2.2) and person (46.3%; Eigenvalue = 6.1). Uniform DIF analysis revealed females scored significantly higher (i.e., greater discrepancy) than males on Item 1 (t = -3.35) ("How well did you remember the location of people and vehicles"). No other item differed significantly.

Predictive Validity

Predictive validity refers to how accurately DSDA predicted drivers who actually passed/failed on-road. Thus, data from patients categorized as "requiring further testing" (n = 20) were excluded from this analysis because the pass/fail category was indeterminate for these participants.

ROC curves revealed optimal cutoff scores of 57 and 72 for DriveSafe and 10 and 13 for DriveAware based on results of the on-road assessment. The resulting driver trichotomization appears in Figure 5.

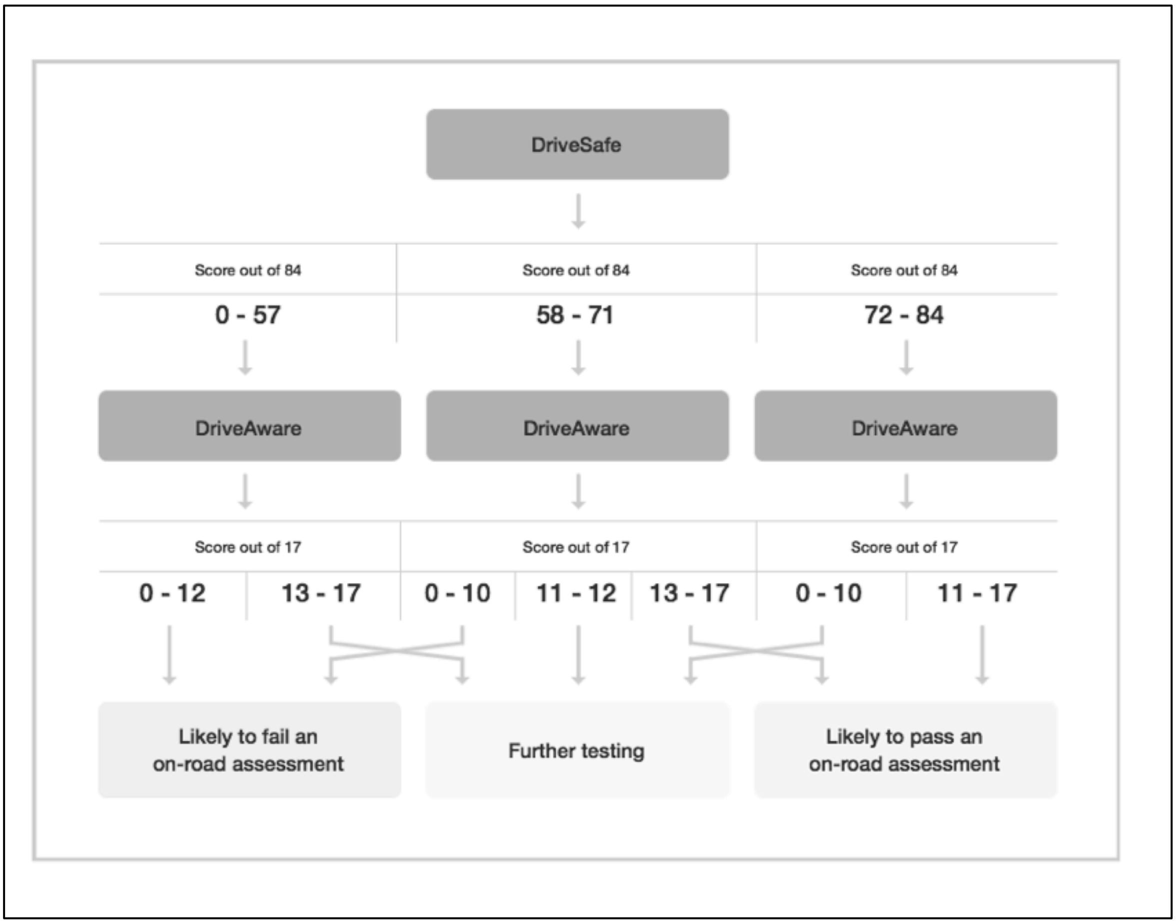


Figure 5. DSDA Cutoff Scores Trichotomy

Specificity, sensitivity, NPV and PPV of the combined and individual tests appear in Table 4. Classification accuracy appears in Table 5.

Table 4.

Descriptive Statistics (%) for DriveSafe and DriveAware Lower- and Upper- Cutoff Scores

	Lower Cutoff		Upper Cutoff		Combined
	DriveSafe	DriveAware	DriveSafe	DriveAware	DSDA
Sensitivity	91	89	63	51	91
Specificity	66	74	94	91	86
PPV	79	83	93	89	83
NPV	84	83	65	57	92
Accuracy of classification	80	83	68	68	88

Note: DSDA = DriveSafe DriveAware; PPV = positive predictive value; NPV = negative predictive value

Table 5.

Accuracy of DriveSafe DriveAware (DSDA) Classification

	Actual Driving Performance		
DSDA	Unsafe	Further Testing	Safe
Unsafe (fail test)	39 (34%) (true positive)	4 (4%)	4 (4%) (false positive)
Safe (pass test)	8 (7%) (false negative)	8 (7%)	49 (44%) (true negative)

Discussion

The purpose of this study was to examine the psychometric properties and predictive validity of data gathered with touchscreen DSDA. Touchscreen DSDA met criteria set by researchers for a useful GP driver screen. Researchers (Anstey et al., 2005; Asimakopulos et al., 2012; Bédard et al., 2008; Kay et al., 2009a; Molnar et al., 2006) agree that the statistics most valuable to clinicians selecting assessment tools are sensitivity, specificity, PPV, and NPV; these statistics should be ≥80% (Asimakopulos et al., 2012). Specificity of touchscreen DSDA was 86%; sensitivity 91%; PPV, 83%; and NPV, 92%. Overall accuracy of classification was 88%. These findings indicate touchscreen DSDA had slightly lower, but acceptable, sensitivity and specificity compared to original DSDA: specificity of touchscreen DSDA was 96-97% in original DSDA; sensitivity was 93-95% in original DSDA (Kay et al., 2009a). Touchscreen DSDA had acceptable evidence for internal reliability and construct validity similarly to original DSDA (Hines & Bundy, 2014; Kay, Bundy, & Clemson, 2008; Kay et al., 2009a, 2009b).

Eight participants passed DSDA but failed on-road (i.e., false negative); we could discern no distinguishing characteristics (such as diagnosis, age, gender, or driving performance). Nonetheless, GPs should use professional judgment and other clinical indicators in addition to touchscreen DSDA to support fitness-to-drive determinations. Further supporting clinical utility, Touchscreen DSDA has two cutoff scores for trichotomising drivers, reducing the problem of overlapping scores (Molnar et al., 2006); places only 11% of patients in the further testing category (Bédard et al., 2008); takes around 10 minutes; and has high face-validity (i.e., patients can perceive its relationship to driving).

The map of items and drivers indicates DriveSafe most precisely differentiates less competent drivers (i.e., there was a ceiling effect for driver who fell into the safe category).

This is not problematic as measurement precision is unnecessary to discriminate between "good" and "excellent" drivers; DriveSafe is designed to identify safe drivers. As expected, participants found direction of movement the most difficult variable to recall and object type the easiest for most items, consistent with previous research (Kay et al., 2009a).

The DriveAware map revealed gaps where there are no items to measure both the least and most aware participants. However, the purpose of DriveAware is to distinguish between poor performers where DriveSafe results are borderline and additional qualitative and quantitative information is required to categorize patients; the test is not used alone. DriveAware Item 5 ("Why have you been asked to complete DriveSafe DriveAware?") had fit statistics outside the acceptable range indicating somewhat erratic responses. Nonetheless, we retained Item 5 because of its contribution to judging patient insight. DIF analysis revealed females tended to underestimate their performance compared to males on item 1 ("How well did you remember the location of people and vehicles?"). Studies that examine gender differences in self-rated driving performance consistently show that women express less confidence in their performance than men (Ackerman et al., 2011; D'Ambrosio et al., 2008; Molnar & Eby, 2008; Morgan et al., 2009). Health professionals should be mindful of gender differences when interpreting DriveAware results.

Limitations

P-Drive allows consistent scoring over variable routes, nonetheless conditions likely varied somewhat. Relatively small numbers in patient groups precluded systematic examination of differences by diagnostic group. However, person fit statistics were within acceptable ranges and results of the principal components analysis indicated no evidence that diagnoses separated groups. Further research confirming the cut-off scores for specific diagnostic groups is needed.

Conclusion

Our results suggest that touchscreen DSDA is sufficiently predictive be used as a first-level cognitive fitness-to-drive screen. Only patients with inconclusive results require referral for further testing. Patients identified as likely to be safe can avoid unnecessary testing and patients identified as likely to be unsafe can be advised to discontinue driving.

Conflicts of Interest

This project was supported by Pearson, who publishes DSDA. The first author was employed by Pearson to project manage the conversion to touch screen. The second author is a DSDA author and receives royalties from sale of the tests. The fourth author, who conducted the statistical analysis, is employed by Pearson as a senior psychometrician.

References

Ackerman, M. L., Crowe, M., Vance, D. E., Wadley, V. G., Owsley, C., & Ball, K. K. (2011). The impact of feedback on self-rated driving ability and driving self-regulation among older adults. *Gerontologist, 51*(3), 367-378. doi:10.1093/geront/gnq082

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., Menard, 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

Bédard, M., Weaver, B., Darzins, 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

Bekiaris, E., Amditis, A., & Panou, M. (2003). DRIVABILITY: A new concept for modelling driving performance. *Cognition, Technology & Work, 5*(2), 152-161. doi:10.1007/s10111-003-0119-x

Blackmore, S. J., Breistaff, G., Nelson, K., & Trościanko, T. (1995). Is the richness of our visual world and illusion? Transsaccadic memory for complex scenes. *Perception, 24*(9), 1075-1081.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model* (2nd ed.). Mahwah New Jersey: Lawrence Erlbaum Associates.

Caird, J. K., Edwards, J. D., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors, 47*(2), 235-249.

Cheal, B., & Kuang, H. (2015). *DriveSafe DriveAware for Touch Screen Administration Manual*. Sydney, Australia: Pearson Clinical.

Crundall, D. (2009). The deceleration detection flicker test: A measure of experience? *Ergonomics, 52*(6), 674-684. doi:10.1080/00140130802528337

D'Ambrosio, L. A., Donorfio, L. K. M., Coughlin, J. F., Mohyde, M., & Meyer, J. (2008). Gender Differences in Self-Regulation Patterns and Attitudes Toward Driving Among Older Adults. *Journal of Women & Aging, 20*(3-4), 265-282. doi:10.1080/08952840801984758

Dalchow, J. L., Niewoehner, P. M., Henderson, R. R., & Carr, D. B. (2010). Test acceptability and confidence levels in older adults referred for fitness-to-drive evaluations. *American Journal of Occupational Therapy, 64*, 252-258.

Di Stefano, M., & Macdonald, W. (2003). Assessment of older drivers: Relationships among on-road errors, medical conditions and test outcome. *Journal of Safety Research, 34*(4), 415-429. doi:10.1016/j.jsr.2003.09.010

Dobbs, A. R., Heller, R. B., & Schopflocher, D. (1998). A comparitive approach to identify unsafe older driver. *Accident Analysis & Prevention, 30*(3), 363-370.

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

Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions, 21*, 1095.

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*, 189-198.

Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A. (2010). *Mini-Mental Stated Examination (MMSE-2) User's Manual* (2nd ed.). Lutz, Florida: Psychological Assessment Resources.

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

Hatakka, M., Keskinen, E., Gregersen, N. P., Glad, A., & Hernetkoski, K. (2002). From control of the vehicle to personal self-control; Broadening the perspectives to driver education. *Transportation Research Part F 5*, 201-215.

Hines, A., & Bundy, A. (2014). Predicting driving ability using DriveSafe and DriveAware in people with cognitive impairments: A replication study. *Australian Occupational Therapy Journal, 61*(4), 224-229. doi:10.1111/1440-1630.12112

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

Jang, R. W., Man-Son-Hing, M., Molnar, F. J., Hogan, D. B., Marshall, S. C., Auger, J., . . . Naglie, G. (2007). Family physicians' attitudes and practices regarding assessments of medical fitness to drive in older persons. *Journal of General Internal Medicine, 22*(4), 531-543. doi:10.1007/s11606-006-0043-x

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(5), 529-546. doi:10.1002/wcs.130

Justiss, M. D., Mann, W. C., Stav, W., & Velozo, C. (2006). Development of a behind- the-wheel driving performance assessment for older adults. *Topics in Geriatric Rehabilitation, 22*(2), 121-128.

Kay, L., Bundy, A., & Clemson, L. (2003). Awareness of driving ability in senior drivers with neurological conditions. *American Journal of Occupational Therapy, 63*, 146-150.

Kay, L., Bundy, A., & Clemson, L. (2008). Predicting fitness to drive using the Visual Recognition Slide Test (USyd). *American Journal of Occupational Therapy, 62*, 187-197.

Kay, L., Bundy, A., & Clemson, L. (2009a). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation, 90*(9), 1514-1522. doi:10.1016/j.apmr.2009.03.011

Kay, L., Bundy, A., & Clemson, L. (2009b). Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation, 31*(13), 1074-1082. doi:10.1080/09638280802509553

Kay, L., Bundy, A., Clemson, L., & Jolly, N. (2008). Validity and reliability of the on-road driving assessment with senior drivers. *Accident Analysis & Prevention, 40*(2), 751-759. doi:10.1016/j.aap.2007.09.012

Koustanaï, A., Van Elslande, P., & Bastien, C. (2012). Use of change blindness to measure different abilities to detect relevant changes in natural driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*(3), 233-242. doi:10.1016/j.trf.2011.12.012

Linacre, J. M. (2003). Data Variance: Explained, Modeled and Empirical. *Rasch Measruement Transactions.* Retrieved from http://www.rasch.org/rmt/rmt173g.htm

Linacre, J. M. (2014). *Winsteps Rasch Measurement Computer Program User's Guide*. Beaverton, Oregon: Winsteps.com.

Lützenburger, M., & Albayrak, S. (2013). *Can you simulate traffic psychology? An Analysis.* Paper presented at the 2013 Winter Simulation Conference, Washingon, DC.

Marshall, S. C., Demmings, E. M., Woolnough, A., Salim, D., & Man-Son-Hing, M. (2012). Determining fitness to drive in older persons: A survey of medical and surgical specialists. *Canadian Geriatrics Journal, 15*(4), 101-119. doi:10.5770/cgj.15.30

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), *Human Behaviour and Traffic Safety* (pp. 485-520). New York: Plenum Press.

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 American Geriatric Society, 54*(12), 1809-1824. doi:10.1111/j.1532-5415.2006.00967.x

Molnar, L. J., & Eby, D. W. (2008). The relationship between self-regulation and driving-related abilities in older drivers: An exploratory study. *Traffic Injury Prevention, 9*(4), 314-319. doi:10.1080/15389580801895319

Morgan, C. M., Winder, S. M., Classen, S., McCarthy, D., & Awadzi, K. D. (2009). Literature review on older adult gender difference for driving self-regulation and cessation. *Topics in Geriatric Rehabilitation, 25*(2), 99-117.

O'Donnell, J. M., Morgan, M. K., & Manuguerra, M. (2018). Patient functional outcomes and quality of life after microsurgical clipping for unruptured intracranial aneurysm. A prospective cohort study. *Journal of Neurosurgery, Published online March 2, 2018*. doi:10.3171/2017.8.JNS171576

O'Regan, J. K., Rensink, R. A., & Clark, J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398*, 34.

Patomella, A., & Bundy, A. (2015). P-Drive: Implementing an assessment of on-road driving in clinical settings and investigating Its internal and predictive validity. *American Journal of Occupational Therapy, 69*(4), 6904290010p6904290011-6904290018. doi:10.5014/ajot.2015.015131

Patomella, A., Tham, K., Johansson, K., & Kottorp, A. (2010). P-Drive on-road: Internal scale validity and reliability of an assessment of on-road driving performance in people with neurological disorders. *Scandinavian Journal of Occupational Therapy, 17*(1), 86-93. doi:10.3109/11038120903071776

Patomella, A., Tham, K., & Kottorp, A. (2006). P-Drive: Assessment of driving performance after stroke. *Journal of Rehabilitation Medicine 2006, 38*, 273-279. doi:10.1080/16501970600632594

Rensink, R. A. (2002). Change detection. *Annual Review Psychology, 53*, 245-277.

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.

Richard, C. M., Wright, R. D., Ee, C., & Prime, S. L. (2002). Effect of a concurrent auditory task on visual search performance in a driving-related image-flicker task. *Human Factors, 44*(1), 108-119.

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., & Vecera, S. P. (2009). Change blindness, aging, and cognition. *Journal of Clinical and Experimental Neuropsychology, 31*(2), 245-256. doi:10.1080/13803390802279668

Selander, H., Lee, H. C., Johansson, K., & Falkmer, T. (2011). Older drivers: On-road and off-road test results. *Accident Analysis & Prevention, 43*(4), 1348-1354. doi:10.1016/j.aap.2011.02.007

Shechtman, O., Awadzi, K. D., Classen, S., Landford, D. N., & Yongsung, J. (2010). Validity and critical driving errors of on-road assessment for older drivers. *American Journal of Occupational Therapy, 64*, 242-251.

Sims, J., Rouse-Watson, P., Schattner, A., Beveridge, A., & Jones, K. (2012). To drive or not to drive: Assessment dilemmas for GPs. *International Journal of Family Medicine, 2012*, 1-6. doi:10.1155/2012/417512

Tennant, A., & Pallant, J. F. (2006). Unidimensionality matters! (A tale of two smiths?). *Rasch Measruement Transactions, 20*, 1048-1051.

Veirk, E., & Kiesel, A. (2008). Change detection: Evidence for information accumulation in flicker pardigms. *Acta Psychologica, 127*, 309-323.

Woolnough, A., Salim, D., Marshall, S. C., Weegar, K., Porter, M. M., Rapoport, M. J., . . . Vrkljan, B. (2013). Determining the validity of the AMA guide: A historical cohort analysis of the assessment of driving related skills and crash rate among older drivers. *Accident Analysis & Prevention, 61*, 311-316. doi:10.1016/j.aap.2013.03.020

Yale, S., Hansotia, P., Knapp, D., & Ehrfurth, J. (2003). Neurologic conditions: Assessing medical fitness to drive. *Clinical Medicine & Research, 1*(3), 177-188.

CHAPTER 8

Determining Suitability of Touchscreen DSDA for Medical Practice (Discontinued Study)

Findings from the literature review described in Chapter 3 established that general practitioners (GPs) need, but do not currently have access to, suitable fitness-to-drive screens. Surveys conducted to determine how GPs are currently tackling driver screening in the absence of suitable tools present a unified picture. Although GPs generally agree they should be responsible for determining medical fitness to drive, they report a lack of confidence due to non-specific medical guidelines, insufficient training, and lack of resources (Classen et al., 2016; Hoggarth, 2013; Jang et al., 2007; Jones et al., 2012; Marshall et al., 2012; Omer et al., 2014; Sims et al., 2012; Wilson & Kirby, 2008). Therefore, GPs tend to take an ad hoc approach, without clear guidelines, an agreed standard, or specific assessment protocols (Braekhus & Engedal, 2009; Jang et al., 2007; Sims et al., 2012; Wilson & Kirby, 2008). The most common approach is to conduct vision screening: often or always performed by 93% to 99% of respondents in 3 surveys (Braekhus & Engedal, 2009; Jang et al., 2007; Wilson & Kirby, 2008). Aside from this procedures vary widely, including a combination of history taking, cardiac examinations, abbreviated neurological examinations, review of medications, physical test (e.g., strength testing), and hearing testing (Braekhus & Engedal, 2009; Jang et al., 2007; Sims et al., 2012; Wilson & Kirby, 2008). Unfortunately these category of tests are unsuitable for determining fitness to drive because the purpose of physical and medical examinations is to detect disease, not determine functional status or fitness to drive (Laycock, 2011).

The percentage of GPs who often or always use formal cognitive tests to screen fitness to drive varies across studies: 22% (Braekhus & Engedal, 2009); 49% (Jang et al., 2007); and, 56% (Sims et al., 2012). Omer et al. (2014) found that less than one third reported "often or always" assessing cognition (Omer et al., 2014). Wilson and Kirby (2008) reported that 26% of GPs (*N* = 99) reported they did not assess cognitive function at all (Wilson & Kirby, 2008). This may reflect the lack of valid, reliable and practical assessments available for GPs to use. The MMSE (Folstein et al., 1975; Folstein et al., 2010) is overwhelmingly the most common cognitive test administered (Dobbs et al., 1998; Hoggarth, 2013; Marshall et al., 2007; Wilson & Kirby, 2008): used by 80% of GPs who formally test cognition when screening fitness to drive in one Norwegian study (Braekhus & Engedal, 2009), and in general practice for 88% of GPs in a New Zealand study (Hoggarth, 2013). Unfortunately the MMSE and similar pen and paper tests (e.g., the MoCA and clock drawing test) are not sufficiently predictive of on-road performance to reliably use in driver screening, as discussed in Chapter 3.

Pearson Clinical Assessment (Australia and New Zealand) conducted market research prior to the DriveSafe DriveAware (DSDA) digital development, to determine how GPs were currently determining fitness to drive, if they perceived a need for a fitness-to-drive screen, and how they preferred the test be designed for medical practice. They surveyed a representative sample of 200 Australian GPs in 2012. Results, reported in Appendix C, are consistent with the survey findings described above. Additionally GPs reported finding the national medical guidelines cumbersome, ambiguous and not user-friendly; receiving insufficient information from patients and family to make an informed decision; and concern about lack of government funding for fitness-to-drive assessments. GPs reported they relied on patient observations (e.g., mobility and grooming), medical examinations, vision assessment, and information provided by family to determine fitness to drive. Despite a reported lack of confidence assessing fitness to drive, most (63%) GPs reported rarely or never referring patients for an occupational therapy on-road assessment. Barriers cited included cost, waiting times, patient reactions, concern about validity of the assessment, and service access.

These findings are consistent with published GP survey results, which indicate GPs are making decisions about fitness to drive without referring to the results of on-road assessment (Braekhus & Engedal, 2009; Hoggarth, 2013; Jang et al., 2007; Omer et al., 2014). Rates of referral for license authority or occupational therapy on-road tests are low, with GPs, often citing concerns about cost, waiting times and uncertainty regarding validity (Jones et al., 2012; Sims et al., 2012). In one Irish survey, 70% of GPs ($N = 257$) reported rarely or never referring for on-road assessment (Omer et al., 2014). A Norwegian study indicated GPs only referred for on-road assessments in exceptional cases (Braekhus & Engedal, 2009). This may be region-specific depending on practice traditions, awareness of referral pathways, and the level and cost of service provided. Occupational therapy driving assessors were considered invaluable in two more recent Australian GP surveys, with much higher referral rates (59% in one case) (Jones et al., 2012; Sims et al., 2012).

In summary, decades of research demonstrated that GPs need, but do not have access to, practical and valid driver screens. The focus of this thesis was on the development of such a screen. Examination of the psychometric properties and predictive validity of data gathered with touchscreen DSDA indicated the tool might be sufficiently predictive and practical for use in medical practice (results presented in Chapter 7). To ensure predictive validity, I conducted the research within the context of occupational therapy driving clinics. However, GPs provided the impetus for developing Touchscreen DSDA. Therefore, I conducted a study to determine if physicians considered touchscreen DSDA practical and valid for medical practice and if they would use it for screening fitness to drive.

I asked a convenience sample of GPs and medical specialists in Sydney and Melbourne if they would like to trial touchscreen DSDA in their practices for 6-weeks, then participate in a brief semi-structured interview to explore their experiences and the experiences of any staff involved (e.g., practice nurses). The physicians had previously expressed an interest in being involved in touchscreen DSDA development and research and provided their contact details through various avenues of contact including at conference presentations, professional development sessions, and via health networks. Physicians who agreed to be involved were provided with an information sheet describing which patients would be suitable for touchscreen DSDA (i.e., drivers aged 60+ with any diagnosis and drivers aged 18+ with a medical condition that has the potential to impact cognition). GPs were advised not to use the test with patients with little or no driving experience, who did not speak English as their primary language, or who had not completed at least year 7 at school; the test is not standardised for these groups. They were also advised that no patient data would be collected. Physicians were provided with an iPad, stand, headphones and stylus for a minimum of 6-12 weeks at the commencement of the study. Physicians kept the iPad and accessories if they completed the participant interview. Touchscreen DSDA was provided free of charge during the trial. Physicians were not paid for their time.

Physicians from six general and specialist medical practices agreed to participate in the study: general practitioners (2), a consultant rehabilitation physician (1), neurosurgeons (2), and a geriatrician (1). Only one GP and one medical specialist used touchscreen DSDA with their patients and completed the interview. The other four physicians did not trial touchscreen DSDA with any of their patients although provided with additional time (8-12 weeks). One GP reported being too busy to implement the app. Two medical specialists reported being very motivated to trial the app but did not do so within the time frame and study closure date. The remaining medical specialist did not provide a reason for not implementing the app. The study was eventually discontinued because I could not recruit enough participants and therefore collect sufficient data. However, findings from the two interviews that were conducted, raised important issues that could be further explored in future research.

Method

The University of Sydney Human Research Ethics Committee approved this study. Participants provided written, informed consent prior to study enrolment.

Participants

One participant was a general practitioner based in a small medical practice in Melbourne, Australia. She was not the business owner. The other participant was a consultant rehabilitation physician with his own medical practices both in Sydney and a regional area of Australia, where he resided. He also conducted hospital rounds.

Procedure

This qualitative study was conducted over an 18-month period. No patient data were collected. Physicians participated in a 20-minute semi-structured interview after using touchscreen DSDA in their medical practices for at least 6 weeks. Two team members conducted the interviews, taking detailed notes and transcribing each interview. A female team psychologist conducted the GP interview in the GP's office. I conducted the other interview over the phone because the Sydney physician lives in a regional area. We used an open-ended interview guide to ensure consistency of wording. The 12 questions were designed to explore frequency of touchscreen DSDA use, type of patients tested, administration processes, physician and patient experiences, whether screening results were trusted, and how test results were applied.

Data Analysis

Because only two physicians participated in the study there was insufficient data for conducting the planned structured inductive thematic analysis procedure described by Braun and Clarke (2009). Therefore, I present findings from the two interviews separately for each participant. I then discuss findings that were consistent across the two.

Results

The consultant rehabilitation physician trialled the app with four patients aged 18 to 65 with potential cognitive impairment due to stroke, age-related changes, multiple sclerosis, and cerebral palsy. He wanted to use the app with a fifth patient aged in his 70's and diagnosed with hypoxic ischemia and short-term memory loss, but was concerned the results might indicate the patient was fit to drive, whilst he and the patient's son were trying to discourage driving. He was concerned the test might give the patient false hope. The physician reported that one of the four patients (diagnosed with a stroke) found it hard to understand the test instructions and took 20 minutes to complete the screen, whereas the other patients had no difficulty with self-administration. The physician reported he was not concerned about the additional administration time but rather that the patient kept looking to him for re-assurance and he had to state: "I can't help you. You just do what you think is best".

The physician reported he still wanted to refer some patients for an occupational therapy driving assessment even though they passed the screen "just to be sure". For example the stroke patient who was slow to complete the test passed, but the physician referred him for a driving assessment, regardless, due to the difficulty the client had with self-administration. The physician reported concern that the test was too easy (as all his patients passed) but stated he would continue using the app because touchscreen DSDA gave him an objective and consistent method of determining fitness to drive. He stated, “sometimes the outcome was different from what I expected but it is still probably helpful…It helps the patient see how they are going…and gets them thinking about return to driving”. The physician was considering referring to an occupational therapist (without driver-training) in the private hospital to administer the test but stated that for his public hospital ward rounds: “I wouldn’t get the registrar to do it on the ward or the practice nurse. I would keep doing it myself. I like watching how they do the test: how they react and how it works for them”. The physician reported he was considering administering touchscreen DSDA to patients on hospital admission and discharge, as an outcome measure unrelated to predicting driving performance.

The second interview was conducted with a GP in Melbourne who had tried touchscreen DSDA with five patients. She reported her first assessment of a dementia patient (MMSE 23/30) had not gone well. The patient had a high level of education (PhD) and was very experienced with a desktop set up. He failed touchscreen DSDA and “got offended” and was “very resistant to the fact that he had failed”. The patient reported the GP had not given him sufficient set-up instructions. He subsequently wrote a complaint letter, stating touchscreen DSDA tested short-term memory rather than driving skill and attributed his poor result to lack of familiarity with an iPad. The GP advised the patient’s medical specialist of the touchscreen DSDA outcome and advised the patient’s wife to arrange an on-road driving assessment, but she was not confident they followed this advice. The GP reported this experience taught her that the medical practice needed documentation to support the test, such as a consent form describing potential outcomes so the patient was aware of what was at stake. She stated, “I was probably a little bit naïve sailing in there with my new little test”. She reported this experience was a good but painful and provided a tough lesson:

> By the time we had that first experience we were seeing this as our job to negotiate this gently rather than just tear in there with the results: and we saw it as a much more complex issue than when we started.

The second patient the GP tested had diabetes and suspected dementia. This patient spontaneously told the doctor he had a lot of scratches on his car and was reversing into things constantly. The patient failed the test then became very abusive (as he had been related to other matters in the past). He also wrote a complaint letter stating he had not been given sufficient instructions, had been tricked, and did poorly because he was not experienced with iPad. The doctor reported both she and another staff member had in fact given detailed instructions and informed the patient he would need to do an occupational therapy driving test regardless of the outcome. The GP attributed both patients’ blaming the poor result on lack of familiarity with an iPad as “bluster” and “an excuse”.

The GP anticipated she had lost one of the patients: “I am sure he is driving very badly and doesn’t want to know about it so he’s avoiding us”. The interviewer asked the doctor how she would have managed this conflict in the past. She replied that she would have referred the patient for a driving test and been faced with the dilemma: “Am I going to report this to VicRoads or am I just going to negotiate this? It has to be one of the hardest bits of general practice”. The GP described a third patient who needed driver screening but had not agreed to do touchscreen DSDA. The GP stated, “I am now at the point where – do I send him to VicRoads anyway? But his elderly wife is relying on him and it’s just a disaster”. She lamented that public transport was very poor in her area.

The GP reported she had learned from these two negative experiences and subsequently developed a consent form, and determined to be clearer when explaining why the test was being administered and potential consequences. She stated, “So I was much more proactive. After that we didn’t have any trouble...We learnt the hard way”. She was very relieved that following patients did well on touchscreen DSDA and could clearly manage self-administration. The third attempt to administer touchscreen DSDA was an uneventful and positive experience. The patient passed: “He became engaged with the process of doing it even though it was unfamiliar territory, so we could see that someone who had not got dementia could actually suss their way around it as a naïve person.”

Touchscreen DSDA was administered to two subsequent patients referred by the police due to unsafe driving. The police also required both patients to do an occupational therapy on-road assessment. One patient was a 70-year-old female who lost concentration and drifted into a car (no formal diagnosis). The GP reported the patient’s positive experience doing touchscreen DSDA restored her confidence in driving and she went on to pass the occupational therapy driving test. The interviewer pointed out that the patient had actually fallen into the further testing category. The GP reported this was only because she sometimes “forgot she could touch the screen”, but the GP was confident she was safe to drive. The second patient referred by the police was a dialysis patient found slumped in his car after fainting, due to a series of medical and personal issues. The GP stated she thought, “Aha, I’ve got just the thing for you” and was pleased that the patient passed touchscreen DSDA. Fitness to drive reflected medical issues for this patient but the GP found touchscreen DSDA useful for determining if the patient was “concentrating”.

The GP administered the app on the first three occasions then asked her practice nurse to administer it for the remaining patients. She reported patients needed a very quiet, separate room so they could concentrate (whereas the Sydney physician reported no difficulty administering the app in a busy medical practice setting). The GP described finding assessing patient fitness to drive very stressful. She stated, “It is probably the hardest thing to do in all of general practice – talking about their driving. It is such a sensitive, tightly-held thing”. She was concerned about the risk of losing a relationship with the patient and losing a customer. She stated, “We need legislation and mandatory reporting”. The GP lamented the lack of tools for determining fitness to drive despite this being a necessary part of their role: “All we can do is go behind their back and report it to VicRoads [the state licensing authority]”. The GP planned to continue using touchscreen DSDA in the group private practice:

> We are planning to introduce this as a standard part of our general practice 75+ assessments but this made me realise that without any backup from the government it’s just really hard. The people who need it are the hardest ones to do it with.

The GP stated her ideal situation would be administering touchscreen DSDA within the medical practice, with government funding. Alternatively she suggested the licensing authority could administer the test to everyone age 75+ and doctors could access the result via a password-protected link online. The GP stated, “this [touchscreen DSDA] helped me feel more supported but getting them to do it – the ones most likely to fail are the ones most likely to be resistant. So I feel a bit uncertain how to proceed.” The GP said she would still refer patients who failed touchscreen DSDA for an occupational therapy driving assessment because they would otherwise not accept the result. She reported feeling confident that patients who passed the test could avoid expensive on-road testing and noted the test administration was a positive experience for both her and the patients when they passed. The GPs concluding remarks were, When you've got someone with dementia, you are never going to get to that point of insight. So, you're always up against it. Even demonstrating to them that they failed didn't invite them to reflect. They just got defensive. So, that's also a lesson in that it is always going to happen. So, ok, what are we going to do to prepare for that? We've got to get them off the roads, eventually.

Discussion and Implications for Future Research

The data collected only represented the views of two physicians. Therefore, the findings cannot be generalised. However, data gathered identified some common areas of concern consistent with those expressed by physicians in the literature, requiring further exploration in research.

Lack of support. GPs consistently report a lack of clarity and support from policy makers and licensing authorities regarding driver-screening procedures (Braekhus & Engedal, 2009; Jang et al., 2007; Marshall et al., 2012; Omer et al., 2014; Sims et al., 2012). Consistent with this, the Melbourne GP express significant concern regarding the lack of "back up" from the government, licensing authorities, and general practice colleagues. She wanted more legislation around driver screening so there would be clear processes that the patient must follow, implemented by an external body. For example she suggested mandatory administration of touchscreen DSDA in licensing authority offices for drivers aged 75+ (there is currently no aged based testing of drivers in Victoria). This is consistent with the findings from the survey of 200 Australian GPs described in Appendix C: some GPs reported their use of touchscreen DSDA would depend on whether it was accepted by licensing authorities and insurers. They were concerned patients would not do the test unless it was mandated (see Appendix C).

Results of the two interviews, and the difficulty recruiting physicians to trial touchscreen DSDA, suggests that physicians may require more support implementing touchscreen DSDA into their practices, such as legislation and policy support from policy makers and license authorities; training in how to administer the test and interpret the results; and, documentation to support administration (i.e., patient information sheets and consent forms). The Melbourne GP found that the information sheets and consent forms she developed reduced patient concerns and conflict. Therefore, it may be beneficial to develop standard patient information sheets to accompany touchscreen DSDA, which can be adapted for the needs of individual medical practices. Further research is recommended regarding whether doctors require touchscreen DSDA administration training in addition to the touchscreen DSDA in-app administration manual because both participants misinterpreted test results for one of their patients. For example, the Melbourne GP considered that her patient who received the ‘further testing’ result had done well and passed. The Sydney physician administered the app to a patient without driving experience as a cognitive screen. He did not seem aware that touchscreen DSDA was not standardised for patients without driving experience: although this is stated in the administration manual and on the automatically-generated report forms. Pearson offers individual support and test administration training webinars for GPs but these findings suggest it may be important to explore additional ways of training and supporting GPs in test administration.

Managing patient conflict. Surveys consistently show that physicians are concerned about telling patients they need driver screening and then communicating the result to patients when they fail (Braekhus & Engedal, 2009; Classen et al., 2016; Jang et al., 2007; Jones et al., 2012; Marshall et al., 2012; Omer et al., 2014; Sims et al., 2012). Physicians report concern that withdrawing driving damages the patient-doctor relationship; may result in losing a patient; and significantly reduced quality of life for the patient (Braekhus &

Engedal, 2009; Classen et al., 2016; Jang et al., 2007; Jones et al., 2012; Marshall et al., 2012; Omer et al., 2014; Sims et al., 2012). The Melbourne GP describing feeling upset regarding the negative reactions of two patients who failed DSDA and lacked confidence regarding how to manage patient reactions: both patients continuing to drive despite her on-road assessment recommendation. The Sydney physician did not describe any concerns about withdrawing driving privileges but did describe deciding whether or not to administer touchscreen DSDA based on how he perceived the patient would react to the outcome.

The Melbourne GP implemented support systems around touchscreen DSDA testing to reduce patient conflict including providing more information about the testing procedure and the implications. She found this significantly reduced conflict for future patients. Whilst there is an in-app consent form for the patient or doctor to complete prior to each administration and a patient outcome report that is given to the patient on completion of the test, these findings suggest additional supports are needed. For example, patients may benefit from a generic information sheet that describes the test background and potential outcomes in detail. Further research is required to determine whether use of touchscreen DSDA will be advantageous for doctors in helping them manage patient conflict when addressing fitness-to-drive screening.

Trusting touchscreen DSDA results. The predictive validity study described in Chapter 7 indicates touchscreen DSDA has sufficient accuracy to be used to by physicians to determine the need for further on-road assessment. However, both physicians raised issues around trusting the test results. The Sydney physician preferred to continue referring for on-road assessment regardless of the touchscreen test result just to be sure. The Melbourne GP seemed willing to accept the results as long as they agreed with her expectation and clinical judgement. She noted a gender difference between males and females in acceptance of driving cessation: “Very few men give up their licenses voluntarily. Most of them are forced to…Women volunteer a lot of the time”. The GP considered that her male patients who failed touchscreen DSDA would still want to undergo on-road testing because they would not accept that the test result related to their actual driving performance. She also noted a difference between her male and female patients’ capacity to self-administer the iPad application; women performing better. However, the gender-based DIF analysis performed in the predictive validity study described in Chapter 7 indicated no statistically-significant difference in performance between males and females in the DriveSafe subtest.

The Sydney physician was not certain he could trust the results with two patients who passed and still wanted to refer for an on-road assessment “just to be sure”. He was concerned that the additional time taken by one patient may indicate difficulty with driving and was concerned that another patient diagnosed with multiple sclerosis may have difficulty with the physical aspects of driving (a valid reason for proceeding with on-road assessment). These findings are preliminary and it is not surprising that both physicians were uncertain regarding trusting results since they were unfamiliar with the test, used it with a low number of patients, and all of the Sydney-based physician’s patients passed. Further research is required to investigate whether doctors will trust touchscreen DSDA recommendations, how they are going to manage results that disagrees with their expected outcomes, and in what circumstances they would override the result.

DSDA feasibility for medical practice. Both physicians found DSDA feasible for medical practice in terms of administration time, ease of use, location of testing, and cost. No practical concerns were reported. The Melbourne GP reported: “This [touchscreen DSDA] helps me feel more supported”. Both physicians intended to continue using touchscreen DSDA. The Sydney physician described it as a “handy”, objective and consistent method of assessing fitness to drive and felt it was a good tool for providing patients with feedback on how they were going. Both physicians intended to continue referring for occupational therapy on-road assessments and to administer touchscreen DSDA in conjunction with other health professionals (i.e., the practice nurse or non-driver trained occupational therapists). This suggests administration of touchscreen DSDA may in fact increase referral to generalist and specialist occupational therapy services as physician awareness of issues related to fitness to drive are raised. This evolution was particularly evident for the Melbourne GP.

The fact that only two doctors successfully implemented touchscreen into their clinics highlighted the difficulty of changing current health professional clinical practice. This may be due to time, cost and training limitations, and suggests some doctors may prefer to outsource touchscreen DSDA test administration. One limitation of the study was that doctors were not compensated for their time in administering the application and participating in the interview. It is recommended that doctors be paid for their time in future studies to potentially increase participation. It is important to conduct further research to identify barriers to adopting the fitness to drive screen, particularly as GPs consistently report they need such a test. The preliminary finding from the two interviews suggests some physicians may prefer to refer to occupational therapists to do both off-road screening and on-road testing. Further investigation is required whether generalist and driver-trained occupational therapists, rather than GPs, should be the target population for administration of touchscreen DSDA.

CHAPTER 9

Summary and Conclusion

The studies included in this thesis contributed to the conversion of original DSDA into a clinical, touchscreen test of cognitive fitness to drive. Results of the study of the psychometric properties and predictive validity of data gathered with touchscreen DSDA, presented in Chapter 7, indicated touchscreen DSDA was sufficiently accurate to be used as a first-level screen to predict on-road performance. Specificity of touchscreen DSDA was 86% and sensitivity 91%. The positive predictive value was 83% and negative predictive value, 92%. Overall accuracy of classification was 88%; 11% of patients were classified as "required further testing". The data showed evidence of construct validity and reliability. Now, revisiting the criteria for the design of a suitable GP fitness-to-drive screen identified in the literature, summarised in Chapter 2, results of the research presented in this thesis indicates touchscreen DSDA meets these criteria (see Table 9.1).

Table 9.1

Comparison of Touchscreen DSDA with GP Fitness-to-Drive Screen Design Criteria

GP driver screen criteria	Criteria references	Touchscreen DSDA
Evidence-based cut-off scores	Bédard et al., 2008; Molnar et al., 2006	Yes
2 cut-off scores (generating a "further testing" category)	Bédard et al., 2008; Langford et al., 2008; Molnar et al., 2006	Yes
Sensitivity, specificity, PPV and NPV reported and above 80%	Asimakopulos et al., 2012; Bédard et al., 2008; Langford et al., 2008; Molnar et al., 2006; Weaver & Bédard, 2012	Yes (specificity 86%; sensitivity 91%; PPV 83%; NPV 92%; overall accuracy of classification 88%)
Criterion measure: a standardised on-road assessment	Classen et al., 2015; Classen et al., 2010; Hargrave et al., 2012; Laycock, 2011; Wheatley & Di Stefano, 2008; Yale et al., 2003	Yes
Small percentage (e.g., 10-20%) of patients classified "further testing"	Bédard et al., 2008; Weaver & Bédard, 2012	Yes (11%)
Face valid for the driving task	Crundall, 2009; Dalchow et al., 2010; Weaver & Bédard, 2012	Yes
Brief (i.e., 10 minutes or less)	Asimakopulos et al., 2012	Yes (around 10 minutes)
User friendly for GPs (i.e., portable, no unique testing consoles, no training, and simple)	Asimakopulos et al., 2012; Dalchow et al., 2010; Fildes, 2008; Molnar et al., 2006	Yes

Nevertheless, results of the discontinued study conducted to determine if touchscreen DSDA was feasible for medical practice, presented in Chapter 8, indicated potential barriers to GPs adopting the test. This is early, incomplete research that examines the views of only two doctors but raises important issues that require further investigation. Physicians in the study reported finding touchscreen DSDA user-friendly and practical for medical practice. However, the findings suggested physicians required further supports around driver screening for touchscreen DSDA to be successfully adopted: for example, additional training, standardised information for patients about the test and potential implications, legislative and policy support, and funding.

The difficulty of recruiting doctors for the study and getting them to try the application with their patients, may indicate that doctors will prefer to refer to external providers for touchscreen DSDA administration. It may be that occupational therapists will be the primary users of touchscreen DSDA and that use of the test among generalist occupational therapists may strengthen the role of occupational therapists in driver-screening and addressing client community mobility needs, as recommended in the Occupational Therapy Practice Framework (American Occupational Therapy Association, 2014). Findings from the discontinued study suggest use of touchscreen DSDA may raise physician awareness of patient fitness to drive concerns, leading to an increase referral rates to occupational therapist: both for off-road administration of touchscreen DSDA by generalist occupational therapists and for on-road testing by driver-trained occupational therapists. Further research is required to identify barriers to physicians adopting touchscreen DSDA and to verify these assumptions.

Findings from the two studies conducted to convert original DSDA into a touchscreen test, presented in Chapters 5 and 6, provided processes and guidelines that healthcare researchers converting standardised assessments for digital administration could follow. We found that a combination of literature review and usability testing, following an iterative process concurrently with app design, programming and evaluation, ensured the test was user-friendly for practitioners and patients, and avoided costly design errors. Additionally we found that our success relied on a clear understanding of the existing knowledge base for measurement of the desired construct and critical test mechanisms, so these are not lost when converting test variables to a new mode of data collection, scoring and interpretation.

Limitations and Future Directions

The research presented in this thesis has several limitations. The criterion measure used in the study of touchscreen DSDA psychometric properties and predictive validity was a standardised on-road assessment. P-Drive (Patomella et al., 2010) was the on-road assessment tool selected as it allowed consistent scoring of driving performance across variable routes and driving conditions and when using vehicle modifications. Despite use of a standard on-road assessment tool, driving in a real-world environment cannot be completely standardised, as drivers must negotiate rapidly-changing situations. We conducted the on-road assessments in diverse locations around Australia and New Zealand. Whilst all driving clinics were based in capital cities and measures were taken to standardise routes, traffic conditions and physical road features would have varied amongst sites to reflect clinical needs, evaluator skill, time of day, traffic density, and weather conditions.

In this study, the same therapist conducted both the touchscreen test and the on-road assessment. Therapists were asked not to observe the client complete the self-administered components of the test but we could not control this. However, therapists were blind to the results of touchscreen DSDA since no scoring categories or cut-off scores existed at that point. Using a separate therapist, blind to the off-road assessment results, might have improved validity of finding but this was not possible within reasonable cost for the 10 clinics.

The need for cut-off scores for specific diagnoses (e.g., dementia or stroke) may be questioned. However, when we examined diagnostic groups separately there was no discernible difference in spread of patients related to diagnosis. Confirming predictive validity of touchscreen DSDA with specific diagnostic groups is an important area for future research. Touchscreen administration of DSDA allowed the collection of precise digital data, which provided scope for future research in areas such as examining differences in performance amongst specific diagnostic groups (e.g., dementia and traumatic brain injury); for people with vision impairments (e.g., visual field loss or glaucoma); or for neurological patients with a hemi-spatial neglect. Touchscreen DSDA may provide important insight regarding observation, attention, touch patterns, and memory recall for these or similar groups. Research regarding the success of touchscreen DSDA as an outcome measure for evaluating the effectiveness of interventions (e.g., pre- and post- neurosurgery) or tracking changes in cognitive capacity over time (e.g., for degenerative conditions such as dementia) would address a knowledge gap in these areas and increase the practical application of the tool.

Touchscreen DSDA has only been standardised for people who speak English as a first language. Standardising administration of the test for drivers in countries where English is not typically spoken is an important area for future research to allow broader access. Touchscreen DSDA currently reflects driving conditions for regions where drivers drive on the left side of the road (e.g., Australia, New Zealand, Singapore, UK, and Hong Kong). The test needs to be standardised for countries where drivers drive on the right side of the road.

This would involve re-photographing the images for the relevant cultural contexts and conducting research to standardise touchscreen administration for drivers in these countries.

Touchscreen DSDA was released to the Apple Store in April 2015 and is now increasingly used by health professionals in Australia and New Zealand in clinical practice and research. Researchers (Dickerson et al., 2017; Gibbons et al., 2017) advise that there is no one best tool for screening fitness-to-drive for all drivers and several should be used. Touchscreen DSDA is one test that can be useful.

Appendix A

DriveSafe DriveAware Administration Manual

I wrote the following technical manual to provide touchscreen DSDA administrators with instructions regarding how to set up and administer the screen and how to interpret the results. Dr Haijiang Huang, Senior Psychometrician at Pearson US, assisted in writing the Chapter 5 statistical analysis and results section. The DriveSafe DriveAware Administration Manual was published by Pearson in-app and online on release of touchscreen DSDA to the Apple Store in April, 2015. The reference for this manual is:

Cheal, B., & Kuang, H. (2015). DriveSafe DriveAware for Touch Screen Administration Manual. Sydney, Australia: Pearson Australia Group Pty Ltd.


DriveSafe DriveAware for Touch Screen
A Screening Tool for Cognitive Fitness to Drive
DSDA
DriveSafe
DriveAware
Administration Manual
Beth Cheal and Haijiang Kuang
ALWAYS LEARNING
PEARSON

DriveSafe DriveAware for Touch Screen

A Screening Tool for Cognitive Fitness to Drive

Administration Manual

Beth Cheal
and Haijiang Kuang

Table of Contents

About the Authors

Beth Cheal, Touch Screen DriveSafe DriveAware Project Manager, is a Driver Trained Occupational Therapist and licensed Driving Instructor who has completed a Bachelor and Master's of Applied Sciences in Occupational Therapy at the University of Sydney (USyd). The conversion of DSDA into a screening tool for cognitive fitness to drive for use by doctors and other health professionals was conducted by Beth Cheal as a PhD (USyd) under the supervision of Professor Anita Bundy, Chair of Occupational Therapy at the University of Sydney and co-author of the DriveSafe and DriveAware computer version. Beth Cheal has over 15 years of experience working as a driving assessor and educator via the University of Sydney and her practice Rehab on Road.

Haijiang Kuang, Senior Psychometrician at Pearson Clinical Assessment, has completed a PhD in Statistics and Experimental Design and a Masters in Counseling and Student Psychology (University of Minnesota), and a Bachelor in school education. She has over fifteen years' experience in psychometrics and as the lead psychometrician at Pearson, has successfully accomplished many key projects for the publication of achievement and ability tests, behaviour scale tests and personality and clinical inventories. Haijiang provided psychometric consultation and data analysis to the DSDA project and worked with the team in writing and editing chapter 5 of the DSDA manual.

Acknowledgements

The touch screen DSDA was developed by the Research and Development team, Pearson Clinical Assessment, Australia. Publication of this product has been a success due to the dedication and hard work of several groups to whom we would like to acknowledge and express our sincere appreciation.

First, we thank the staff and research participants from the driving clinics across Australia and New Zealand who supported the development of the DSDA touch screen version. Individuals within these clinics willingly shared their expertise, collected data, and provided valuable feedback on the test. Their commitment to developing better clinical tools for use in driver assessment and rehabilitation is indicative of their professionalism in clinical practice.

- All Ages Occupational Therapy, Victoria
- Australian Unity Retirement Living Services, Constitution Hill, Sydney
- Calvary Health Care Sydney Driving Program, Sydney
- Rehab on Road, Sydney
- St. Joseph's Hospital Driving Program, Sydney
- Evolution OT, Brisbane
- Home & Driving Occupational Therapy Services, Perth
- Independent Living Centre WA, Perth
- Kevin O'Leary & Associates, Wellington
- Organisation of Therapy and Rehabilitation Services (OTRS), Hamilton & Auckland

We acknowledge the Pearson project team members Dr. Fiona Brown (Project Director), Sabrina Mooney (Business Services Manager), and Melinda Cooper (Occupational Therapy Consultant). Each of these individuals invested significant time, expertise, and hard work to bring the DSDA touch screen version to life. We also acknowledge the following Pearson staff for their contribution to the development and design of this tool: Christopher Wilson, Nicholas Bendrups, Katherine Ritchens, and Ivan Tang.

Dr. Jianjun Zhu (Director of Psychometrics), Haijiang Kuang, (Senior Psychometrician), and Xuechun Zhou (Psychometrician) deserve our special recognition for their expert statistical consultation and analysis conducted throughout the project. We would like to thank Dawn Dunleavy (Managing Editor) for editing this manual.

Both Anne Madden (Vice President Australia, New Zealand, Singapore), and Lois Gregory (Director, Clinical Content Development for North America, Australia/New Zealand, and the UK) deserve our special thanks for entrusting this project to us and for their unwavering support throughout. We are also grateful to Dr. Larry Weiss (Vice President of Clinical Test Development) and Dr. Aurelio Prifitera (Managing Director, Clinical Assessment), for their permission to conduct the conversion of DSDA to touch screen.

We also acknowledge Professor Anita Bundy, University of Sydney, for her advice regarding research design, conversion from computer to touch screen testing, statistical analysis and other critical aspects of the project; and Dr. Ann-Helen Patomella, Karolinska Institute (Sweden), for her advice regarding research design and standardisation of the on-road assessment process. We thank Dr. Patomella for allowing us to use the P-Drive (Performance Analysis of Driving Ability), of which she is the author, and for assisting in the training of research site staff. We thank Joan O'Donnell, Occupational Therapist and PhD Candidate at the Australian School of Advanced Medicine, Macquarie University, for her assistance in the development phase of the DSDA touch screen version.

Last, we would like to acknowledge the high quality work of Very Livingston, the company responsible for design and programming of the application.

Beth Cheal
Project Manager

Dr. Nicki Joshua
Project Director

Chapter 1.
Introduction

Purpose

DriveSafe DriveAware (DSDA) is a cognitive screening tool that measures a driver's awareness of the driving environment and their own abilities related to driving. The test can be used when ability to manage the cognitive aspects of driving may be impaired by a medical condition, injury, or the ageing process. General practitioners and other health professionals can use the touch screen DSDA to accurately predict which patients with a cognitive impairment require an on-road assessment. Occupational therapists who specialise in driver assessment and rehabilitation can also use the touch screen DSDA as part of their off-road assessment of fitness to drive.

The touch screen DSDA is used to assess two areas that are critical for safe driving—global awareness of the driving environment and awareness of one's own abilities related to driving. Awareness is necessary for a driver to be able to monitor his or her own performance and employ compensatory strategies where necessary (e.g., avoid driving at night or on unfamiliar roads). Awareness is, therefore, critical for safe driving (Kay, Bundy, & Clemson, 2009b).

The touch screen DSDA consists of three subtests:

1. DriveSafe
2. DriveAware
3. Intersection Rules (Optional)

The touch screen DSDA was designed so that most items can be self-administered by the majority of patients. The DriveSafe subtest, two of the seven DriveAware subtest questions, and the optional Intersection Rules subtest, can all be self-administered. The remaining five questions from the DriveAware subtest are administered by a general practitioner or health professional in a brief interview.

As part of the development research, use of the touch screen DSDA was examined in a prospective study with 134 older (60 + years) and/or cognitively impaired drivers across Australia and New Zealand (see Chapter 5 for study results). The statistical analysis undertaken within this study revealed the Intersection Rules subtest (which was included as part of DriveAware in previous versions of the test) should be removed from the DriveAware subtest and included in the test only as an optional, separate addition. This optional subtest provides a brief screening of knowledge regarding right-of-way rules at intersections and qualitative information, such as ability to follow instructions and plan responses. This subtest can be skipped to reduce test time.

Testing Time

Test time is approximately 10 minutes, depending on the patient and whether or not the optional Intersection Rules subtest is included. The DriveAware clinician interview takes about 3 to 5 minutes to administer.

In the development research, the DriveSafe subtest and self-administered DriveAware questions took an average of 6 minutes to complete. The optional Intersection Rules subtest took an average of 3 minutes to complete.

Test Components

The touch screen DSDA is currently available only on iPad®. You need the following test components to administer the test as it was standardised:

- iPad with iOS operating system 7 or later (an iPad mini or iPhone® cannot be used)
- DSDA touch screen version
- Tablet stand angled to 20 degrees
- Stylus (optional depending on patient preference and ease of use)
- Headphones (optional, depending on sound distractions in the room, hearing impairment, and patient preference)
- Desk or table
- Suitable upright chair (e.g., office or kitchen chair)

Downloading DSDA and Purchasing Reports

Download DSDA from the Apple® App Store directly to the iPad. Then you (the administrator) must register as a test user with Pearson Clinical Assessment. Follow this link to complete the registration form: https://www.pearsonclinical.com.au/accounts/create. You will receive a welcome email confirming your username and password.

To purchase report usages online, go to www.pearsonclinical.com.au/sessions/create. You can administer the test without report usages, but you will not receive results or a report. If you do not have access to WiFi, telephone the client services team and order additional credits at the following contact numbers:

Australia: 1800 882 385 (toll free)
New Zealand: 0800 942 722 (toll free)

When you log into the test with your DSDA username and password, you are prompted to create a 4-digit PIN. The PIN security ensures that the patient only ever has access to the DSDA test. If the patient tries to access other information within the test after it has begun, the test locks and can be unlocked only by entering the PIN. Never give the PIN to the patient and keep it private. The practitioner administering the test is the only person who can enter the PIN.

Who Can Administer DSDA?

Only general practitioners, medical specialists, occupational therapists, speech pathologists, physiotherapists and psychologists may administer the DSDA and interpret the test results. A practice nurse may administer the test under the direction of a registered medical practitioner.

Who Should Complete DSDA?

DSDA should be administered to patients when cognitive capacity for driving could be impaired by a medical condition, injury or the ageing process. Practitioners should use caution when considering touch screen DSDA administration with patients who have:

- physical deficits that may impact driving,
- reduced literacy or English language skill,
- been diagnosed with a psychiatric condition, and / or
- are learner or beginning drivers.

Patients With Physical Deficits

All patients with physical deficits that could impact ability to operate car controls should be referred for an occupational therapy driving assessment. Vehicle modifications or licence and vehicle restrictions may be required for these drivers. However, the DSDA test may still provide useful information on the new learning capacity for this group if cognitive impairment is also present.

If a patient has an upper limb deficit that impacts their ability to use touch screen technology, they may use a stylus or the Administrator-Assisted Method may be used.

Patients With Reduced Literacy or English Language Skills

The touch screen DSDA was validated with patients who attended school to year 7 or higher. The touch screen DSDA is easier to administer to patients who have reduced English language skills, reduced literacy, or a medical condition impacting communication (e.g., aphasia) as compared to the computer version. The test has not, however, been validated for these groups. Results may be used qualitatively but should be interpreted with caution.

If you indicate that the patient does not speak English at home, the following notation appears on the clinical report: "Drivesafe DriveAware has been standardised for people who speak English as a first language. If English is not spoken at home the test results may be used qualitatively, but should be interpreted with caution".

Patients Diagnosed With a Psychiatric Condition

DSDA was developed to determine a patient's ability to manage the cognitive aspects of driving. Psychiatric conditions may impact cognitive functioning, however, often the primarily difficulty with driving is due to disturbances of emotion, behaviour, and perception. If you indicate that a patient has an identified psychiatric condition, the following statement appears on the report: "DriveSafe DriveAware has not been standardised for drivers with psychiatric conditions. The purpose of the test is to predict cognitive capacity for driving. Results may be used qualitatively, but should be interpreted with caution".

Learner Drivers

All learner or beginner drivers who have a medical condition that may affect their driving need to undergo an occupational therapy driving assessment if there is concern regarding their fitness to drive. The DSDA test is not valid as a standardised test for this group as prior driving experience is assumed and the test was validated with experienced drivers.

Chapter 2.
General Administration Guidelines

Administration Methods

The touch screen DSDA can be self-administered or practitioner-administered. The touch screen development phase revealed that 9 out of 10 senior adults (ages 65 to 93) were able to successfully self-administer the application. There will always be a small number of individuals who cannot self-administer the test due to a medical condition or difficulty with the technology. The Administrator-Assisted Method (described on page 23) was developed for use with these such individuals. It is recommended that patients who are unable to complete the DSDA touch screen version either independently or via the Administrator-Assisted Method, be referred to a driving clinic for further assessment of fitness to drive.

Rapport / Approach

The goal is to administer the test in as non-threatening a way as possible so the patient can give his or her best effort. Use your clinical experience and common sense to determine the most effective way of establishing rapport. A confident, friendly, and relaxed approach can elicit cooperation and ease uncertainty or anxiety about testing.

Instructions

Administer the test in a professional and unhurried manner, following the instructions given. This manual includes standard phrases that must be used when prompting is required. **The standard phrases appear in bolded, red type**. These are the only words that should be used. The test includes audio and written instructions to enable self-administration.

Rest Periods / Breaks

If the patient requests a break during testing, allow the break after the patient has completed a subtest and not in the middle of a subtest. Stopping testing in the middle of a subtest may impact test validity. The device only saves data upon completion of a subtest. Data is not be saved if a subtest is aborted part way through.

Subtests can be completed on different days or with breaks in between. Subtests can be administered by different practitioners. Because the administration time is short, it is unlikely that breaks will be required.

Discontinuing Testing

If the patient is using the self-assessment method and struggling, allow him or her time to repeat items and practice. If the patient continues to struggle, switch to the Administrator-Assisted Method. If the patient continues to struggle with the test to the point that the results may not be valid, discontinue testing. The test can be discontinued at any time by tapping the top left corner of the screen twice. Enter the PIN number to exit the test.

If the patient is unable to complete the test independently or with the Administrator-Assisted Method, refer him or her to a driver assessment service to determine fitness to drive.

Frequency of Testing (Test-Retest Reliability)

Test–retest reliability has not been examined for the touch-screen version of DSDA, however, a test–retest reliability study was conducted for the computer version of the test (O'Donnell, 2013). Retesting was undertaken at 6 week and 12 month intervals for patients diagnosed with an unruptured aneurysm. The computer version was found to be test-retest reliable at both time intervals (results being prepared for publication). There is no clinical reason to expect a difference in test-retest reliability between the two digital delivery methods, however, this is an area for future research.

Use Caution in Applying the Results

Suitably qualified clinicians are expected to use DSDA as a decision support tool in conjunction with their clinical judgment and other clinical indicators when assessing a patient's capacity to resume or continue driving.

Prior to advising a return to driving, the general practitioner or driver trained occupational therapist must ensure the patient's vision is legal for driving according to the national medical guidelines Assessing Fitness to Drive (2012). These guidelines apply to both Australia and New Zealand and are available online at; https://www.onlinepublications.austroads.com.au/items/AP-G56-13. The general practitioner must also ensure the patient does not have a medical condition that is contraindicated for driving (e.g., seizures; blackouts; some heart conditions). Medical clearance is required for driving for patients who have some long term medical conditions. These notifiable conditions are outlined in Assessing Fitness to Drive guidelines.

Assessing Fitness to Drive indicates that a practical driving assessment may be required for some people to determine fitness to drive (Austroads, 2012). Referral to a driving assessment service is advised when;

- Results of DSDA place the patient in the category 'requires further testing,' and the patient wishes to continue to drive.
- Results of DSDA place the patient in the category 'unlikely to pass an on-road driving assessment,' but the patient still wishes to drive.
- The patient has long-term physical deficits that may impact their ability to operate car controls (e.g., amputations, paralysis, and incoordination). Vehicle modifications are likely to be required.
- Information supplementary to the clinical assessment is required for borderline cases where fitness to drive is not clear.

Occupational therapists conducting driving assessments must be 'driver trained'. This means they must have completed postgraduate training in driver assessment and rehabilitation and have a registration number allocated by the training institution.

Australian driver-trained occupational therapists can be contacted via the relevant state Occupational Therapy Association (OT Australia) website: http://www.otaus.com.au.

New Zealand driver-trained occupational therapists can be contacted via Occupational Therapy New Zealand–Whakaora Ngangahau Aotearoa website: www.nzaot.com.

Setting Up the Testing Environment

Room Set Up

Administer the touch screen DSDA in a quiet room that is free from distractions and interruptions (e.g., an office or clinical treatment room). If the test must be administered in a busy area due to lack of space (e.g., patient waiting area), have the patient use the headphones and seat them as privately as possible (e.g., behind a screen or in a separate area). You should administer the test without family or friends present. If this is not possible, instruct any person accompanying the patient to sit silently and avoid inadvertently helping or distracting the patient.

Table / Chair Set Up

Seat the patient comfortably in an upright chair at a table or desk. Do not place the tablet on the patient's lap. Make sure the desk is free of other materials to minimise distractions. Position the patient away from windows or lights that may cause glare on the tablet screen. Close window coverings or turn off lights if the patient prefers it.

- Seat the patient on an upright chair at a table.
- Place the stand on the table, directly in front of the patient (in their midline) at a 20-degree angle.
- Place the tablet on the stand and make sure the patient can reach it comfortably, without allowing the patient's hand to rest on the screen. This will trigger unwanted responses.
- Adjust the brightness of the tablet image to full to ensure there is no glare on the screen and there is adequate contrast in the image.
- The volume of the tablet should be turned to full.

Note: It is acceptable to turn down volume or brightness of the tablet if requested by the patient for his or her comfort. The patient may adjust the proximity of the stand for comfort.

The goal of the tablet set-up is to ensure that the patient has the best view of the screen for his or her focal length, with good image contrast for the time of day and light in the room.

Use of Headphones and Stylus

Offer the patient the opportunity to use the headphones if they have difficulty hearing or a stylus if they have difficulty responding during the familiarisation process. The decision to use these aids is up to the patient.

Administrator's Position

Sit beside (rather than opposite) the patient on a separate chair during the set up and practice stage. Sit on the patient's non-dominant side so that you can easily access the tablet without touching the patient, in case you need to provide assistance.

DriveSafe

The DriveSafe subtest is self-administered with instructions presented in written and audio form. DriveSafe consists of 10 images of a 4-way intersection (an example intersection image is presented in Figure 1.). Each intersection includes a number of people and vehicles (ranging from 2 to 4 objects in total). These objects are presented for 4 seconds then disappear. For each object presented, the patient is prompted to recall 3 pieces of information:

1. Type of object (e.g., car, pedestrian, couple walking together, truck, or bicycle)
2. Object location
3. Direction of movement

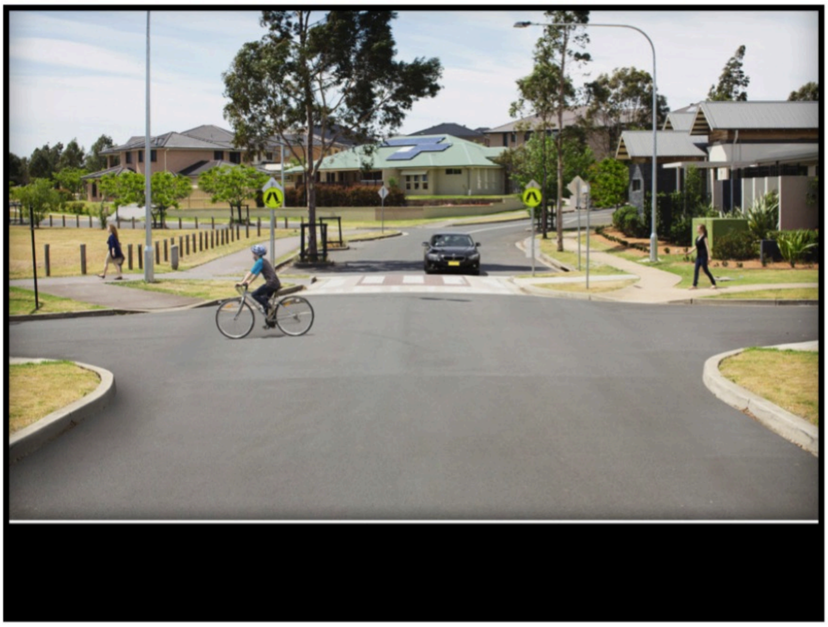


Figure 1. Example DriveSafe Intersection Image

After the objects disappear, the patient is prompted with "touch the screen to show your responses". When the patient touches the screen, a yellow location point appears with an object menu attached (see Figure 2.). The patient selects one of the five possible object icons to indicate the type of object previously observed.

Figure 2. Object Location Point with Object Type Menu

After the patient selects the object type, an arrow appears (see Figure 3.). The patient must drag the arrow in the direction the object was travelling.

Figure 3. Object direction icon

DriveAware

The DriveAware subtest consists of seven questions. The patient rates his or her perceived performance on the DriveSafe subtest in two self-administered questions (see Figure 4.). Self-administered questions are presented via written and audio instructions.

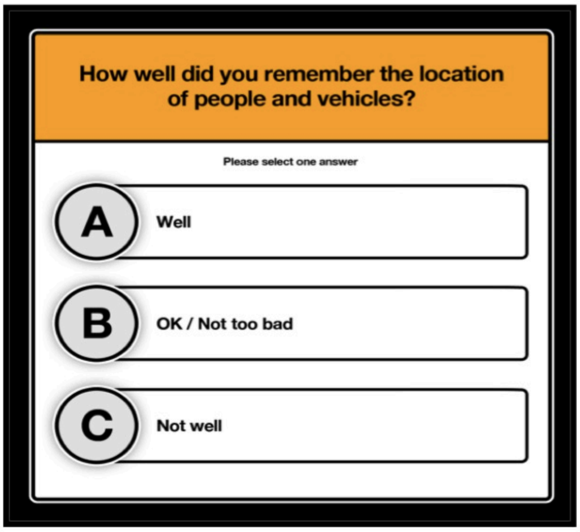


Figure 4. Example DriveAware question

The remaining five DriveAware questions are part of an interview that a general practitioner or health professional conducts at the conclusion of the test. The administrator reads aloud four questions to the patient and enters the responses via the touch screen. The final question requires the administrator to rate their level of concern for the patient's fitness to drive, based on clinical or personal indicators.

Results of the DriveSafe and DriveAware subtests are used to classify drivers in three categories: likely to fail an on-road assessment, requires further testing, and likely to pass an on-road assessment. For more information on scoring and categorisation, see Chapter 4.

Intersection Rules (Optional)

Intersection Rules is an optional, self-administered subtest with instructions presented in written and audio format. The Administrator-Assisted Method could also be used if required. If the healthcare professional selects it for inclusion, the subtest automatically begins after the two DriveAware self-administered questions, but before the remaining five questions included in the brief interview.

The Intersection Rules subtest presents eight intersections with two to four vehicles in each intersection (see Figure 5.). Four intersections have road sign symbols. For each intersection, the patient is prompted with the phrase "tap the vehicles in order". The patient touches the screen to indicate the order in which vehicles in the intersection should proceed, according to the road markings and symbols presented.

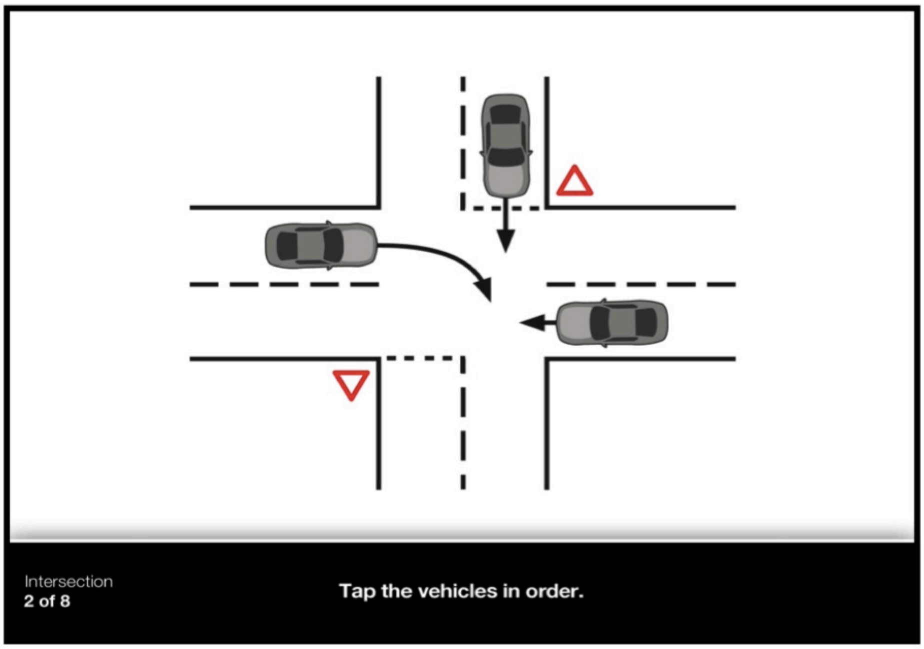


Figure 5. Example Intersection Rules item

Chapter 3.
Administering DSDA

Preparing the Tablet for Testing

The DSDA home screen is shown in Figure 6. The home screen has three columns. Use the left column to search patient records by name, status, or date of testing. Touch 'Patient set-up' to set up a new test. The number of report usages available is shown at the bottom of this left column.

The centre column lists records from newest to oldest and colour codes records based on status. Red indicates the test is ready to start or resume. Yellow indicates the test has been completed, but the report has not yet been created. Green indicates the test has been completed and the report created. The right column lists all records for patients previously tested. You can retrieve old records from this column or by using the search fields in the left column.

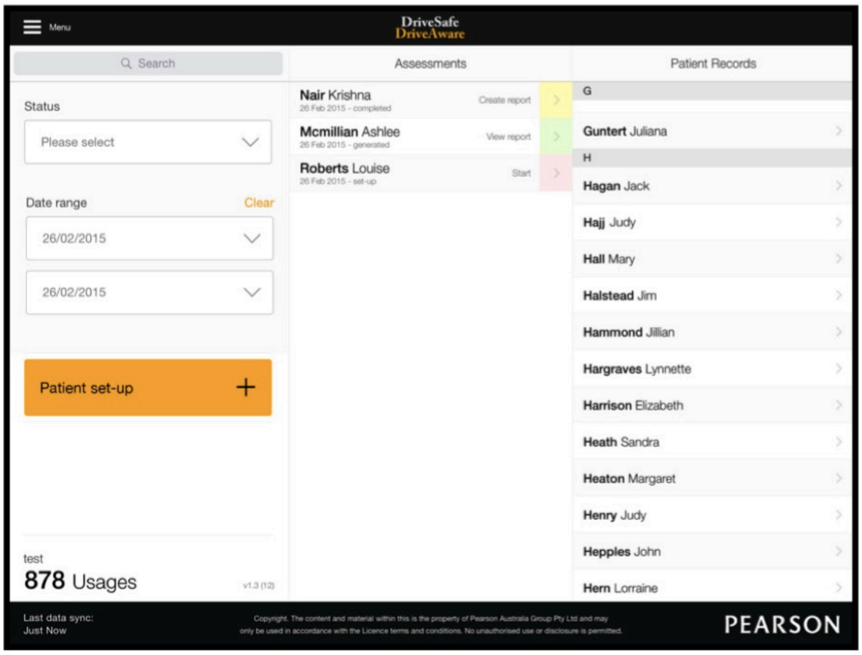


Figure 6. DSDA Touch Screen Home Screen

Set up the patient record and test prior to the patient's appointment. This process takes 1–2 minutes and ensures that the test is ready when the patient arrives for the assessment. Select 'Patient set-up' to start a new test for either a new patient or to re-test a previous patient (see Figure 7.).

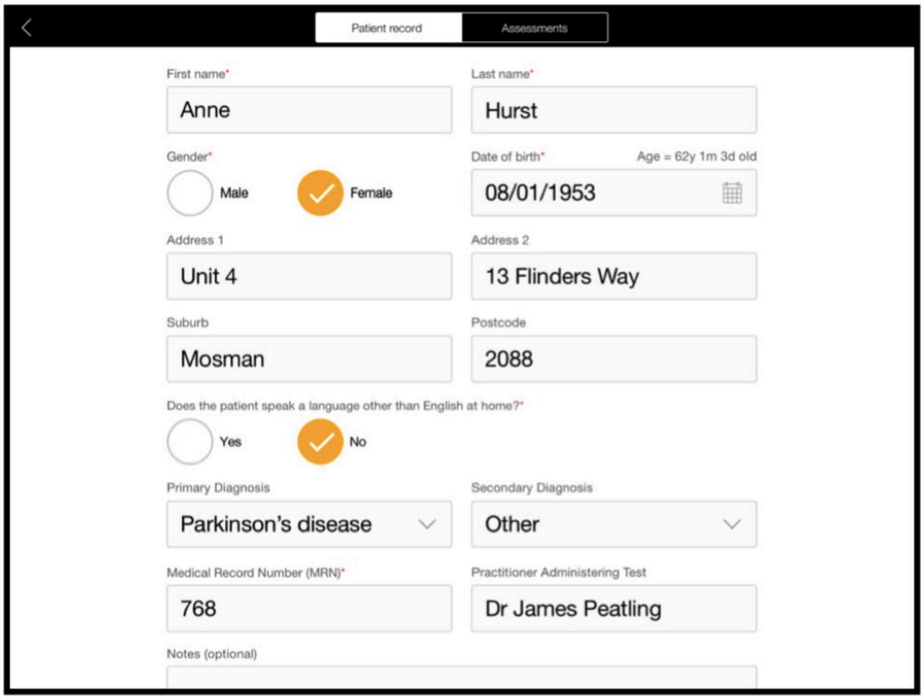


Figure 7. DSDA Patient Set-Up Screen

The administrator must enter the patient details, then select 'Save' or 'Start Test' as preferred. Select 'Start Test' to begin testing. Select 'Save' to return to the home page.

Consent Forms

Prior to beginning the test, either you OR the patient must provide consent. The patient consent form provides a brief summary of the test purpose and content, explains where the information is being sent, and states that the participant may withdraw consent at any time until testing has been completed and results calculated. The patient must select 'yes' in response to "I have read and agree to the conditions". Alternatively you can consent on behalf of the patient by selecting 'yes' to "I have given the patient a verbal explanation of the DSDA test, its procedures and risks and I believe that the patient has understood that explanation".

Administering the Demo and Practice Sections

Use the following procedure for the demonstration and practice sections, for both administration methods.

Test Start

Once consent has been given, the test start screen appears (see Figure 8.).

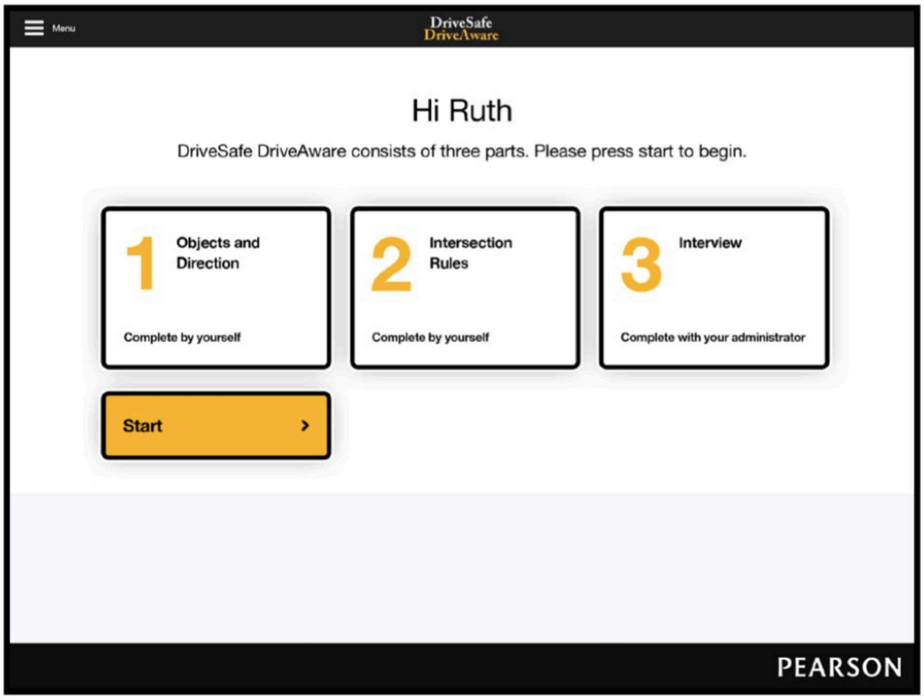


Figure 8. DSDA Patient Start Screen

The patient touches 'Start' when ready to begin. Allow the patient to self-administer the test, following the prompts. The patient is taken through audio and vision checks to ensure the volume is appropriate and the screen is the correct brightness. When the checks are complete, the patient is presented with the DSDA demonstration introduction screen, which describes the basic premise of the test via audio and written instructions (see Figure 9.).

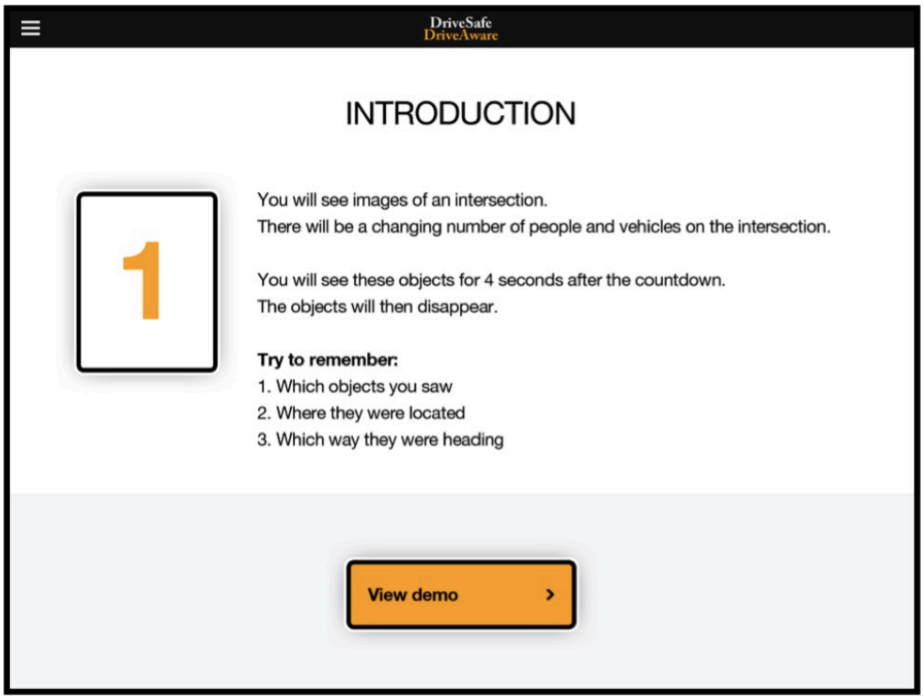


Figure 9. DSDA Demonstration Introduction Screen

The patient begins the demonstration by touching 'View demo'. **Remain in the room until the patient successfully completes the demonstration and practice items.** This enables you to:

1. provide assistance if needed; and
2. decide whether or not the Administrator-Assisted Method is required.

Demo Section

The demonstration section consists of one item to teach the patient the test functions (tap, arrow swivel, object menu and undo). Watch the patient complete the demonstration item (with the red car driving forwards).

- If the patient struggles to move the arrow, say: **touch the arrowhead and move your finger to show which way the car was driving.** You may also demonstrate this, but allow the patient to try it first.

Repeating the Demo Section

The patient can repeat the demonstration item **only once** by selecting 'Repeat demo'. Have the patient repeat the item if he or she is having significant difficulty and you think it may help.

The patient moves from the demonstration to the practice section by selecting 'Start practice' (see Figure 10.).

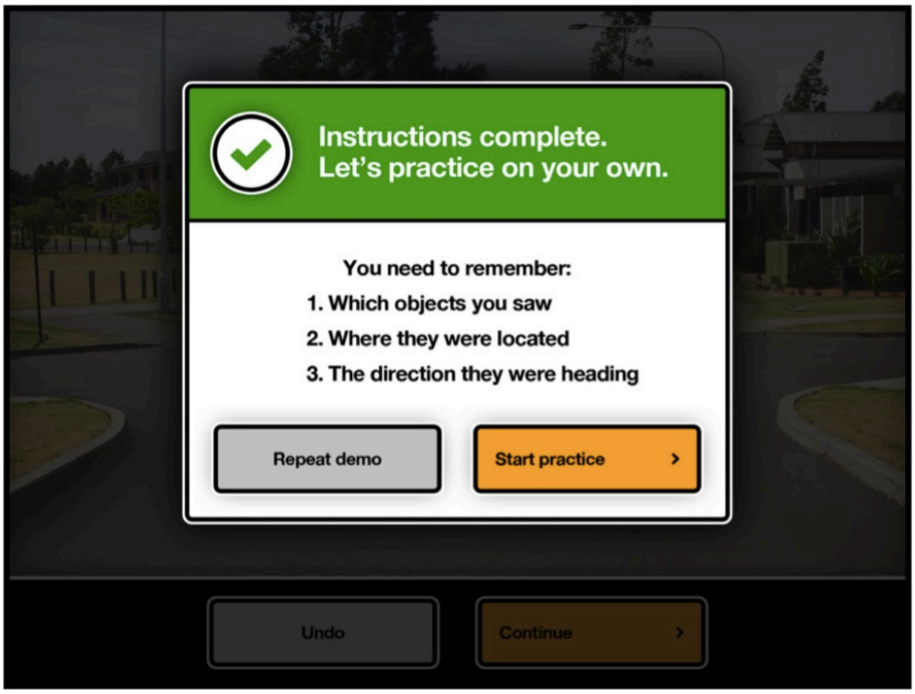


Figure 10. DSDA Practice Start Screen

Practice Section

The practice section of DriveSafe consists of three items. The purpose of the practice section is to teach the patient the basic premise of the test. The items provide various levels of assistance as required via error messages and additional instructions.

The patient must complete each of the three practice items correctly before the test allows them to proceed. The patient can have a maximum of four attempts per item. If the patient fails to answer correctly on the fourth attempt, they receive a message instructing them to see their administrator for help. The development and research phase of the test revealed no improvement in performance beyond four attempts and/or the patient became too frustrated to succeed beyond this point.

Repeating the Practice Section

The patient may repeat the practice section **only once** if you judge this as beneficial in facilitating independence. Repeated practice beyond this was not found to be beneficial in the development and research phase of the test.

If a Patient has Difficulty and Does Not Improve With Practice

If a patient has significant difficulty with the test and does not improve with practice, consider what you observed during the practice section of DriveSafe and ask yourself:

> **Was the patient able to complete at least one practice item without assistance (i.e., indicate object location, type and direction of movement without help)?**

If the answer is 'no' then use the Administrator-Assisted Method outlined on page 23. Figure 11. represents the Administration Method Decision Tree.

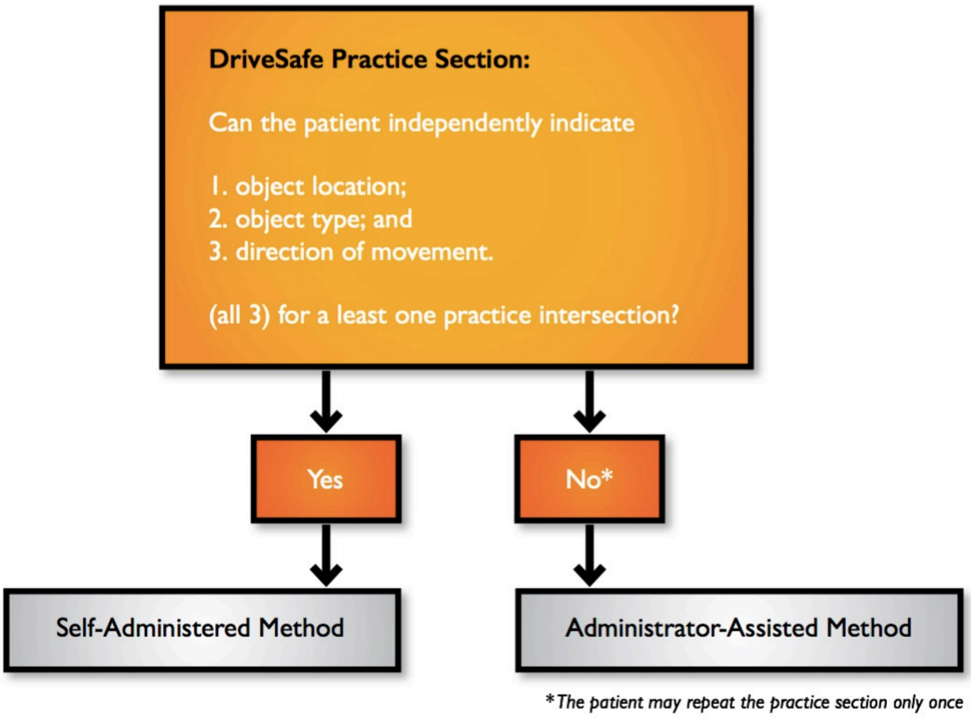


Figure 11. Administration Method Decision Tree

DSDA Self-Administered Method

DriveSafe

After the patient has successfully completed the demonstration and practice sections, the DriveSafe start test screen appears (see Figure 12.). The patient should select 'Start test' when he or she is ready to begin (see Figure 12.).

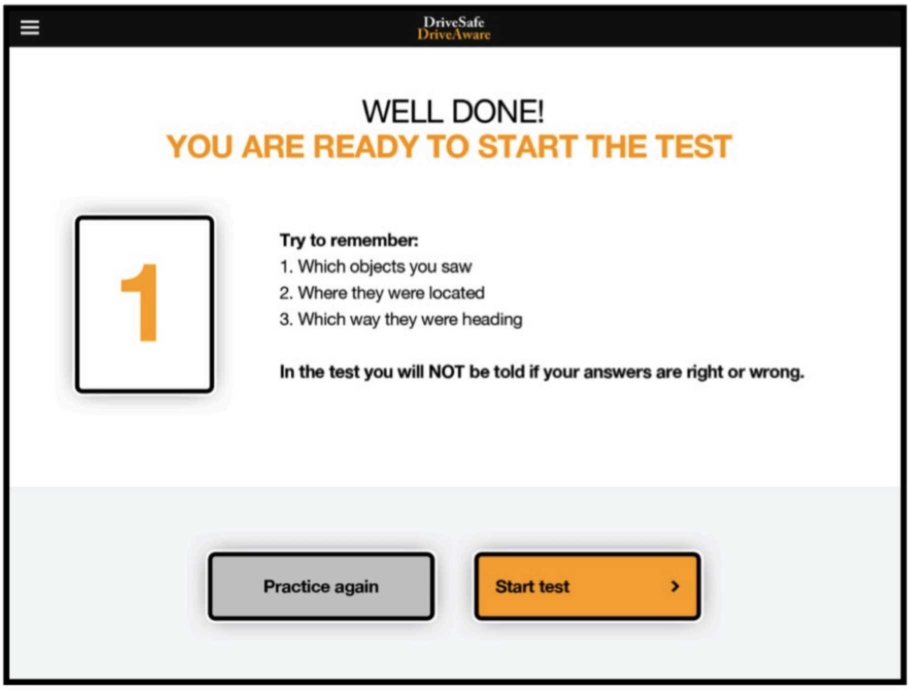


Figure 12. DriveSafe Start Test Screen

No further prompting or assistance may be given. At this point, the administrator may leave the room but should be available if the patient calls for help. If the patient asks for assistance, you can only say:

Sorry, I cannot provide assistance during the test. Just do your best.

You cannot physically assist with test functions, provide verbal prompts, or provide feedback on performance if the patient is self-administering the test.

DriveAware

DriveAware—Self-Administered Questions

The patient completes the two self-administered DriveAware questions according to written and audio instructions. The patient can only choose one response. Do not provide any prompting or assistance during the questions. If the patient asks for help, say:

Sorry, I cannot provide assistance during the test. Just touch the response that seems most correct to you.

DriveAware—Clinician Interview

The remaining five DriveAware questions comprise an interview that you are to conduct the same way for all patients. Conduct the interview in a private area, such as a treatment room or office. You and the patient should be comfortably seated. You must read the questions verbatim and touch the response that is most like the one the patient gives. Following are two examples of how to enter responses for questions in the DriveAware interview.

Example 1 - DriveAware Question 6 (see Figure 13.).

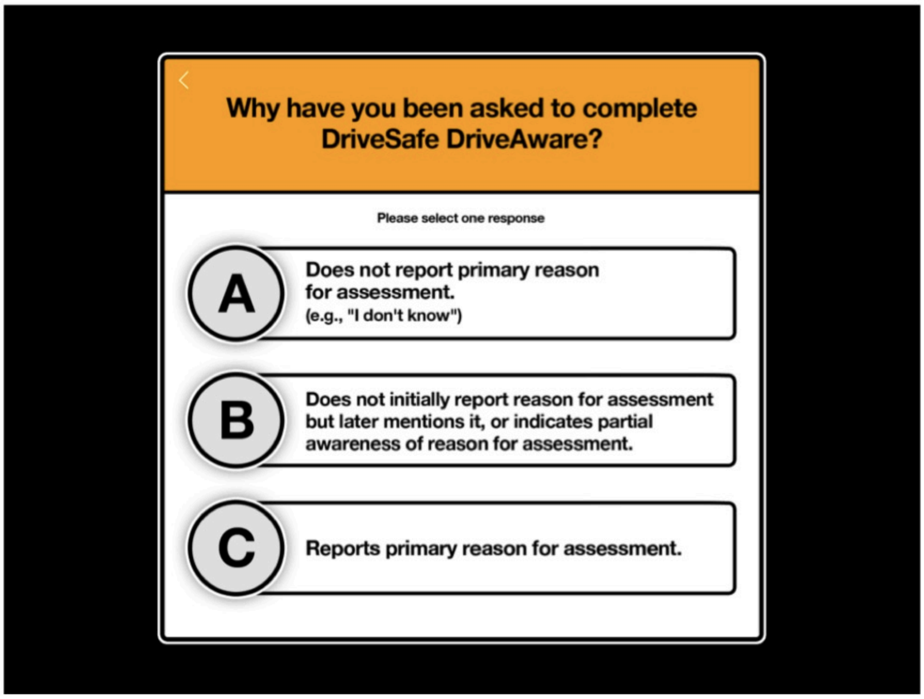


Figure 13. DriveAware Question 6.

Table 1. presents examples of responses to DriveAware Question 6.

Table 1. Examples of responses to DriveAware Question 6.

Example of Patient Responses	Example of Clinical Contexts	Appropriate Responses
"I don't know"	Has been diagnosed with dementia. Diagnosis discussed with patient in the past.	A–Does not report primary reason
"My GP just told me to come."	Referred due to memory loss. Family concerned about forgetfulness.	A–Does no report primary reason
"Doctor can you please fill out my license form to say I am OK to drive"	License authority has posted the patient a medical form to be completed by their doctor due to their age (75).	C–Reports primary reason
"I agreed to be involved in a research project"	Test being administered as part of a research project	C–Reports primary reason
"I'm just getting old"	Had a stroke and recent hospital admission	A–Does no report primary reason
"I don't know. My family are worried about my memory but I'm OK"	Diagnosis of dementia has been discussed with patient in previous appointments	B–Indicates partial awareness
"I have had a brain injury so people are worried this might affect my driving. I think I'm OK."	Has had a brain injury. There is uncertainty regarding whether this will affect driving safety.	C–Reports primary reason
"I've had a stroke"	Had a stroke last month but has largely recovered.	C–Reports primary reason
"I don't know – I have no problems with my driving. I have been driving for 60 years"	Told practice nurse earlier he had hit a pole in a car park so his daughter had suggested seeing the doctor.	B–Indicates partial awareness

Example 2 - DriveAware Question 7 (see Figure 14.).

Figure 14. DriveAware Question 7. (Clinician Rating)

For DriveAware Question 7., select A, B or C depending on your judgment for this patient (see Figure 14.). **Base your decision on the clinical information you have available** (e.g., medical diagnosis, vision report, medical test results, referral information) **and other information reported by family, caregivers or the patient** (e.g., difficulty with daily activities, reduced mobility, recent collisions, report of reduced frustration tolerance).

Intersection Rules (Optional)

The Intersection Rules subtest is an optional subtest. When the administrator first sets up the patient, they will be required to indicate whether this subtest should be included for each individual.

The administrator does not need to be present during the Intersection Rules subtest. If the patient calls for help, assistance can be provided only for the practice item. This may include demonstration of how to number vehicles or select 'undo'. The administrator may repeat instructions but only as written (e.g., **You need to number all of the vehicles**). The patient can repeat the practice item as many times as they like (if at all). If the patient asks for help deciding right of way, say:

Sorry, I cannot provide assistance during the test. Just do your best.

DSDA Administrator-Assisted Method

Sit at the table next to the patient on the patient's non-dominant side. This position enables you to take over test functions without touching the patient. The patient should have an unimpeded view of the tablet.

- The tablet should be muted so that there is no sound.
- The administrator enters or selects test functions (e.g., selection of 'start' and 'next', and indication of DriveSafe responses).
- The patient may continue to enter responses as long as it does not disrupt the flow of the test. Use your clinical judgment to decide.

DriveSafe

Starting the Test

Begin testing with the DriveSafe start test screen (see Figure 15.). Clearly state the following instructions to the patient:

You will see images of an intersection. Try to remember

1. which objects you saw,
2. where they were located, and
3. which way they were going.
Are you ready to start?

Select 'Start test' when the patient is ready to begin. You may repeat the instructions verbatim as many times as required. Do not use any additional words.

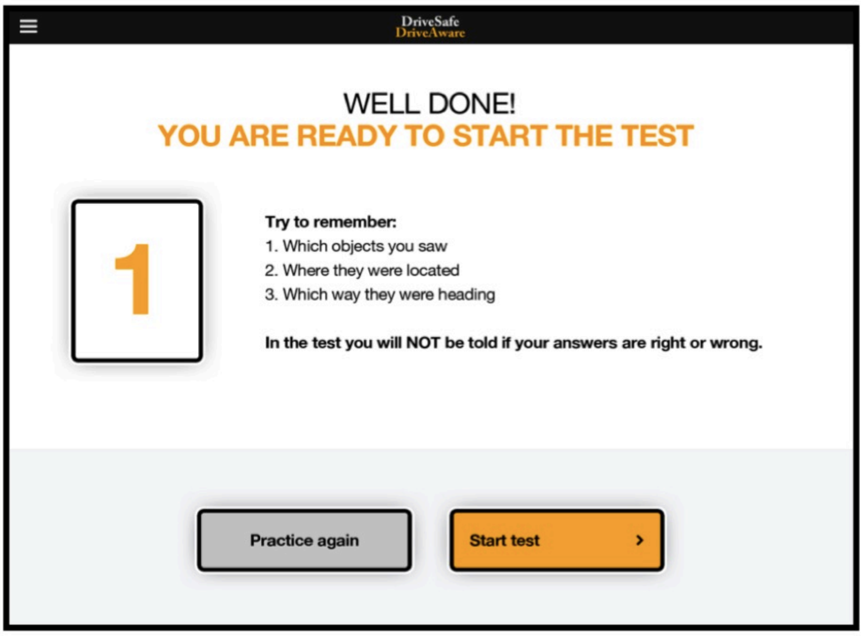


Figure 15. DriveSafe Start Test Screen

During Countdown

During the countdown, say: **Get ready to look.**

After Objects Disappear

When the objects have disappeared, say: **Tell me what you saw**.
If the patient says "There is nothing there" or does not understand the test requirements, say:
Try to remember the objects you saw before they disappeared. Tell me what you saw.
You may repeat these words many times as required, but do not use any additional words.

Entering Responses in DriveSafe

Your role during DriveSafe is simply to provide assistance with response entry. No prompting should be given to elicit information. For example, the administrator is not permitted to say "which way was the car going?" or "where did you see the car?"

Enter responses exactly as the patient has stated them. Pointing and gestures are accepted to indicate responses (e.g., direction of movement).

Keep in mind the three categories of information that need to be recalled when entering patient responses (name of object, object location, and direction of movement). If the patient says "there was a car over there", enter only the object type (car) and object location (as indicated by the patient) and, not the direction of movement. Table 2. provides examples of interpreted patient responses.

You may allow the patient to touch the tablet **if you judge this as effective** (e.g., the patient enters the location and object type, and you enter the direction of movement). Ask the patient to locate the object as specifically as possible by say: **Is this what you mean?** No other prompting or instruction is allowed. If the patient says "no", delete the response and re-enter it as indicated.

If the patient accidentally triggers unwanted or additional responses (e.g., as the result of a tremor), simply delete the responses. When using the Administrator-Assisted Method, the administrator responds to all pop-ups and messages such as, **Are you sure you have completed this screen?** and **Do you want to delete this response?** The goal is to facilitate a smooth testing flow and limit disruptions.

Patients with significant cognitive deficits may attempt to tell the administrator their responses while objects are still displayed on the screen. In this case, say: **Wait until the objects are gone**. When they have disappeared, say: **Now tell me what you saw**.

Table 2. Example responses with corresponding response entry

Patient Response	Response Entry
"I saw something here (pointing but not touching screen) but I can't remember what it was."	**Location:** Administrator touches location point indicated by patient but first asks: **"Is this what you mean?"** **Object:** No entry (leave menu on screen) **Direction of Movement:** No entry
"I saw two people over here (touching screen)."	**Location:** Location point entered by patient **Object:** Administrator selects couple icon **Direction of Movement:** No entry
"I saw a thing here on the road but I can't remember if it was a car or a truck. I'm not sure."	**Location:** Administrator touches location point indicated by patient but first asks: **"Is this what you mean?"** **Object:** No entry **Direction of movement:** No entry
"I saw a person over here" (pointing to pole on left footpath but not touching screen).	**Location:** Administrator touches location point indicated by patient but first asks: **"Is this what you mean?"** **Object:** Administrator selects person icon **Direction of movement:** No entry
"There were two people here" (pointing to right footpath). "They were walking this way" (gesturing left) "and a bike over here" (touching screen) "coming this way" (gesturing toward self).	**Location 1**: Administrator touches location point indicated by patient but first asks: **"Is this what you mean?"** **Object 1**: Administrator selects 2 person icon **Direction of movement 1:** Administrator moves arrow to approximately 270º. **Location 2:** Entered by patient **Object 2:** Administrator selects bike icon **Direction of movement 2:** Administrator moves arrow to approximately 180º.

When the patient is ready to move to the next image, ask: **Are you ready to move on?** If the patient says yes, select Continue.

DriveAware

DriveAware—Two Self-Administered Questions

Read the questions exactly as written and ask the patient to respond. Allow the patient to see the screen and read the text, but do not allow the patient to touch the screen. You must enter the responses for the patient (this is to avoid incorrect responses inadvertently being selected as responses cannot be changed). If the patient asks for help, say: **Sorry, I cannot provide assistance during the test. Just tell me the response that seems most correct to you.**

DriveAware—Clinician Interview

The remaining five DriveAware questions are conducted in the same way for all patients, whether the test has been self-administered or administrator-assisted. This procedure is described on page 19.

Intersection Rules (Optional)

The Intersection Rules subtest runs if you opted to include it in the set up stage.

Practice Item

You must read the instructions exactly as written in (see Figure 16.). Mute the sound for this part of the test. Select 'View practice' when the patient is ready.

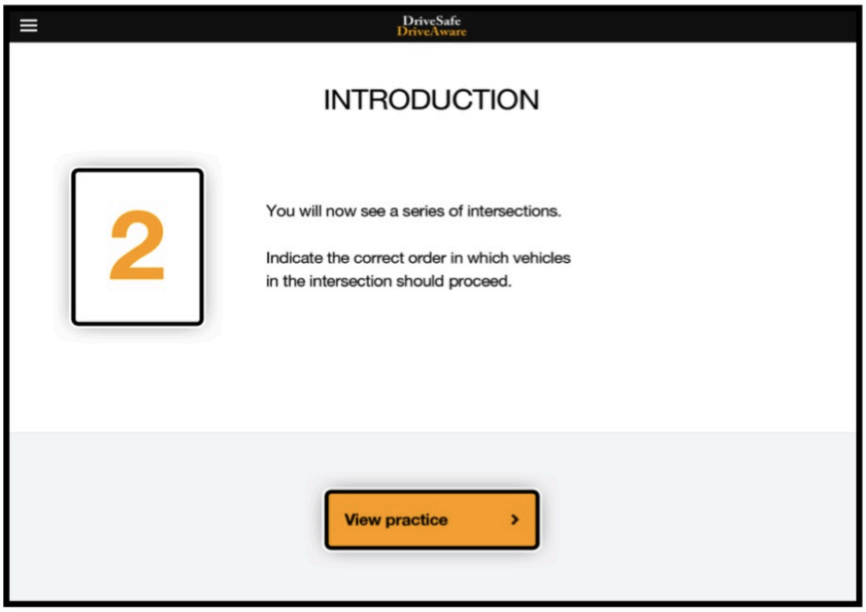


Figure 16. Intersection Rules Instructions

When the practice intersection appears, ask: **Which car goes first?** The patient may touch the desired car him- or herself, or you may touch the car on his or her behalf. Next, ask: **Which car goes second?** Again, the patient may touch the desired car or you may touch it on his or her behalf. Allow the patient to touch cars only if it does not disrupt the flow of testing. If the response is incorrect, read the error messages verbatim and then select Continue or Try again. The practice item can be completed as many times as you think it is necessary.

When the patient is ready to start the test, read the instructions on the Intersection Rules start test screen (see Figure 17.) verbatim and then select 'Start test'.

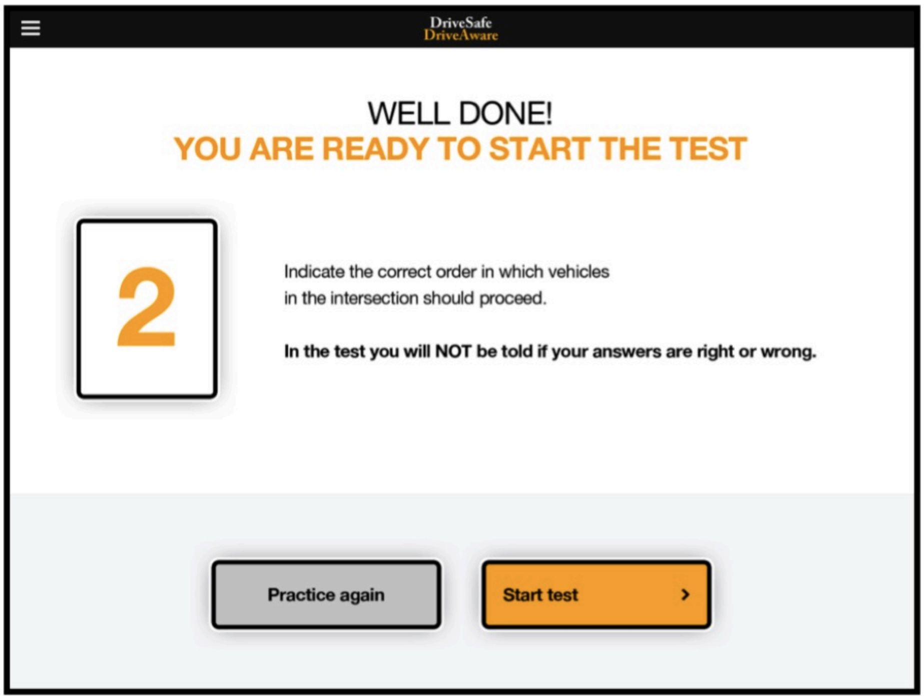


Figure 17. Intersection Rules Start Test Screen

Test Items

Follow the same procedure for selecting vehicles as you did with the practice items.

Ask: **Which car goes first?**

Enter the response (or allow the patient to touch the desired vehicle). Do this for all remaining vehicles (e.g., **Which car goes second? Which car goes third? Which car goes fourth?**).

Chapter 4.
Scoring and Reporting

Scoring

The touch screen DSDA automatically scores all test items and records all subtest timings. These are presented in a clinical report and patient letter at the conclusion of the test (see Reporting).

DriveSafe Scoring

The DriveSafe subtest consists of 10 intersections, with a total of 28 objects presented. All possible independent and combined variable options were examined throughout the development and research phase. Results indicated that the variables most predictive of driving performance were object type, object location, and direction of movement, scored as independent variables. As a result, these variables were selected as the basis for the DriveSafe subtest scoring.

Each correctly recalled piece of information (object type, location and direction of movement) in the DriveSafe subtest receives one point. The maximum score for 28 objects is 84. The total number of objects missed, the total number of details missed (location and directions), and the number of additional objects the patient indicates (that were not actually present) are also recorded. Time taken to complete the test is automatically recorded.

The total DriveSafe subtest score (maximum 84) is used to categorise patients three ways (see Figure 21.). A score of 57 or below indicates the patient is likely to fail an on-road assessment. A score of 72 or above indicates the patient is likely to pass an on-road assessment. A score of 58 to 71 indicates further assessment is required to determine fitness to drive. The patient's DriveAware results are then applied to this scoring categorisation.

DriveAware Scoring

The DriveAware subtest consists of 6 items. A discrepancy score is calculated for each item based on the difference between the patient's self-rating and actual performance on the DriveSafe subtest, along with the clinician's ratings. A final score is generated by converting the discrepancy score (–2 to 2) into a final score via a 5-point ordinal scale. This results in a maximum total score of 17. A score of 10 or below indicates impaired awareness of abilities related to driving and a high score of 13 or above indicates awareness of abilities related to driving. A score of 11 or 12 indicates further assessment is needed.

Final categorisation in DriveAware depends on the DriveSafe subtest outcome. Both DriveSafe and DriveAware scores are combined to further classify patient ability and awareness into the three categories: Pass, Fail, and Further Testing. For example, given a high DriveSafe score (≥ 72) and a low DriveAware score (≤ 10), a patient is classified as Further Testing. A low DriveSafe score (≤ 57) and a high DriveAware score (≥ 13) is also classified as Further Testing. Awareness is necessary for accurately self-monitoring driving performance and implementing strategies to compensate for any reduction (e.g., avoiding complex traffic conditions). In the first scenario, awareness is reduced, so further assessment is advised regarding fitness to drive. In the second scenario, the patient performed poorly in DriveSafe, but DriveAware results indicated a high level of awareness, so this patient has more potential for implementation of compensatory strategies. Further testing is therefore advised. The DSDA scoring system and decision flowchart is illustrated below (see Figure 18.).

Patient Categorisation

DriveSafe DriveAware subtest scores are used to identify patients who are likely to fail an on-road driving assessment, those who are likely to pass an on-road assessment, and those who require further testing to determine fitness to drive. The cut scores and categorisation flow chart for both subtests are indicated in Figure 18.

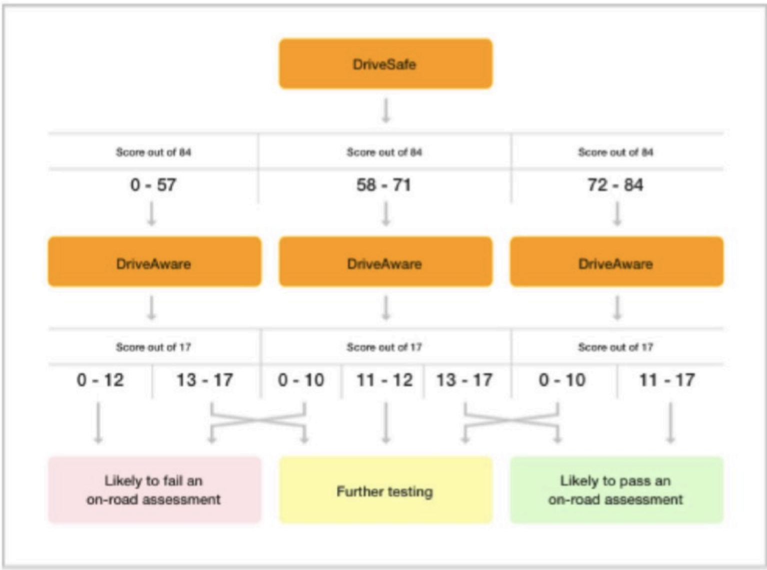


Figure 18. DriveSafe DriveAware classification flow chart

Intersection Rules Scoring (Optional)

The Intersection Rules subtest consists of 8 items. One point is awarded per item only when all vehicles are placed in the correct order of right of way. This results in a maximum score of 8. Time taken to complete the test is recorded.

Intersection Rules was initially investigated in the context of DriveAware in the statistical analysis during the development and research phase. Results indicated the subtest was not useful as a predictor of driving performance and was not correlated with performance on the DriveSafe subtest, the DriveAware subtest, or with the on-road assessment outcome. Intersection Rules was therefore altered to be an optional subtest. Patient performance on this subtest does not contribute to DriveSafe or DriveAware scoring. The subtest is retained primarily for driver trained occupational therapists to use as a brief screen of knowledge regarding right of way at intersections and for the qualitative information provided (e.g., ability to plan responses and follow instructions). This subtest may still be useful for other clinicians including general practitioners and can be selected for inclusion at the patient set-up stage.

Reporting

When the patient has completed DriveSafe DriveAware, you can generate the report. You can have the report emailed to yourself or print it directly from the tablet. The report consists of three sections:

1. Clinical Report: Summary Report
2. Clinical Report: Extended Report
3. Patient Letter

You may generate only the Summary Report or both clinical reports if additional information is required (e.g., the patient wants evidence of why they fell into a particular category). The Patient Letter provides a brief summary of performance and the implications of this performance for the patient. A sample Clinical Report (Summary and Extended) and Patient Letter is provided on page 37. A 9-square, colour-coded graphic provides a quick visual representation of performance in both the Summary Report and the Patient Letter.

Clinical Report: Summary Report

The Summary Report is a brief, one-page summary of performance that includes demographic information and any notes that were entered. The Summary Report includes the following information (see Figure 19.):

- Demographic information
- DSDA diagram illustrating test results
- Total scores for the DriveSafe DriveAware subtests and the Intersection Rules subtest (if completed)
- Outcome category (likely to pass on-road, likely to fail on-road, requires further testing)
- Recommendation
- Notes entered by the administrator into the open notes fields

Clinical Report: Extended Report

The Extended Report provides more detail on performance. The Extended Report includes the following information (see Figures 20. and 21.):

- Summary of the number of objects missed and details missed in the DriveSafe subtest
- Additional objects indicated (that were not present) in the DriveSafe subtest
- DriveAware score in relation to cutoff scores
- Performance on the Intersection Rules subtest (if completed)—simple intersections compared to complex intersections
- Time taken to complete DriveSafe and Intersection Rules (if included)
- Categorisation diagram illustrating cutoff scores and categorisation process
- Caution notices if the test flags that the person speaks English as a second language or has a psychiatric condition.

Patient Letter

The Patient Letter includes the following information (see Figure 22.):

- Purpose of the test
- Total score on the DriveSafe subtest as well as indication of the number of objects and details missed
- Total score on the DriveAware subtest
- Total score on the Intersection Rules subtest (if completed)
- Outcome category (likely to pass on-road, likely to fail on-road, requires further testing)
- DSDA diagram illustrating test results
- Recommendation

DriveSafe DriveAware - Summary Report

Name	MRN
Ruth Allan	**417**
DOB / Age	Assessment Date
03 Aug 1945 - 69 years	**03 Mar 2015**
Gender	Test Administrator
Female	**Beth Cheal**
Diagnosis	
Parkinson's disease	

DriveSafe	**78**/ 84
DriveAware	**11**
Intersection Rules	**7** / 8

Outcome*

Likely to pass an on-road assessment

*Please consult manual for interpretation of category.

Recommendation

Results indicate the patient is likely to pass an occupational therapy on-road assessment and is likely to be able to manage the cognitive aspects of driving. Please consider this recommendation in the context of your clincal judgement.

PEARSON

Figure 19. Example Summary Report

DriveSafe DriveAware - Extended Report

Name
Ruth Allan

MRN
417

1. DriveSafe (Objects & Directions)

DriveSafe determines awareness of the driving environment.

Table 1: Missed information

Information	No. Missed / Incorrect
Objects	1/28
Details (Location / Direction)	5/56

Total score **78** / 84

Additional objects **0**

Time taken to complete test **4** minutes and **6** seconds

Research indicates the median time taken to complete DriveSafe for people who pass an occupational therapy on-road assessment is **4 minutes and 46 seconds**, and for people who fail the assessment, **6 minutes and 49 seconds**.

2. Intersection Rules

Right of way is determined for 8 intersections.

Table 2: Performance based on intersection complexity

Intersection No.	Road Signs	No. of Vehicles in Image	Score
1, 3, 4, 7	Without road signs	2	4/4
2, 5, 6, 8	With road signs	3-4	3/4

Total score **7** / 8

Time taken to complete test **0** minutes and **33** seconds

Research indicates the median time taken to complete Intersection Rules for people who pass an occupational therapy on-road assessment is **2 minutes and 38 seconds**, and for people who fail the assessment, **3 minutes and 49 seconds**.

2 of 4

PEARSON

Figure 20. Example Extended Report—Page 1

DriveSafe DriveAware - Extended Report

Name
Ruth Allan

MRN
417

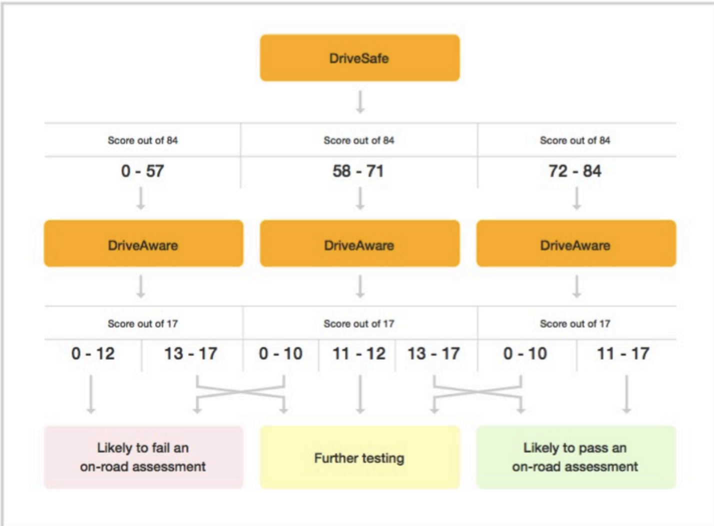
3. DriveAware

DriveAware determines awareness of own abilities related to driving.

Score **11**

A score of 10 or below indicates impaired awareness.
A score of 11 or 12 indicates further assessment is needed.
A score of 13 or above indicates intact awareness.

Figure 1. DriveSafe DriveAware Categorisation Cut-off Scores



PEARSON

Figure 21. Example Extended Report—Page 2

DriveSafe DriveAware - Patient Letter

Name
Ruth Allan

MRN
417

The purpose of this test is to determine a driver's awareness of the driving environment and their own abilities related to driving. Results are used to indicate if further assessment of ability to manage the cognitive aspects of driving (thinking skills) is required.

Your results were:

DriveSafe	**78** / 84	Missed 1 object and 5 details (location/direction).
DriveAware	**11**	A score of 10 and under indicates reduced awareness of abilities related to driving.
Intersection Rules	**7** / 8	

These results place you in the category:

Likely to pass an on-road assessment

Your results indicate you are likely to pass an occupational therapy on-road assessment (provided you do not have any physical impairment that may affect your driving ability).

Note: If you have a medical condition that can affect physical capacity to drive, you may be required to undergo an occupational therapy driving assessment even if you fall into the 'green' section on this graph.

If you would like more information about DriveSafe DriveAware please visit the following website: https://www.pearsonclinical.com.au/.

Disclaimer

The recommendations above were based on information provided by you, clinical information, performance on the day of the assessment, and results of research. The results of this test are for clinical and diagnostic purposes only, and are intended for use with a licenced clinical practitioner. The results should not be relied upon as an indicator of being collision-free (or otherwise).

PEARSON

Figure 22. Example Patient Letter

Chapter 5.
Test Development and Research

Background

Responsibility for determining fitness to drive typically falls to the primary healthcare physician (referred to as general practitioners in Australia and New Zealand). General practitioners report concern about their role in assessing patient fitness to drive, including the impact of withdrawing driving on the patient's quality of life and the patient-doctor relationship. General practitioners frequently report a lack of objective, valid, and reliable tools for predicting driving ability to assist them in this role (Schattner, Jones, Beveridge, Sims, & Rouse-Watson, 2010).

Improved survival rates after stroke or brain injury and an ageing population mean greater numbers of people with cognitive impairment who wish to resume or retain their ability to drive. Appropriately identifying "at risk" drivers is a growing challenge for society, general practitioners, and licensing authorities.

A standardised off- and on-road driving assessment conducted by a driver-trained occupational therapist is considered the gold standard for determining fitness to drive (Kay, Bundy, Clemson, & Jolly, 2008). This method of testing, however, is time consuming and costly, which is usually borne by the driver. Due to a shortage of specialist occupational therapists, access to testing in remote areas can be limited and even urban areas can have long waiting lists (Kay, Bundy, & Clemson, 2009a).

For more than 25 years, researchers have examined a variety of clinical tests to identify an off-road assessment that can accurately predict driving performance without taking drivers on the road. The computer version of DriveSafe and DriveAware (DSDA) is the only test that has shown sufficient sensitivity and specificity to predict on-road performance accurately (Kay, Bundy, Clemson, Cheal, & Glendenning, 2012). This test has been used by occupational therapists as part of a clinical assessment of fitness to drive for more than 20 years. Computer administration of the DSDA is limited to driver-trained occupational therapists because verbal responses need to be interpreted by trained professionals.

The touch-screen version of the DSDA was developed as a portable, user-friendly test that can be administered in a medical or clinical setting by general practitioners or other health professionals without specialised training. Touch screen technology enables patient responses to be precisely captured and interpreted based on results of the development research.

Precursor Tests—VRST and DSDA Computer Version

The precursor test for DriveSafe DriveAware was the Visual Recognition Slide Test (VRST), developed by Becky Zropf, an occupational therapist who established driver assessment and rehabilitation training for occupational therapists at the University of Sydney in 1989. Driver-trained occupational therapists in several Australian jurisdictions used VRST for many years as part of a clinical assessment of fitness to drive (Kay, Bundy, & Clemson, 2008). VRST is conceptually different from other driver screening tests currently being used, in that it assesses global awareness of the driving environment rather than the component visual processing and cognitive skills (Kay, Bundy, & Clemson, 2008).

Dr. Kay examined the psychometric properties of VRST in 2007 via a large retrospective study (N = 838). The VRST was found to have sound psychometric properties and was a promising screening test with a specified cutoff score (Kay, Bundy, & Clemson, 2008). The test yielded sensitivity and specificity of 81% and 89% respectively. VRST was updated, shortened and renamed DriveSafe, in light of these results.

The research also identified the importance of awareness of driving ability for safe driving. In the absence of suitable measures of driving awareness at the time, Dr. Kay developed the Driving Awareness Questionnaire (DriveAware). DriveAware was found to have good construct validity; however, more items were needed to assess the least aware drivers (Kay, Bundy, & Clemson, 2009b). The test was subsequently modified to include more items, improving its psychometric properties.

The construct validity and internal reliability of DriveSafe and DriveAware were re-examined following modification, in a study of 115 drivers with cognitive impairment using Rasch modelling (Kay, Bundy, & Clemson, 2009a). Results indicated that when DriveSafe and DriveAware are used together, it is possible to separate drivers into unsafe, safe, and further testing categories. Fifty percent of drivers were clearly categorised as either unsafe or safe. The optimal lower cut score identified unsafe drivers with a specificity of 97% and the optimal upper cut score identified safe drivers with a sensitivity of 93% (Kay, Bundy, & Clemson, 2009a).

In 2009, Pearson Clinical Assessment published the computer version of DriveSafe DriveAware (DSDA). Pearson converted DSDA to touchscreen in 2015.

DSDA Conversion to Touch Screen

Conversion to touch screen technology provided opportunity for a number of key simplifications to be made to the test to suit the new delivery mode (tablet) and testing context.

Market research was conducted by Pearson prior to development of the touch screen version; a representative sample of 250 general practitioners across Australia was surveyed to determine how they would use a touch screen version of DSDA in medical practice. Most general practitioners reported they would prefer the practice nurse to set up and monitor the self-administered sections of the test; then the general practitioner would conduct the brief interview and provide test results to the patient. Further details on setup and administration can be found in Chapter 3.

DriveSafe

The computer version of the DriveSafe subtest relies on the clinician's interpretation of the patient's verbal responses. The touch screen test is able to capture more information in the one response. For instance, the required response inputs of 'object location' and 'side of screen' (required in the computer version of the test), could be collapsed in the touch screen version, as touching the tablet screen provided both types of information within the one response. Furthermore, images from the previous version of the DriveSafe subtest were updated to facilitate international adaptation of the test. A four-way intersection replaced the roundabout image. With the exception of this change, new images replicated the computer version as closely as possible. Suitable boundaries for the 28 location zones and 28 direction ranges were set by comparing responses from 30 study participants with a sample of participants with no impairments (N=30).

DriveAware

Five out of seven DriveAware subtest questions were retained from the computer version of the test, some with minor word changes where words were no longer applicable (e.g., 'slide test' was adjusted to 'test'). Two questions were dropped from the subtest because they related to items that could not be measured in the new testing context (i.e., anxiety about the assessment process and difficulty with memory). One recommendation from Dr. Kay's research was that additional items be included in DriveAware. The ability of the touch screen version to capture precise information regarding performance, allowed the addition of 3 new DriveAware questions relating to patient performance on the DriveSafe and Intersection Rules subtests. As discussed statistical analysis indicated the two items related to performance in Intersection Rules were not successful and therefore these were dropped. The version of the DriveAware subtest used in the research included 8 items and the final published version includes 6 items (see Table 3.).

Table 3. Final DriveAware Questions

Item Order	DriveAware Question
1	**How well did you remember the location of people and vehicles? (discrepancy score calculated according to actual performance in DriveSafe subtest)**
2	**How well do you remember the direction of travel? (discrepancy score calculated according to actual performance in DriveSafe subtest)**
3	**Do you have any concerns about your driving ability? (discrepancy score calculated according to clinician rating)**
4	**How often do you get surprised by vehicles / pedestrians appearing out of nowhere? (discrepancy score calculated according to actual performance in DriveSafe)**
5	**Why have you been asked to complete DriveSafe DriveAware? (discrepancy score calculated according to actual reason for completing DSDA)**
6	**How well do you think you did in the test? (discrepancy score calculated according to actual performance in DriveSafe)**

Intersection Rules

The computer version of the Intersection Rules subtest consisted of 8 items. Two items were dropped due to differences in road rules between Australia and New Zealand. These items were replaced with 3 new intersections, with the same number of vehicles and level of complexity as the removed items. Following statistical analysis, one of the new intersections was dropped to shorten the test and return it to the original 8 items. An additional item was included in the research in case there was an unforeseen problem with one of the items.

Research Aim

A prospective study was conducted to determine if the DSDA touch screen version was a valid, user friendly tool that general practitioners and other health professionals could administer to determine if older and cognitively impaired patients were able to manage the cognitive aspects of driving or if they required referral to a specialist driving service for further assessment.

The test was administered to a convenience sample of 134 older (60 years +) and cognitively impaired (18 years +) drivers referred over a 7-month period to ten driver assessment and rehabilitation clinics across Australia (Sydney, Melbourne, Perth and Brisbane) and New Zealand (Wellington, Auckland and Hamilton). The purpose of the study was to examine the psychometric properties of the test and its predictive validity as compared to the criterion measure of a standardised occupational therapy on-road assessment. Sixteen driver trained occupational therapists were involved in administering the test and conducting the on-road driving assessments.

Quality Assurance Procedures

Ethics Approval

The University of Sydney Human Research Ethics Committee provided approval for this study. Two of the research sites were public hospitals and therefore required further ethical approval via the National Ethics Application Form (NEAF). St Vincent's Hospital Sydney Human Research Ethics Committee provided ethical approval for the two hospital sites.

Assessor Qualifications and Training

An occupational therapy driving assessment is considered the gold standard for determining fitness to drive as actual driving performance is observed (Kay, Bundy, Clemson, & Jolly, 2008). Therefore, an on-road driving assessment was used as the criterion measure for this study. In Australia, an occupational therapy driving assessment must be conducted by a registered driver-trained occupational therapist. A qualified driving instructor must be present in the vehicle and it is considered best practice for the assessment to be conducted in a dual controlled vehicle. Accordingly, the research sites used in this research were all existing driving clinics where assessments were conducted by appropriately qualified occupational therapists and driving instructors in the course of their usual work.

Training was provided to all research sites via two, one-hour interactive webinars conducted by Dr. Ann-Helen Patomella, Driver Trained Occupational Therapist, Lecturer at the Karolinska Institute in Sweden and author of PDrive, and Beth Cheal, DSDA Project Manager, Driver Trained Occupational Therapist and Driving Instructor. Content was summarised in a manual provided to each site. Training included the standardised administration of DSDA, criteria for the standardised on-road assessment route, the standardised administration of PDrive (the on-road assessment tool used), and how to implement study protocols. After the first webinar, individuals at each site were required to undertake an evaluation of driving performance of an impaired driver using PDrive (presented in a DVD recording). Results were compared in the second webinar to ensure consistent marking.

After the training was completed, all research sites began data collection. In addition to the DSDA assessment and PDrive on-road assessment, occupational therapists were asked to complete a Mini Mental State Examination Version 2 (MMSE-2) (Folstein, Folstein, McHugh & Fanjiang, 2010), for each participant to determine cognitive status. Driving Instructors were asked to complete a route evaluation checklist and driving clinics were asked to submit a map of their standardised on-road assessment route as part of the standardisation process. Occupational therapists were asked to rate on-road awareness as 'intact', 'absent' or 'partial' using standard definitions.

Quality Checks Performed

After assessments were completed, Beth Cheal, the DSDA Project Manager, reviewed all data and participant driving reports to ensure participants met the study criteria and that the assessment had been conducted according to the standardised study protocols. This ensured consistent classification of participants based on the definitions of on-road assessment outcomes used in standardisation of the computer version of the test (Kay, Bundy, & Clemson, 2009a):

- **Pass**—Safe and legal driving with no further intervention required
- **Conditional Pass**—Safe and legal driving with restrictions on the license (e.g., kilometer or time of day restrictions)
- **Intervention**—A series of lessons required to improve driving techniques or to learn to use vehicle modifications
- **Fail**—Failure to meet criteria for safe and legal driving or substantial errors and/or driving instructor intervention required for safety

It is critical that driver-trained occupational therapists who use the DSDA as a part of their off-road screen adhere to these definitions of pass, fail, and intervention when taking the patient on-road, to ensure patients are catagorised correctly. In particular, if driving-instructor intervention was required for safety (including both verbal and physical intervention), this was not considered a pass for this study regardless of the circumstances.

Statistical Analysis Methods

Construct validity and internal reliability of the DriveSafe DriveAware subtests were examined using a Rasch modelling technique (Bond & Fox, 2007) via Winsteps Version 3.72.2 (Linacre, 2014a). Rasch modelling constructs a linear measure from ordinal scores by converting raw scores into logit scale scores and assessing goodness-of-fit for both items and participants along the same measure continuum. An item and participant map is generated in which items are arranged in order of difficulty and participants are arranged in order of competence. The analysis generates two pairs of goodness-of-fit statistics; infit and outfit, expressed in two forms as mean square fit statistics (MnSq) and standardised fit statistics (ZStd). These statistics indicate how well data from each item and participant fit to the Rasch model with the assumption that easy items are easy for all participants and more competent participants perform better on all items.

Items or participants with fit statistics outside the acceptable range should be considered for removal from the test. Point measure correlation coefficients should be positive and large enough to show a strong relationship between the item and the construct. Both item fit statistics and point measure correlations are used to verify the unidimensionality of the test (i.e., to ensure items are only included for measuring a single underlying concept.) The unidimensionality of the test was also examined via a principal component analysis with Winsteps Rasch software (Linacre, 2014a). When the empirical variance closely matches the modelled variance and when the percentage of unexplained variance from the first factor is much less than the percentage of explained variance by the Rasch model, the test fits the expectations of the model as the evidence of unidimensionality and the construct validity of the test (Linacre, 2014).

Rasch modelling produces reliability estimates for both items and participants. A separation statistic provides evidence of internal reliability (ability of the test to separate groups of participants into levels of ability). In order to conclude that differences in the measure are due to real differences in the extent to which participants possess the trait and not due to error of measurement, the separation statistic should be 2.00 or greater. The participant reliability index (Cronbach's alpha equivalent) and the item reliability index should be 0.80 or higher (Linacre, 2014).

To ensure that the test items function similarly for male and female participants, a differential item function analysis (DIF) is completed. Items that function significantly different ($p \leq .5$) and with a DIF larger than 1.5 logit could be considered for removal from the test (Linacre, 2014).

Finally, the predictive validity of the DriveSafe DriveAware subtests was examined. Optimal lower and upper cut scores were determined using descriptive statistics. The lower cut score was set to identify those who were unsafe (i.e., Sensitivity and Positive Predictive Value) and minimising the proportion of drivers falsely categorised as unsafe (i.e., False Positive). The upper cut was set to identify those who were safe (i.e., Specificity and Negative Predictive Value) and minimising the proportion of drivers falsely categorised as safe (i.e., False Negative). Sensitivity and Specificity were calculated with a confidence interval (CI) of 95%.

Results

The touch-screen DSDA was administered to a convenience sample of 134 older (60+ years) and cognitively impaired (18 years +) drivers referred to ten driving clinics. For the Rasch analysis, two cases were excluded due to missing DriveAware data (N=132). For the sensitivity and specificity analysis, four cases were excluded because intervention had been recommended only for training in the use of vehicle modifications and another sixteen cases were excluded for not having the classifications Pass or Fail (N=112).

DriveSafe

The Rasch analysis indicates that all DriveSafe subtest items have infit and outfit statistics within the acceptable range. The map of items and participants (see Figure 23.) illustrates the spread of participant ability from more competent to less competent in standard deviation units. The range of item difficulty was comparable to the range of participant ability, except for the most competent drivers. This is acceptable as the point of the test is to identify participants as "pass" and not the level of competency within that category. The item–person map demonstrates the test mainly assesses the least competent drivers, which is the group of most concern. The purpose of the test is to identify and classify the "at risk" group of drivers.

Item fit statistics range from .63 to 1.38 (Median = .97), which are in the acceptable range for fit to the model. Item-to-total correlation coefficients are all positive ranging from .44 to .75 (Median = .63), which supports the validity of the DriveSafe subtest. The principal component analysis yielded a high modelled variance (56.4%) and closely matches the empirical variance (55.1%). The percentage of unexplained variance by the first contrast is 4.6%, which is much less than the variance explained by item (14.8%) or by person (40.4%). All the statistics provide strong evidence for the unidimensionality and the construct validity of the DriveSafe subtest.

The DriveSafe subtest initially consisted of 11 items, however, one item was dropped because the analysis suggested this item did not add value to the test. This resulted in a slightly shorter test, which feedback from research sites indicated was favourable. The final version of the DriveSafe subtest therefore consists of 10 items with strong internal consistency of Cronbach Coefficient Alpha (.94).

The analysis indicated the test person separation is high (a model separation of 3.76), with a participant reliability index (Cronbach's alpha equivalent) of .93. The item separation is also high (3.30) with an item reliability index of .92. These results indicate the test is sensitive enough to distinguish high and low performers and verifies the item difficulty hierarchy—all provide evidence for the internal reliability of the DriveSafe subtest.

The DIF analysis on gender revealed no significant differences in item performance between male and female participants.

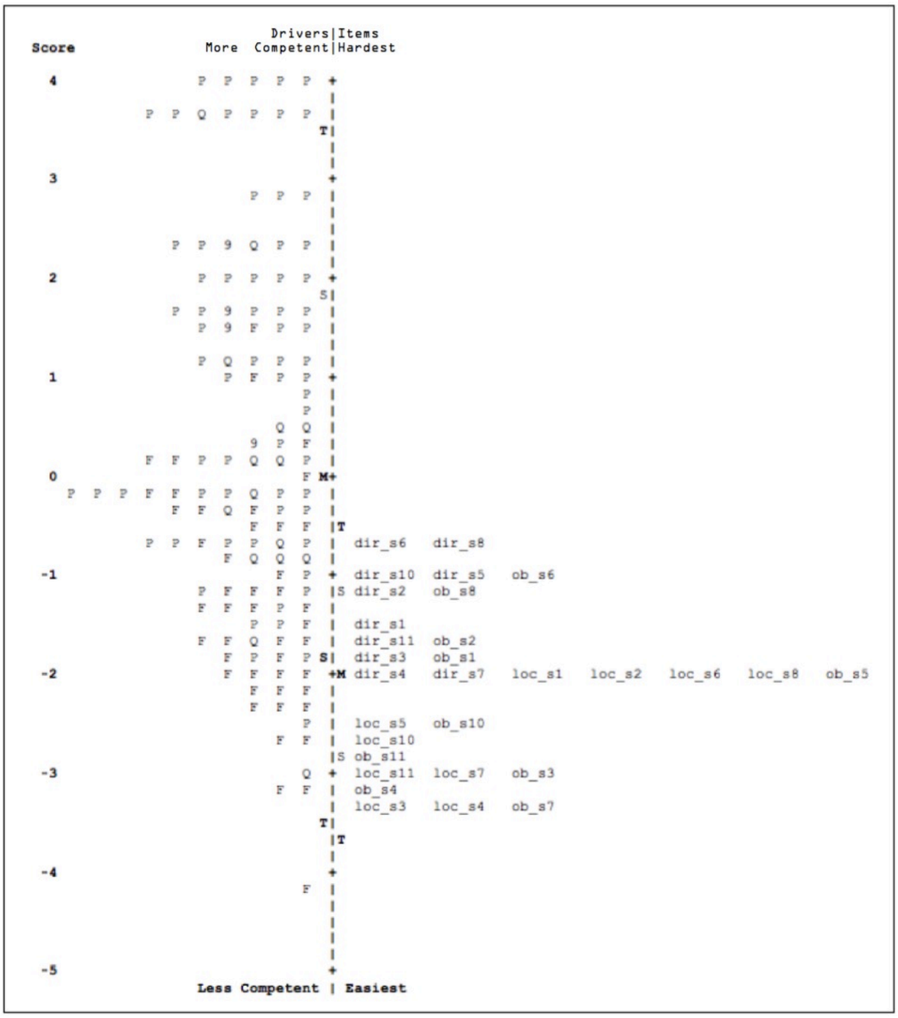


Table 23. Map of drivers and items for the DriveSafe subtest (M=mean, S=one standard deviation, T=two standard deviations, F=failed on-road, P=passed on-road, dir=direction, loc=location, ob=object, s=item number)

DriveAware

The initial research version of the DriveAware subtest consisted of 8 items. A discrepancy score was generated between the patient's performance in the DriveSafe subtest, their self-rated performance and the clinician's rating (-2 to +2). A 5-point ordinal scale was used to generate a final score for each item. To improve the score interpretability, the scale was further adjusted to a 4-point scale of 0-3: where initial scores -2, -1, 0, 1, 2, were recoded to the respective 3, 3, 2, 1, 0 new scores. The final scale of 0–3 removes the negative values and results in a lower score that represents lower awareness and a higher score that represents greater awareness. (In the computer version of the test a higher score represents lower awareness, which may be counter intuitive; thus the scale was reversed).

Two of the initial DriveAware items related to awareness of performance on the Intersection Rules subtest. Item analysis revealed that these items did not perform the same way as others. This, coupled with the optional nature of the Intersection Rules subtest, resulted in the decision to remove the two items from the DriveAware subtest. The total maximum score of the final DriveAware subtest is 17, as five of the six items have a maximum score of 3 and one item has a maximum score of 2.

Among the six items, five have acceptable infit statistics (Infit Mnsq) from .67–1.10 and moderate-to-high item–measure correlations from .64 to .82 (Median = .70). One item had the smaller item-measure correlation ($r = .54$) and the larger Infit statistics (Infit MnSq=1.67) but within the acceptable range (Linacre, 2014). The item behaviour may reflect the fact that responses to this question: "Why have you been asked to complete DriveSafe DriveAware?" were less varied than the responses to other questions. Item analysis results support the construct validity of the DriveAware subtest. The principal component analysis yielded a modelled variance (59.3%) that closely matched the empirical variance (59.4%). The percentage of unexplained variance by the first contrast was 11.4% less than the percentage of variance explained by item (16.1%) or by person (46.3%). All statistics show that there is no clear secondary dimension and no excessive amount of misfitting items, indicating the DriveAware subtest measures a unidimensional construct.

Figure 24. is a comparison of the spread of participant awareness (from intact to absent) with subtest item difficulty. The map reveals gaps in that there were no items to measure participant awareness. As expected, this distribution of items confirms that the DriveAware subtest is not a stand-alone test and should be used in conjunction with another test, such as the DriveSafe subtest. The hierarchy of items is conceptually logical and reflects a progression of awareness.

The item separation of 6.87 and score reliability of .98 imply that the sample was able to confirm the item difficulty hierarchy of the test. However, the person separation was lower (1.98) and the reliability was .80, which further confirms that the DriveAware subtest by itself may not be sensitive enough to distinguish between high and low performers.

Among most of the DriveAware subtest items, DIF analysis revealed no statistically significant difference between male and female participants except for Item 1, "How well did you remember the location of people and vehicles?" For this item, female participants had better awareness than males. There might be various causes for this finding, such as gender differences in self-rating of performance. Because the sample was less controlled than in educational and other experimental testing settings, it is not known whether the DIF analysis results are caused by the sample or by the nature of the test. Because the question is important in tapping the patient's awareness of driving, it was determined that the item should remain in the test.

```
                                      Drivers | Questions
Score                        Intact Awareness | Least Discrepant
  5                                       F F +
                                              |
                                              |
                                              |
                                              |
                                              |
  4                                           +
                                    P P P P F |
                                              |
                                              |
                                             T|
                                              |
  3                                           +
                                              |
                                              |
                                              |
                                F F P P P P P |
                                              |
  2                                           +
                                              |
                                             S|
                P P P P F P F P P P P P P P P |
                                              |
                                              |T
  1       F P P P P P P P P P F P F F P P P P +  Concern re: driving
                                              |
                                              |
      F P P F P P P F P F P P F P P P P P P P |
                                              |
                                              |S
  0                     P P P P P P P P F P P M+
                                              |
                          F F F F F F F F P F |  Surprised by hazards
                                              |  Why asked to do DSDA
                                              |
                            F F F F F F P P F |M
 -1                                           +  Performance on DSDA
                                        F F F |
                                              |
                          F P P F F F F F F F |  DS direction recall
                                             S|
                                  P F F F P P |S
 -2                                           +
                                              |  DS location recall
                                  F F F F F F |
                                              |
                                              |
                                  F F F F F P |T
 -3                                           +
                                              |
                                          P F T|
                                              |
                                              |
                                              |
 -4                                       F F +
                             Absent Awareness | Most Discrepant
```

Table 24. Map of drivers and items for the DriveAware subtest (M=mean, S=one standard deviation, T=two standard deviations, F=failed on-road, P=passed on-road)

Predictive Validity of DriveSafe DriveAware

Among the sample of 112 participants for the predictive validity study, a total of 53 participants (47.3%) passed the driving assessment, 47 (42.0%) failed, and 12 (10.7%) participants required driving lessons to learn to use modifications or to improve driving performance.

To calculate predictive validity, the DriveSafe DriveAware subtests were used together to categorise participants who were predicted to fail an on-road assessment, those who were predicted to pass and those who required further testing to determine fitness to drive (such as referral to a driving clinic).

The optimal cutoff scores on the DriveSafe subtest were 57 and 72. The optimal cut-off scores on the DriveAware subtest were 10 and 13. The test identified unsafe drivers at the low cutoff score with Sensitivity of 91% (95% CI: 84 to 96) and 89% (95% CI: 82 to 97) for DriveSafe and DriveAware, respectively. The test identified safe drivers at the upper cutoff score with Specificity of 94% (95% CI: 87 to 99) and 91% (95% CI: 84 to 99) for DriveSafe and DriveAware, respectively.

The positive predictive value for the lower cutoff was 79% and 83% indicating that participants predicted to be unsafe had a high probability of being unsafe. The negative predictive values generated by the upper cutoff (65% and 57% respectively) indicated that participants predicted to be safe had a slightly high probability of being safe. However, it is worth noting that positive and negative predictive values are influenced by the prevalence of safe/unsafe drivers in the population that is being tested. If we test in a high prevalence setting, it is more likely that persons who test unsafe truly have unsafe outcome than if the test is performed in a population with low prevalence.

The predictive validity results provide the empirical evidence for DriveSafe DriveAware test with Specificity of 86% and Sensitivity of 91%, Positive Predictive Value of 83%, Negative Predictive Value of 92% and the overall Accuracy of Classification of 88% as combining both DriveSafe and DriveAware subtest cut scores in the 3-by-3 categorisation figure.

Conclusion

The results of this study present evidence that supports the clinical utility of the DSDA touch screen version in predicting with substantial accuracy, which patients with a cognitive impairment require an on-road assessment. People who are not a good candidate for an on-road assessment (i.e., those who will likely "fail") can be advised not to drive and can be redirected to use their time and monetary resources in other ways. The research evidence supports the conclusion that the test has retained the strong psychometric qualities of the computer version of the test, including internal consistency, predictive validity and ability to classify drivers into 'pass', 'fail' and 'further testing' categories.

Appendix A

DSDA Demo and Practice Administration

Procedure

Test Set-Up

- Place tablet on a stand angled to 20 degrees, on a table
- Ensure brightness and volume are set to full (no glare / adequate image contrast)
- Seat patient in an upright chair at the table
- Place the tablet directly in front of the patient (midline and within comfortable reach)
- Offer opportunity to use headphones and/or stylus if difficulties noticed in hearing or touching the screen during the familiarisation process (decision to use is the patient's).

Demonstration Section

- If the patient struggles to move the arrow say; **Touch the arrowhead and move your finger to show which way the car was driving.** You may also demonstrate this, but allow the patient to try first.
- **'Repeat demo' may be selected only once.** Repeat the demonstration only if the patient is having significant difficulty and you think it may be beneficial to repeat the demonstration.

Practice Section

- **'Practice again' can be selected only once.** Repeat the practice only if you think it may be beneficial in the patient achieving independence.
- **If the patient is not able to complete at least 1 of the 3 practice intersections independently** (i.e., enter the three pieces of information without assistance), use the **Administrator-Assisted Method'** (Appendix B).

Appendix B

DSDA Self-Administered Method

DriveSafe

If the patient asks for assistance only say:

Sorry, I cannot provide assistance during the test. Just do your best.

Provide no assistance with test functions, verbal prompting, or feedback on performance.

DriveAware—Self-Administered Questions

Provide no prompting or assistance during the two self-administered questions. If the patient asks for help, say:

Sorry, I cannot provide assistance during the test. Just touch the response that seems most correct to you.

DriveAware—Clinician Interview

Read the first 4 questions verbatim to the patient. Tap the response that is closest to the one provided. Enter your rating for the last question. (This question is not read aloud).

Intersection Rules (Optional Subtest)

If the patient asks for help, you may provide assistance on the **practice item**, such as demonstrating how to number vehicles or select 'undo'. You may repeat instructions but only as written (e.g., **you need to number all of the vehicles**). The patient can repeat the practice item as many times as they like. If the patient asks for help only say:

Sorry, I cannot provide assistance during the test. Just do your best.

Appendix C

DSDA Administrator-Assisted Method

Procedure

- Sit on the patient's non-dominant side for easy access to the tablet
- Mute tablet sound
- Take over data entry
- Do not prompt for responses and only use words provided in red.
- The patient may continue to enter responses if not counterproductive (administrator to use clinical judgment to decide)
- If unwanted / additional responses are triggered, delete them. Take over responding to all pop-up messages. Aim: Allow a smooth flow and limit disruption.

Verbatim Instructions

DriveSafe

Start-Up

You will see images of an intersection
Try to remember
1. Which objects you saw
2. Where they were located and
3. Which way they were going
Are you ready to start?

During Countdown: **Get ready to look**
When the objects have disappeared: **Tell me what you saw**
If additional help needed: **Try to remember the objects you saw before they disappeared. Tell me what you saw.**
If tries to respond too soon: **Wait until the objects are gone.**
When gone: **Now tell me what you saw.**
When finished: **Are you ready to move on?**

In-Test

Entering Responses

- Enter patient responses exactly as stated (based on the 3 categories – location, object and direction. If a category is excluded by the patient do not enter)
- Pointing / gestures are accepted to indicate location and direction.
- If clarification is needed, ask: **Is this what you mean?** (adjusting response as required)

Appendix C (cont.)

DriveAware

Self-Administered Questions

Administrator to read DriveAware questions exactly as written and ask the patient to respond. Allow the patient to see the screen and read the text for the self-administered questions. The patient or the Administrator may tap the selected response. If the patient asks for help, say: **Sorry, I cannot provide assistance during the test. Just tell me the response that seems most correct to you.**

Clinician Interview

Read the first 4 questions verbatim and tap the response that is closest to the one given. Enter your clinician rating for the last question. This question is not read aloud.

Intersection Rules

Read any instructions or error messages verbatim as written on the screen, with no additional words.

In Intersection Demonstration / Test Items

Ask: **Which car goes first?** Either the patient or you may tap the selected car (as works best). Continue for remaining vehicles: **Which car goes second? Which car goes third? Which car goes fourth?**

References

Austroads Ltd. (2012). *Assessing fitness to drive for commercial and private vehicle drivers* (4th ed.). Sydney, Australia: Author.

Bond, T., & Fox, M. (2007). *Applying the Rasch Model* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Kay, L., Bundy, A., & Clemson, L. (2009a). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine & Rehabilitation, 90*, 1514–1522.

Kay, L., Bundy, A., & Clemson, L. (2009b). Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation*, 31, 1074–1082.

Kay, L., Bundy, A., & Clemson, L. (2008). Predicting fitness to drive using the Visual Recognition Slide Test (USyd). *American Journal of Occupational Therapy*, 62, 187–197.

Kay, L., Bundy, A., Clemson, L., Cheal, B., & Glendenning, T. (2012). Contributions of off-road tests to predicting on-road performance: A critical review of tests. *Australian Occupational Therapy Journal*, 59, 89–97.

Kay, L., Bundy, A., Clemson, L., & Jolly, N. (2008). Validity and reliability of the on-road driving assessment with senior drivers. *Accident Analysis & Prevention*, 40, 751–759.

Linacre, J. M. (2014). *Winsteps® Rasch measurement computer program user's guide*. Beaverton, Oregon: Winsteps.com.

Folstein, M., Folstein, S., McHugh, P., & Fanjiang, G. (2010). *Mini-Mental State Examination, 2nd Edition™ (MMSE®-2™)*. Lutz, Florida: Psychological Assessment Resources, Inc.

O'Donnell, J., Morgan, M., Savage, G., and Bundy, A. (2013). DriveSafe/DriveAware: a normal population test retest reliability study. *Australian Occupational Therapy Journal, 25th National Conference, 60* (1).

Schattner, P., Jones, K., Beveridge, A., Sims, J., & Rouse-Watson. (2010). *To Drive or not to Drive: Assessment Dilemmas for General Practitioners* (Final Report). Victoria, Australia: Monash University, Department of General Practice.

Appendix B

Are Age Effects, Experience, and Health Status Measured by the Change Blindness Model?

The change blindness effect is stable and sustained across a wide variety of environments. However, the nature of the task and the stimuli presented (task effects) and individual factors (observer effects) can impact change blindness detection rates (Jensen et al., 2011). A model that is capable of measuring the impact of 'task effects' (e.g., object location, relevance, and brightness) and 'observer effects' (e.g., age, experience, medication effects, and personal goals) on performance would have significant advantages for predicting driving performance because it is widely recognised that these factors impact safe driving. Therefore, I reviewed the literature to determine if there was sufficient evidence that the change blindness model was capable of measuring these impacts to enable discrimination among individuals. The online databases searched and the search terms used are described in Chapter 2.

Task Effects

The impact of object features (e.g., location and relevance) on change detection has been widely studied. For example, research results indicate object deletion is more likely to be detected than object addition (Agostinelli, Sherman, Fazio, & Hearst, 1986; Beck, Levin, & Angelone, 2007a; Pearson & Schaefer, 2005) and probable change is more likely to be detected that improbable change (Beck, Angelone, & Levin, 2004). The variables measured in touchscreen DSDA are accuracy of ability to recall object type, location, and direction of movement. Therefore these task effects will be further examined.

Impact of object type / relevance. Viewers seem only sensitive to changes in the structures that are relevant to the task they are currently performing (Rensink, 2002). Within the context of driving, viewers are much more likely to notice changes to the scene that are driving related (e.g., traffic light changes rather than a building change) (Galpin et al., 2009; Richard et al., 2002). Looked-but-failed-to-see events may occur because objects with atypical properties, such as a bicycle entering from an unexpected direction, may be perceived as marginal for the task at hand and therefore may go unnoticed (Koustanaï et al., 2012). There is significant experimental evidence that objects of central interest are detected significantly faster than objects of peripheral interest (Galpin et al., 2009; Hollingworth & Henderson, 2000; Pearson & Schaefer, 2005; Pringle et al., 2001; Rensink et al., 1997; Velichkovsky et al. 2002; Wallis & Bulthoff, 2000)

This included a series of mudsplashes experiments which provided important insights regarding the nature of attention (O'Regan et al., 1999). Observers identified central interest changes almost immediately but took significantly longer to identify peripheral interest changes, missing them completely in 13-30% of cases (O'Regan et al., 1999). Results indicated the observer was unable to make a comparison of the current view with the previous view once the transients had subsided, even though the change was not masked and the disturbance was minor. The authors proposed that these results demonstrated how attention-grabbing transients in the overall scene prevent attention from being focused on the change location, even when in full view (O'Regan et al., 1999). In a subsequent experiment, a small textured rectangle briefly covered the change location to cue attention to the change without disclosing the nature of the change (O'Regan et al., 1999). Viewers could immediately identify the change if made to central-interest objects, demonstrating the mask had not “wiped” representations from this part of the scene and that the central interest elements had been encoded (O'Regan et al., 1999). Observers were often unable to identify marginal interest object changes, indicating the marginal-interest elements had not been encoded. Results of both experiments confirmed that internal representations of the visual world are actually quite sparse and largely only contain information on things of central interest (O'Regan et al., 1999). Only the aspects of the scene that the observer attends to and encodes as interesting, are available for comparison (O'Regan et al., 1999).

Impact of object location. Research results consistently show that the further the changed object is from the eye fixation point, the less people are likely to notice it, regardless of meaningfulness (O'Regan, Deubel, Clark, & Rensink, 2000; Pearson & Schaefer, 2005; Wallis & Bulthoff, 2000). Only one driving-related study was found that contradicted these finding: indicating detection of far-relevant targets was faster than near-relevant targets in driving scenes (Galpin et al., 2009). The authors suggested this might be due to the different scanning processes required for driving compared to other tasks, including a horizontal scanning bias due to the reasonable expectation that important changes will occur along the horizontal (Galpin et al., 2009). This bias is evident in real-world driving, in driving simulators, and even when people are watching videos of driving (Crundall & Underwood, 1998; Crundall, Van Loon, & Underwood, 2006). Drivers have a wider attentional search, which may explain why the periphery was well attended to by the participants (Galpin et al., 2009). Galpin et al.'s (2009) findings are consistent with research that shows novice drivers rely on foveal or central vision in driving, with long fixations centrally (Mourant & Rockwell, 1972; Summala, Nieminen, & Punto, 1996). Experienced drivers perform better than novices with peripheral vision, showing longer fixations to the periphery and other specific areas of the scene (Mourant & Rockwell, 1972; Summala et al., 1996). These scanning patterns are not the same for non-driving tasks (Summala et al., 1996). Finding from this research supports the application of the change blindness protocol to assessment of fitness to drive.

Impact of object orientation. Few change blindness studies have examined the impact of object orientation on change detection in isolation. Object orientation is important in driving so that the driver can determine direction of movement in order to judge risk. Only one driving-related study was found that examined the impact of object orientation from the perspective of a driver. Kostanaï et al. (2012) presented a group (n = 60) of novice drivers (i.e., less than 2 years' experience) and experienced drivers (n = 60) with 45 naturalistic driving scenes on a data projector screen, via a one-shot task. The changed object was a car that presented a risk to the driver from 3 possible orientations in intersections (i.e., car turning left, on-coming car, or car entering from the right). One group of participants were given a driving related task and the control group were given a non-driving related task. Results indicated that participants given a driving goal acted like drivers when viewing the scene, tapping into different abilities than just detection (Koustanaï et al., 2012). They performed better that the control group and change detection was strongly correlated with the difficulty of the task. Surprisingly, participants had more difficulty detecting a car on the right at an intersections in less detailed, rural settings, which contrasted with findings from other studies showing increased difficulty with visual complexity (Beck & Levin, 2003; Beck et al., 2007a; Wright, Green, & Baker, 2000). The authors suggested this might be because low contrast cars were used, which are harder to see in rural settings. Also the rural settings provided less structure and context for the driving task (e.g., traffic lights and road markings) (Koustanaï et al., 2012). This study did not show any particular impact of variously orientated vehicles on performance, which is intuitive, considering the complexity and wide potential variation inherent in the driving task.

Observer Effects

Change detection requires encoding items into memory then comparing them over time (Jensen et al., 2011). Individual differences in working memory, selective attention and information processing speed may therefore impact change detection performance (Jensen et al., 2011). Jensen et al. (2011) proposed that an assessment tool that could measures these aspects might be useful in predicting "noticing" in real driving situations.

Impact of chronological age. It is well recognised in fitness to drive literature that advancing age impacts driving performance. A model capable of measuring the impact of age related changes on performance could be practically applied to assessment of cognitive fitness to drive. Results from change blindness experiments indicate older drivers are slower to detect change (Caird et al., 2005; Pringle et al., 2001; Rizzo et al., 2009), demonstrate reduced scanning compared to younger drivers (Caird et al., 2005), and have particularly low accuracy for detecting pedestrians and traffic sign changes, resulting in incorrect decisions (Caird et al., 2005). Bédard et al. (2006) proposes that attention and perception are the key determinants of driving performance among older drivers: impacted by task complexity and the need to divide attention. However, results form a number of studies indicate that only some aspects of attention are compromised with age: primarily visual search, memory retrieval, and selective attention (Folk & Hoyer, 1992; Folk & Lincourt, 1996; Foster, Behrmann, & Stuss, 1995; Hasher & Zacks, 1998; Hoffman et al., 2006; Scailfa & Joffe, 1997). These aspects of attention are particularly compromised where there is a high number of distractors or a task time pressure (Becker & Rasmussen, 2008; Brink & McDowd, 1999; McCarley et al., 2004), as in the task of driving.

Hasher and Zacks (1998) conclude in their review of literature on aging and memory, that older adults respond to age-related deficits in memory retrieval and selective attention inhibitory mechanisms by relying more heavily on environmental cues and the immediate array; and by making greater use of personal experience and knowledge for interpretation. This is consistent with results of a change blindness study conducted by Caird et al. (2005), indicating older drivers rely heavily on traffic control devices such as traffic lights to make decisions, often missing other important objects such as pedestrians. Caird et al. (2005) suggested this might be due to older drivers developing strategies to cope with complex intersections by focusing on the immediate most relevant objects (e.g., traffic devices) but at the risk of missing other important information. Hasher and Zacks (1998) propose that similar effects would likely be involved for younger adults who have been diagnosed with medical conditions affecting the same mechanisms.

Impact of medical conditions. The change blindness model would have important application in the assessment of fitness to drive if the underlying mechanisms allowed reliable measurement of the impact of medical conditions on attention and change detection in driving scenes. A limited number of studies have applied the change blindness model to specific diagnostic groups. The change blindness model has been applied to investigate visual awareness in children diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) (Maccari et al., 2013; Turkan, Amado, Ercan, & Percinel, 2016), children and adults diagnosed with autism spectrum disorder (ASD) (Fletcher-Watson et al., 2012), adults diagnosed with schizophrenia (Grandgenevre et al., 2015), and older drivers diagnosed with Alzheimer's disease (Rizzo et al., 2009).

ADHD is one of the most common childhood psychiatric conditions and inattention is the most commonly studied symptom of ADHD (Maccari et al., 2013; Turkan et al., 2016). Early research indicates that the change-detection model is useful for studying differences in attentional preferences between typically developing children and children diagnosed with

ADHD. Maccari et al. (2013) reported finding the change blindness model more effective for assessing visual search efficiency and focused attention because traditional psychometric tests are lengthy and disengaging for children with ADHD, resulting in loss of interest and demotivation, thus causing an over-estimation of difference compared to typically developing children (Maccari et al., 2013). Conversely video games utilised by other researchers to measure attention are highly engaging, arousing, and interesting, potentially minimizing the gap (Maccari et al., 2013). Maccari et al. (2013) found that children with ADHD showed specific impairment in the top-down search strategies typically useful for solving complex tasks, which may contribute to their limited attentional resources. Additionally children with ADHD were also slower and less accurate in detecting change compared to typically developing children (Maccari et al., 2013). In a study of eye movements during change detection tasks, Turkan et al. (2016) found that typically developing children had longer fixations on the change area and longer fixation maintenance compared to children with ADHD. These findings suggests children with ADHD may have greater difficulty detecting change due to their attentional deficits and difference in voluntary eye movement control (Turkan et al., 2016).

Change detection has been used in a series of studies of attentional preferences of children and adults diagnosed with autism spectrum disorder (ASD). The methodology has been used to study the presumed local-processing bias of people with autism (i.e., increased attention to detail, particularly of insignificant aspects) (Fletcher-Watson et al., 2012). Findings are mixed with some studies showing slower detection time for marginal interest items (Fletcher-Watson, Leekam, Turner, & Moxon, 2006) and lack of the impact of context (Loth, Carlos Gomez, & Happe, 2008), and others showing an advantage in detection for people diagnosed with ASD, particularly for marginal interest items. Attention to social information, assumed to be impaired in ASD, has also been studied, with the surprising findings that there is no difference in detection rates for social versus non-social information for adults with ASD (Fletcher-Watson, Leekam, Findlay, & Stanton, 2008; Smith & Milne, 2009). Further research is required to reconcile this with the reduced attention to social information revealed via other methods of research (Fletcher-Watson et al., 2012).

Rizzo et al. (2009) investigated whether advanced age and Alzheimer's disease (AD) increased change blindness in driving scenes. An increase was anticipated due to likely degenerative changes in areas of the brain responsible for vision, allocation of attention, and change detection. A unique change detection method was applied where participants were shown a static, naturalistic driving scene (taken from the perspective of a driver) with a slow fading in and out of the changed object. Only one change occurred per image pair (e.g., appearance of a car in the left lane). The observer was asked to identify and touch the location of the change as quickly and accurately as possible on the screen. Performance was measured via hit rate, percentage of true positives, false positives rate, number of catch trials where a change was reported and response time. Results from the cognitively health group ($N = 68$, aged 20 - 84) indicated that as age increased, hit rate decreased, with an acceleration from age 68. False positives did not increase and there was no significant correlation in hit rate. Response times were significantly worse for the AD group ($M = 7.42$s) and the older controls ($M = 5.72$s; $M = 1.54$ age adjusted). Change blindness was correlated with poor performance on other cognitive screens including TMT (A & B) and UFOV^{TM} for both the older and AD groups. Results suggest aging reduces capacity to perceive visual change and AD decreases this further (Rizzo et al., 2009). Results of this study indicate the fade in and out change blindness method is applicable and useful for driving related research.

The study of the impact of medical conditions on awareness via the change blindness model is in its early stages, particularly for driving. However, results from the studies highlighted indicate that change detection tasks may successfully differentiate between individuals with cognitive impairment in the context of a driving related task.

Impact of experience. Individuals differ in their level of prior knowledge and experience related to the tasks they perform. Much evidence from the change-detection model indicates that observers are able to draw on previous knowledge and experience to more efficiently guide attention (Beck, Martin, Smitherman, & Gaschen, 2013; Crundall & Underwood, 1998; Jones, Jones, Smith, & Copley, 2002; Reingold et al., 2001; Werner & Thies, 2000; Zhao et al., 2014). One experiment compared experts and novices in the domain of football, finding that experienced observers could attentionally scan scenes more quickly to identify changes meaningful to play, whereas rates of detection for non-meaningful change were the same as for the novice group (Werner & Thies, 2000). Comparable results were found for chess players (Reingold et al., 2001), veterinary medicine students (Beck et al., 2013), veterinarian radiologists (Bass & Chiles, 1990; Beck et al., 2004; Beck et al., 2013) and drug users: who were more likely to notice changes to drug paraphernalia in photographs than non-drug users (Jones et al., 2002).

Similarly, for the domain of driving, the impact of experience on change detection for drivers versus non-drivers has been investigated showing drivers search a larger area more efficiently, with fewer eye movements and more focus on relevant objects that require monitoring (Crundall & Underwood, 1998; Crundall, Underwood, & Chapman, 1999; Summala, Lamble, & Laakso, 1998; Summala et al., 1996; Zhao et al., 2014). These studies indicate awareness of visual information is conditioned by real-world experience. People do not intuitively know where to expect hazards in a driving scene but must be trained in effective search patterns as a learner driver, then practice.

Intuitively we expect people to continue improving the longer they perform a task. However, evidence suggests the benefit of experience plateaus within the first 1-3 years of practice (Beck et al., 2013; Miglioretti et al., 2009). The largest change in ability occurs in the first few years, after which lower-order functions become more automated (Beck et al., 2013; Miglioretti et al., 2009). This is consistent with the task of driving where eventually the task becomes over-learned and easy, resulting in experienced drivers tending to share time and attentional resources with other activities (e.g., mobile phone use or eating), increasing risk of accident (Summala et al., 1996). The importance of practical training was highlighted in a study conducted by Becks et al. (2013). The effect of experience was only noticeable among undergraduate radiology undergraduates once they had attended practical and specific training in radiology (i.e., senior medical students who had attended a radiology rotation). Performance was not impacted by years of experience after this stage when compared to experienced radiology staff (Beck et al., 2013).

Findings from these studies indicate awareness of visual information is learned and training can impact change detection performance on the specific level but there is no evidence yet that training can be generalized (Beck et al., 2004; Gaspar, Neider, Simons, McArely, & Kramer, 2013). However it is clear that experience has a top-down effect on change detection (Caird et al., 2005; Jensen et al., 2011). It is well recognised that driving experience impacts driving performance. The capacity of the change blindness model to measure the impact of experience highlights the utility of the model for assessing driving performance.

Impact of goals / task. Change blindness experiment results indicate the intention of the observer will impact the degree to which they expect the change, which will affect the mechanisms used in change detection (Koustanaï et al., 2008; Rensink, 2000). Researchers

(Caird et al., 2005; Hoffman et al., 2006; Koustanaï et al., 2012; Pringle et al., 2001; Schömig, Metz, & Krüger, 2011) who apply the change blindness model to driving often introduce a driving related activity (e.g., search for hazards) so the participant views the scene from the perspective of a driver and the task is more representative of real-world driving. Most change blindness research uses an intentional approach where the observer is told to search for a change and is expecting it. Rensink (2002) advises this approach for assessment of perceptual capacities. Alternatively, a divided-attention approach involves giving the viewer a secondary task, such as judging when it is safe to go in traffic (Caird et al., 2005; Crundall, 2009). The search for change may be part of the driving related task (e.g., detecting the car in front stopping) or incidental, with no warning that a change will occur. Research shows change detection is much harder with an incidental approach, although some change is still noticed (Rensink, 2002).

Some researchers (Schömig et al., 2011; White & Caird, 2010) have applied the change blindness model to explore the impact of the driver's goal on driving safety. Schömig et al., (2011) demonstrated that drivers were able to engage in a secondary task in an aware manner and appropriately defer the task in highly demanding traffic situations. However, they questioned whether these results would be the same for highly motivating tasks (e.g., checking a phone text) or if the driver underestimated the risk of engaging in the task (e.g., adjusting the radio). White & Caird (2010) explored these issues further using a change blindness experimental design to evaluate the impact of passenger conversation on hazard detection, measuring the impact of passenger attractiveness and the driver's level of extraversion. The authors found that drivers where more distracted by conversation with attractive passengers and extraverted drivers were more easily distracted. Error rates and frequency of look-but-failed-to-see accidents (measured in a driving simulator) increased in these situations (White & Caird, 2010).

Conclusion

There is clear evidence that the change blindness model is able to measure the impact of factors critical to safe driving such as age-related changes, experience, and health status, supporting the application of this model to the development of fitness-to-drive tests for health professionals (Caird et al., 2005; Crundall, 2009; Hoffman et al., 2006; Wetton et al., 2010). However, further research is required to examine the psychometric properties and predictive validity of such tests so they are evidence based and valid for clinical practice.

Appendix C

DriveSafe DriveAware Digital Development:

General Practitioner Survey Findings

Pearson Clinical Assessment (Australia and New Zealand) conducted market research prior to the DriveSafe DriveAware (DSDA) digital development, to determine if GPs perceived a need for a fitness-to-drive screen and how they preferred the test be designed for medical practice. AFS Smart Askers, a market research company with a large database of GPs who provide paid opinions regarding various topics, conducted the research. I commenced employment with Pearson in May 2012 and enrolled in my PhD in March 2013. The market research commenced prior to my student enrolment and was organised and funded by Pearson. Therefore, I did not include it as part of my thesis. However, I did participate in the survey design, analysis of findings, and write-up of results.

Study Authors: Fiona Brown, Beth Cheal, Melinda Cooper, & Nicki Joshua

Affiliation: All authors were Pearson Clinical Assessment employees (Sydney, Australia)

Report Date: May 2013

AFS Smart Askers sent 866 surveys to Australian general practitioners (GPs) registered in their database. A representative sample was selected based on average GP age, gender, and location by state, nationally. The survey closed once 200 responses were received. Incomplete surveys were not included in the sample (n = 13). Surveys were completed via an online questionnaire with a short presentation of the proposed DriveSafe

DriveAware digital test design presented part pay through the questionnaire. No identifying information was collected. Demographic information collected included location by state (see Figure 1), age, gender, occupation, years of employed in general practice, practice size, patient age demographic, type of patients, and number of patients seen per week.

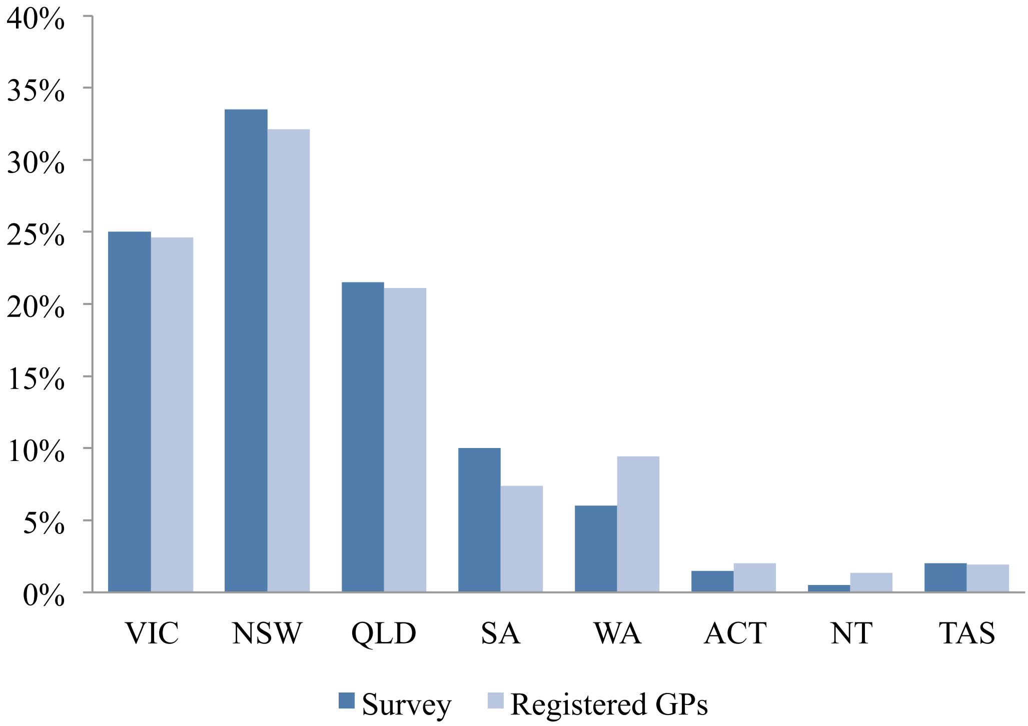


Figure 1. Participant location by state, in comparison to location of registered GPs nationally.

The sample consisted of 76 females (38%) and 124 males (62%). Participant age and years of experience working as a physician are presented in Table 1.

Table 1.

Participant Age and Experience

Age (years)	Sample (%)	Experience (years)	Sample (%)
30-39	15	2-5	3
40-49	28	6-10	15
50-59	41	11-20	28
60-69	11	>20	54
70+	4		
Prefer not to answer	1		
Total	100		100

The average number of GPs per practice was six. GPs reported seeing an average of 149 patients per week and performing an average of 87 fitness-to-drive assessments per year (*M* = 50; Mode = 50). GPs estimated that 25% of their patients were aged 60 to 74 and 16% were aged 75+ on average. They estimated that around 6% experienced memory loss or were diagnosed with dementia and around 5% had been diagnosed with other neurological conditions affecting cognition (e.g., stroke or brain injury) on average. GPs reported 22% of their patients had an enhanced primary care plan on average.

Methods of Assessing Fitness to Drive

GPs were asked to rate how frequently they relied on the listed indicators (professional judgement, patient observations such as mobility and grooming, reference to medical guidelines, family report, formal assessment results, and outcome of an occupational therapy driving assessment) to determine fitness to drive. They were also provided with an "other" free-text category. The three indicators with the highest level of agreement were: (a) reliance on professional judgement ("always" = 64%; "often" = 30%); (b) observations of patient factors such as mobility, grooming and behaviour ("always" = 53%; "often" = 34%); and, (c) reference to Austroads medical guidelines (Austroads & National Transport Commission, 2016) ("always" = 13%; often = 37%). GPs also reported questioning the patient's family about their concerns ("often or always" = 39%; "sometimes" = 43%). Some doctors (*n* = 14) reported in the free-text field that they relied on patients or their family to provide information about driving safety but noted this was not a reliable source of information: "A lot depends on hearsay from family and the patient which is not always reliable"; "No penalty for false or lack of information by patients"; and, "Difficult to assess an older patient with some physical impairments…without a family member to verify that they are not having problems". Other methods listed in the "other" category were general medical examinations (*n* = 17); specialist opinion (*n* = 11); vision assessment (*n* = 8); liaison with the licensing authority or license authority testing (*n* = 4); use of standardised assessments (*n* = 3); and, evaluating use of alcohol and illegal drugs (*n* = 2). Most GPs (63%) reported rarely or never referring for an occupational therapy driving assessment; 32% reported sometimes referring. Only 6% reported often referring. Comments from some participants (*n* = 9) indicated that cost and access might be a barrier to occupational therapy driving assessment referral. Two respondents stated they did not agree with the assessment outcome. Others stated that occupational therapy driving assessments should be paid for by the government and be mandatory for drivers aged 75+.

Use of standardised assessment. Only 27% of GPs agreed that they used standardised assessments to determine fitness to drive. GPs were asked in a separate question to list the standardised assessment they used generally in their medical practices. The most commonly listed test was MMSE (Folstein et al., 1975): 79% of doctors reported using it.

Few other standardised assessments were listed. These included the Anxiety and Depression Checklist K10 (*n* = 8); the Epworth Sleepiness Scale (*n* = 4); Geriatric Depression Scale (GDS) (*n* = 3); General Practitioner assessment of Cognition (GPCOG) (*n* = 3); Clock Drawing Test (*n* = 2); Maze test (*n* = 2); DriveAble (*n* = 2); MoCa (*n* = 1); and Romberg's test (*n* = 1). Nearly 30% of doctors reported using formal vision tests such as the Snellen's eye chart and visual fields tests; 4 doctors reported using audiometry test.

Satisfaction with Methods of Assessing Fitness to Drive

GPs were asked to rate their satisfaction with their current methods of determining fitness to drive: 51% were somewhat satisfied; 26% reported little to no satisfaction. Only 23% were very satisfied. Most (77%) GPs agreed that they were confident in their ability to evaluate patient fitness to drive and had the tools they needed (60%); 40% did not agree they had the tools required. However, when GPs were asked to describe the shortcomings of their current methods of assessing fitness to drive in open-text responses commonly (*n* = 76) described a lack of sufficient information for making fitness-to-drive determinations. Some participants expressed concern about the medical guidelines and license authority forms (*n* = 38). The medical guidelines were described as "cumbersome", "not very user friendly" and "ambiguous". Typical responses included: "Too focused on broad disease categories and not on cognitive function"; "difficult to use in the setting of a GP consultation. There are often time constraints…the guidelines are not very user friendly and there can be scope for misinterpretation"; and, "Chronic conditions which are not on the list but may cause the driving to be unsafe, cause stress for the doctor when assessing fitness to drive". GPs recommended simplification of forms and better information to guide their decision making: "I am satisfied with the methods of formal OT assessments but I am not satisfied with the lack of uniformity in guidelines as to which drivers are to be assess or referred".

Perceptions Regarding Driving Assessor Role

GPs were provided with an open-text field to provide further comment at the conclusion of the survey. Some GPs (*n* = 22) responded that a third party should be doing the fitness-to-drive assessments or that a practical driving assessment should be given; others commented that GPs should not be determining fitness to drive at all (*n* = 15). Rather, an independent body, such as doctors employed by the licensing authority, should be making the determinations. Following are examples of common responses: "Demand the bureaucrats provide their own doctors to do the testing, as we often have an interest conflict in doing these tests"; "I feel I don't have the expertise to accurately assess driving ability in some circumstances"; and, "It is ridiculous to expect that a GP will be able to decide who is fit or otherwise to drive". Most (85%) of GPs agreed or strongly agreed with the statement that they were concerned regarding their legal liability in assessing patient fitness to drive.

Most GPs disagreed with the statement "I am concerned that advising cessation of driving results in loss of business for my practice" ("Disagree" *n* = 50%; "Strongly Disagree" *n* = 32%). However, most (83%) agreed with the statement "It concerns me that revoking a patients' license often leads to negative consequences for them". In the free-text field's 4 GPs reported reluctance to withdraw driving due to the potential impact on the patient's quality of life. Responses included: "There is a tendency of the GP to sign the patient as fit to drive based on social circumstances, e.g., if patient is not able to drive, he/she cannot come for medical check-ups…do everyday shopping" and "If I feel the patient is not fit I do not want to be seen as the person denying the person the privilege, convenience". Some GPs reported concern regarding certifying someone as fit to drive when they were seeing them as a first-time patient without an adequate medical history: "Difficult for first time unfamiliar patients…patients tend to downplay their symptoms of medical problems".

Most (65%) of GPs agreed that they felt unduly pressured to find their patients fit to drive. Free-text responses from 28 doctors indicated they were concerned about negative reactions from patients if they were advised to cease driving or attend driving or medical specialist assessments. Typical responses included: “Too much pressure on me to allow driving”; “Lot of pressure on doctors to oblige the wishes of the patient”; and “The worst is the pressure put on me to agree to a license when I do not agree they are fit”. The following response describes the tension often felt by doctors: “As the patient’s trusted doctor it can be really difficulty to broach the possibility of not being able to drive any longer. They are often extremely offended and resent you for breaking their trust as they see it”; “Patients get angry when you certify them as unfit to drive”; and, “Causes huge difficulties in doctor-patient relationship”. Some GPs (n = 14) reported a perceived conflict of interest in assessing fitness to drive: “It is a major decision to make that impacts on my patient’s quality of life and since they almost always people who are well known to me. I am sometimes sure that I may not remain impartial and objective”; “there is always a tendency for doctors to be over lenient”; and, “GPs don’t like failing their patients. Many of those patients have seen their GPs for many years”. Three doctors reported that if they did not certify their patient as fit to drive they would “just go down the road to the next doctor and get a license!”

Perceptions Regarding Digital DriveSafe DriveAware

GPs were given a brief demonstration of the proposed format of a digital version of DSDA then asked if and how they would use the test. Most participants indicated they would be likely to use the test (84%); 16% were undecided or indicated they would not use the test. Comments included: “Looks good, close to real driving situations without being in the car. Better than the current method which is nothing”; “Standardises the assessment…would help in convincing the patient they need further assessment – without blaming the doctor for this”;

“I am amazed about how old people like iPads as compared to laptops”; and “Keen to embrace objective assessment tools especially those that have been proven as this reduces my liability and improves safety”.

GPs were asked what factors would determine their decision to adopt a fitness-to-drive test like DriveSafe DriveAware. Around 30% (*n* = 59) identified the time taken to complete the test. GPs were also concerned about the assessment cost and how to pass this on (*n* = 43). Typical comments included: “If a Medicare rebate is involved I can see it working. Most of the assessments are on pensioners and they would not pay for the test if it was not mandatory” and “I would expect Transport to pay for the tool: I would advise patient that this service will NOT be bulk billed. Many will be singularly unimpressed”. Some doctors compared the cost of the test to an occupational therapy driving assessment and considered it would be feasible: “Far less expensive and problematic that formal road driving assessments”. Other doctors (*n* = 28) commented that they would need to be convinced regarding the validity and reliability of the test before they would use it and or commented that it must be user-friendly (*n* = 28): “Not sure if elderly driver can handle this”, “Concerned older patients may find it threatening” and “I like the idea but older people are afraid of new technology and are not very quick to do something in 3 seconds”.

GPs specifically addressed how the test would fit in the context of license authority guidelines and their legal obligations in free-text responses (*n* = 25). Some doctors wanted assurance that DSDA would be approved by licensing authorities and insurers (e.g., “I am interested to hear what the roads licensing authorities think of it”; “Looks interesting and good but I would need to trial it in-clinic and assess VicRoads opinion in regards to the test”; and, “It would depend on the legal acceptance of the test”. GPs wanted DSDA formally integrated into license authority guidelines and practice: “Patients would be unlikely to do it if it was not mandated"; "Great idea. However this should be a mandatory annual test for 75 plus drivers – not just the overworked, overused GP and nurse duo"; and, "Really should be a community based requirement for older drivers".

Most (84%) GPs reported they had access to a practice nurse who could administer the test. They preferred in-house administration, with the doctor, a practice nurse, or a combination of both administering the test (84%). Few participants (10%) reported they would prefer to refer to an external provider to complete the test and only 6% considered the test would not be feasible for them. Most GPs (78%) reported access to an iPad. Only 7% preferred android administration.

Conclusion

Results of the survey indicated that GPs perceive a need for a fitness-to-drive screen and would consider using a digital version of DSDA if it was brief, valid, user-friendly and practical for medical practice. GPs saw the test as advantageous because it potentially reduced conflict with the patients by shifting the fitness-to-drive determination from them, and because it provided additional medico-legal support. However, GPs were concerned about how the test would be funded and wanted it to be covered by Medicare. It was also important to them that the test to be integrated within current licensing authority standards and guidelines, so the patient would agree to undergo testing.

Survey Questions

1. Approximately how many assessments of fitness to drive would you do in a year?
2. How frequently do you use each of the following to assess fitness to drive?
    a. Your professional judgement
    b. Observed factors such as patient mobility, grooming and behaviour
    c. Refer to Assessing Fitness to Drive Medical Guidelines (Austroads)
    d. Question family about their concerns
    e. Formal assessment (e.g., Mini-Mental State Examination)
    f. Referral for an occupational therapy driving assessment
3. How satisfied are you with your current methods of assessing fitness to drive (*5 point Likert scale*).
4. Why? What are the shortcomings of the current methods of assessing fitness to drive?
5. Please indicate your level of agreement with the following statements:
    a. I am concerned about my legal liability in assessing fitness to drive.
    b. Advising cessation of driving negatively impacts the doctor-patient relationship.
    c. I feel unduly pressured by patients to report them as fit to drive.
    d. It concerns me that revoking a patients' license often leads to negative consequences for them.
    e. I have the assessment tools I need to assess fitness to drive.
    f. I am concerned that advising cessation of driving results in loss of business for my practice.
    g. I am confident in my ability to evaluate a patient's fitness to drive.

Brief demonstration of proposed DSDA fitness-to-drive screen

6. How long would you be willing to take in administering this test face-to-face? (Insert 0 if you would not administer the test).
7. How long would you be willing for your practice nurse to spend administering this test? (Insert 0 if not applicable).
8. Which of the following scenarios would be feasible within your practice (select all that apply)
    a. Doctor administers the entire test
    b. Practice nurse administers the entire test
    c. Reception / administration staff administer the entire test
    d. Practice nurse supervises the tests and the doctor conducts the interview
    e. Reception / administration staff supervises the tests and the doctor conducts the interview
    f. Patient completes the tasks unsupervised in the waiting area and the doctors conducts the interview
    g. Patient completes the tasks unsupervised in a consulting room and the doctor conducts the interview.
    h. I would refer patients to an external provider (e.g., occupational therapist or psychologist) to administer the entire test and forward the results to me.
    i. None of the above, but I may still use this test in my practice.
    j. None of the above, this product is not feasible for my practice.
9. How would you handle the cost of the assessment?
    a. No direct cost to the patient (I would absorb the cost)
    b. Integrate into the patient consultation fee
    c. Charge a separate fee on top of the usual consultation fee

d. Refer to an external service who will bill the patient directly (e.g., Occupational therapist or psychologist)
e. Other – please specify

10. Please mark all that apply:
    a. I own or would be willing to buy an iPad that patients could use to do this test.
    b. The practice owns or would be willing to purchase an iPad that patients could use to do this test.
    c. I would use this test if it was on an android device as I own or am likely to buy one.
    d. None of the above.
11. Do you currently use any of the following devices in your practice? Please select all that apply: laptop computer, desktop computer, iPad (full size), android tablet (comparable to iPad size).
12. Please feel free to add any additional comments you have about the test.

(Additional questions were asked related to marketing)

Appendix D

Ethics Application Documentation

Research Integrity
Human Research Ethics Committee

Wednesday, 21 August 2013

Professor Anita Bundy
Health Systems and Global Populations; Faculty of Health Sciences
Email: anita.bundy@sydney.edu.au

Dear Professor Anita Bundy,

I am pleased to inform you that the University of Sydney Human Research Ethics Committee (HREC) has approved your project entitled **"DriveSafe and DriveAware - A Valid iPad Application for Physicains to Use to Determine Fitness to Drive?"**.

Details of the approval are as follows:

Project No.:	**2012/2812**
Approval Date:	**20 August 2013**
First Annual Report Due:	**21 August 2014**
Authorised Personnel:	**Bundy Anita; Patomella Ann-Helen; Cheal Beth; Rowe Dominic; Brown Fiona**

Documents Approved:

Date Uploaded	**Type**	**Document Name**
09/08/2013	Questionnaires/Surveys	Questionnaire - Doctors
09/08/2013	Questionnaires/Surveys	Questionnaire - Seniors
09/08/2013	Advertisements/Flyer	Questionnaire - Doctors Phase 3
09/08/2013	Participant Info Statement	Adjusted PIS - Patients
09/08/2013	Participant Info Statement	Adjusted PIS - Students
09/08/2013	Participant Info Statement	Adjusted PIS - Doctors
09/08/2013	Participant Info Statement	Adjusted PIS - Doctors Phase 3
09/08/2013	Participant Info Statement	Adjusted Participant Information Sheet - Seniors
09/08/2013	Participant Consent Form	Adjusted Consent Form - Doctors Phase 1
09/08/2013	Participant Consent Form	Adjusted Consent Form - Doctors Phase 3
09/08/2013	Participant Consent Form	Adjusted Consent Form - Patients Phase 2
09/08/2013	Participant Consent Form	Adjusted Consent Form - Seniors Phase 1
09/08/2013	Participant Consent Form	Adjusted Consent Form - Students Phase 4
09/08/2013	Safety Protocol	Safety Protocol Signature 2
09/08/2013	Safety Protocol	Safety Protocol National & International

HREC approval is valid for four (4) years from the approval date stated in this letter and is granted pending the following conditions being met:

<u>Condition/s of Approval</u>

- Continuing compliance with the National Statement on Ethical Conduct in Research Involving Humans.

Research Integrity
Research Portfolio
Level 2, Margaret Telfer
The University of Sydney
NSW 2006 Australia

T +61 2 8627 8111
F +61 2 8627 8177
E ro.humanethics@sydney.edu.au
sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

THE UNIVERSITY OF SYDNEY

- Provision of an annual report on this research to the Human Research Ethics Committee from the approval date and at the completion of the study. Failure to submit reports will result in withdrawal of ethics approval for the project.

- All serious and unexpected adverse events should be reported to the HREC within 72 hours.

- All unforeseen events that might affect continued ethical acceptability of the project should be reported to the HREC as soon as possible.

- Any changes to the project including changes to research personnel must be approved by the HREC before the research project can proceed.

Chief Investigator / Supervisor's responsibilities:

1. You must retain copies of all signed Consent Forms (if applicable) and provide these to the HREC on request.

2. It is your responsibility to provide a copy of this letter to any internal/external granting agencies if requested.

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

S. J. Assinder

Dr Stephen Assinder
Chair
Human Research Ethics Committee

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007), NHMRC and Universities Australia Australian Code for the Responsible Conduct of Research (2007) and the CPMP/ICH Note for Guidance on Good Clinical Practice.

THE UNIVERSITY OF SYDNEY

Research Integrity
Human Research Ethics Committee

Tuesday, 15 July 2014

Prof Anita Bundy
Health Systems and Global Populations; Faculty of Health Sciences
Email: anita.bundy@sydney.edu.au

Dear Anita

Your request to modify the above project submitted on 30/6/2014 was considered by the Executive of the Human Research Ethics Committee at its meeting on **9 July 2014**

The Committee had no ethical objections to the modification/s and has approved the project to proceed.

Details of the approval are as follows:

Project No.:	**2012/2812**
Project Title:	**DriveSafe and DriveAware - A Valid iPad Application for Physicains to Use to Determine Fitness to Drive?**

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

Helen Mitchell

Chair
Human Research Executive Committee

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007), NHMRC and Universities Australia Australian Code for the Responsible Conduct of Research (2007) and the CPMP/ICH Note for Guidance on Good Clinical Practice.

Research Integrity
Research Portfolio
Level 2, Margaret Telfer
The University of Sydney
NSW 2006 Australia

T +61 2 8627 8111
F +61 2 8627 8177
E ro.humanethics@sydney.edu.au
sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

St Vincent's Hospital

A facility of St Vincents
& Mater Health Sydney

St Vincent's Hospital Sydney Ltd
ABN 77 054 038 872
390 Victoria Street
Darlinghurst NSW 2010
Australia

T + 61 2 8382 1111
F + 61 2 9332 4142
www.stvincents.com.au

18 September 2013

Prof Anita Bundy
Room J010
C42 - Cumberland Campus
University of Sydney
Lidcombe NSW 2141

Dear Anita

SVH File Number: 13/196
Project Title: DriveSafe and DriveAware iPad Application: A Valid Fitness to Drive Cognitive Screening Tool for Physicians?
HREC Reference Number: HREC/13/SVH/294

Thank you for submitting the above **multi centre** project which was first considered by the St Vincent's Hospital HREC at its meeting held on **12 September 2013**. This HREC has been accredited by NSW Ministry of Health as a Lead HREC under the model for single ethical and scientific review and Certified by the NHMRC under the National model for Harmonisation of Multicentre Ethical Review (HoMER).This lead HREC is constituted and operates in accordance with the National Health and Medical Research Council's *National Statement on Ethical Conduct in Human Research* and the *CPMP/ICH Note for Guidance on Good Clinical Practice*. No HREC members with a conflict of interest were present for review of this project.

Your response will be reviewed by the HREC Executive Committee.

In order to make a determination of the ethical acceptability of your project, please respond to the following request for additional information and/or modification.

With regard to the project application and design, the committee request that the following be addressed (National Statement Section 1 – Values and Principles of Ethical Conduct and Chapter 3.3):

1. The Committee noted that the iPad version of DriveSafe and DriveAware is partially self-administered by the patient. Please clarify what measures are in place to ensure that the test is actually performed by the patient and not a third party.
2. The Committee noted that the aim of this study phase is to assess whether doctors are able to administer the iPad version, however, this phase of the study involves administration only by driver-trained occupational therapists. Please clarify how ease and validity of administration by clinicians will be assessed.
3. Please clarify if results of the DriveSafe and DriveAware test, or the MiniMental Status Examination, will be made available to participants.

With regard to the Participant Information Sheet and Consent Form – Main, please address the following (National Statement Section 2: Themes in Research Ethics: Risk and Benefit, Consent):

1. Please reformat this document using the approved St Vincent's template Participant Information Sheet and Consent Form, using standard sections and wording. A template is available via the Research Office website.
2. In reformatting the document, please ensure that a description the DriveSafe and DriveAware test is included, for example, how and why it is administered, and how it differs from standard practice. Please also include information as to why the MiniMental State Examination is being performed, as this is not part of standard evaluation.
3. Please delete the sponsor logo from the header of the document.
4. Further modification may be requested following resubmission of the document.

Please refer to the National Statement for relevant guidelines.

Please contact the Research Office if you require a Word version of the above points in order to assist in preparation for your response. In order to facilitate the Committee's consideration of your project, **only responses that address all the above issues will be accepted.**

Your response should be submitted as ***one* hard copy** of all amended documents *in addition* please email documents to the Executive Officer (research@stvincents.com.au) quoting the SVH file reference in the subject heading. **Amended versions of documents must be submitted in both tracked changes format and clean copies. Please include version number and date on all submitted documents.**

Please note that if the requested information is not received within 3 months, the project will be dismissed and you will be required to re-submit the project at a later date.

Should you have any queries about your project please contact the Research Office, Tel: 8382-2075, email research@stvincents.com.au. The HREC Terms of Reference, Standard Operating Procedures, *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice* and standard forms are available on the Research Office website: www.stvincents.com.au/researchoffice or internal at http://exwwwsvh.stvincents.com.au/researchoffice

Please quote **SVH file number: 13/196** in all correspondence.

Yours sincerely

Ellie Pratt
Acting HREC Executive Officer
St Vincent's Research Office
Level 6 deLacy Building

cc. Beth Cheale

D/2013/51465

St Vincent's Hospital

A facility of St Vincent's
& Mater Health Sydney

St Vincent's Hospital Sydney Ltd
ABN 77 054 038 872
390 Victoria Street
Darlinghurst NSW 2010
Australia

T + 61 2 8382 1111
F + 61 2 9332 4142
www.stvincents.com.au

18 March 2014

Prof Anita Bundy
Room J010
C42 - Cumberland Campus
University of Sydney
Lidcombe NSW 2141

Dear Anita

SVH File Number: 13/196
Project Title: DriveSafe and DriveAware iPad Application: A Valid Fitness to Drive Cognitive Screening Tool for Physicians?
HREC Reference Number: HREC/13/SVH/294

Thank you for providing correspondence from Ms Laijing Lee, Pearson Legal Counsel, dated **20 February 2014**, responding to issues raised regarding the above project at a meeting of the HREC Executive on **24 January 2014.** This HREC has been accredited by NSW Ministry of Health as a Lead HREC under the model for single ethical and scientific review and Certified by the NHMRC under the National model for Harmonisation of Multicentre Ethical Review (HoMER). This lead HREC is constituted and operates in accordance with the National Health and Medical Research Council's *National Statement on Ethical Conduct in Human Research* and the *CPMP/ICH Note for Guidance on Good Clinical Practice*. No HREC members with a conflict of interest were present for review of this project.

The comments regarding Pearson's privacy statement were reviewed by the HREC Executive at a meeting on **11 March 2014** and I am pleased to advise that the Committee has granted ethical and scientific approval of the above **multi centre** project.

You are reminded that this letter constitutes *ETHICAL* and *SCIENTIFIC* approval only. You must not commence this research project at a site until a completed Site Specific Assessment Form/Access Request and associated documentation have been submitted to the site Research Governance Officer and authorised. A copy of this letter must be forwarded to all site investigators for submission to the relevant Research Governance Officer.

The project is approved to be conducted at:
- St Joseph's Hospital, Auburn
- Calvary Rehabilitation and Geriatric Service (CRAGS), Kogarah, NSW

If a new site(s) is to be added please inform the HREC in writing and submit a Site Specific Assessment Form (SSA) to the Research Governance Officer at the new site.

The following documentation has been reviewed and approved by the HREC:
- Study protocol version 1, 4 September 2013
- Participant Information Sheet and Consent Form version 2, 16 December 2013

The following document was noted:
- Safety protocol version 1, 9 August 2013

The National Ethics Application Form (NEAF) document reviewed by the HREC was NEAF **AU/1/A55416**.

Please note the following conditions of approval:

- HREC approval is valid for **5 years** from the date of the HREC Executive Committee meeting and expires on **11 March 2019.** The Co-ordinating Investigator is required to notify the HREC 6 months prior to this date if the project is expected to extend beyond the original approval date at which time the HREC will advise of the requirements for ongoing approval of the study.
- The Co-ordinating Investigator will provide an annual progress report beginning in **March 2015**, to the HREC as well as a final study report at the completion of the project in the specified format.
- The Co-ordinating Investigator will immediately report anything which might warrant review of ethical approval of the project in the specified format, including unforeseen events that might affect continued ethical acceptability of the project and any complaints made by study participants regarding the conduct of the study.
- Proposed changes to the research protocol, conduct of the research, or length of HREC approval will be provided to the HREC for review, in the specified format.
- The HREC will be notified, giving reasons, if the project is discontinued before the expected date of completion.
- Investigators holding an academic appointment (including conjoint appointments) and students undertaking a project as part of a University course may also be required to notify the relevant University HREC of the project. Investigators and students are advised to contact the relevant HREC to seek advice regarding their requirements.

Please note it is the responsibility of the sponsor or the co-ordinating investigator of the project to register this study on a publicly available online registry (eg. Australian Clinical Trial Registry www.actr.org.au).

Should you have any queries about your project please contact the Research Office, Tel: 8382-2075, email research@stvincents.com.au. The HREC Terms of Reference, Standard Operating Procedures, *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice* and standard forms are available on the Research Office website: www.stvincents.com.au/researchoffice or internal: http://exwwwsvh.stvincents.com.au/researchoffice

Please quote **SVH File Number: 13/196** in all correspondence.

The HREC wishes you every success in your research.

Yours sincerely

Sarah Charlton
HREC Executive Officer
St Vincent's Research Office
Level 6 deLacy Building

cc. Beth Cheal
D/2014/14744

St Vincent's Hospital

A facility of St Vincent's
& Mater Health Sydney

St Vincent's Hospital Sydney Ltd
ABN 77 054 038 872
390 Victoria Street
Darlinghurst NSW 2010
Australia

T + 61 2 8382 1111
F + 61 2 9332 4142
www.stvincents.com.au

19 May 2014

Prof. Anita Bundy
Room J010
C42 – Cumberland Campus
University of Sydney
Lidcombe NSW 2141

Dear Anita

SVH File Number: 13/196
Project Title: DriveSafe and DriveAware iPad Application - A Valid Fitness to Drive Cognitive Screening Tool for Physicians?
HREC Reference Number: HREC/13/SVH/294

Thank you for your letter **dated 31 March 2014, received by the Research Office 12 May 2014**, submitting a request to extend HREC approval to additional sites. This HREC has been accredited by NSW Ministry of Health as a Lead HREC under the model for single ethical and scientific review and Certified by the NHMRC under the National model for Harmonisation of Multicentre Ethical Review (HoMER).This lead HREC is constituted and operates in accordance with the National Health and Medical Research Council's *National Statement on Ethical Conduct in Human Research* and the *CPMP/ICH Note for Guidance on Good Clinical Practice*. No HREC members with a conflict of interest were present for review of this project.

I am pleased to advise that the HREC Executive at a meeting on **13 May 2014** approved this request. HREC approval has been extended to the following additional **site:**

- **Monash Health, Driver Assessment & Advisory Service**
  Kingston Centre
  400 Warrigal Road
  CHELTENHAM VIC 3192

  Principal Investigator at the site: Jody Sheree White

You are reminded that this letter constitutes *ETHICAL* and *SCIENTIFIC* approval only. You must not commence this research project at a site until a completed Site Specific Assessment Form/Access Request and associated documentation have been submitted to the site Research Governance Officer and Authorised. A copy of this letter must be forwarded to all site investigators for submission to the relevant Research Governance Officer.

Should you have any queries about your project please contact the Research Office, Tel: 8382-2075, email research@stvincents.com.au. The HREC Terms of Reference, Standard Operating Procedures, *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice* and standard forms are available on the Research Office website: www.stvincents.com.au/researchoffice or internally at http://exwwwsvh.stvincents.com.au/researchoffice.

Yours sincerely

Sarah Charlton
HREC Executive Officer
Research Office
Level 6 de Lacy Building

cc. Beth Cheal

TRIM document: D/2014/24625

Continuing the Mission of the Sisters of Charity

REFERENCES

Ackerman, M. L., Crowe, M., Vance, D. E., Wadley, V. G., Owsley, C., & Ball, K. K. (2011). The impact of feedback on self-rated driving ability and driving self-regulation among older adults. *Gerontologist, 51*(3), 367-378. doi:10.1093/geront/gnq082

Adler, G., & Silverstein, N. M. (2008). At-risk drivers with Alzheimer's disease: Recognition, response, and referral. *Traffic Injury Prevention, 9*(4), 299-303. doi:10.1080/15389580801895186

Agostinelli, G., Sherman, S. J., Fazio, R. H., & Hearst, E. S. (1986). Detecting and identifying change: Additions versus deletions. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 445-454.

Al-Hassani, S. B., & Alotaibi, N. M. (2014). The impact of driving cessation on older Kuwaiti adults: Implications to occupational therapy. *Occupational Therapy In Health Care, 28*(3), 264-276. doi:10.3109/07380577.2014.917779

Allan, C., Coxon, K., Bundy, A., Peattie, L., & Keay, L. (2015). DriveSafe and DriveAware assessment tools are a measure of driving-related function and predicts self-reported restriction for older drivers. *Journal of Applied Gerontology, 35*(6), 583-600. doi:10.1177/0733464815570666

American Occupational Therapy Association. (2014). Occupational Therapy Practice Framework: Domain and Process (3rd Edition). *American Journal of Occupational Therapy, 68*(Supplement_1), S1-S48. doi:10.5014/ajot.2014.682006

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

Aparasu, R. R. (2011). *Research Methods for Pharmaceutical Practice and Policy* (1st ed.). London, UK: Pharmaceutical Press.

Asimakopulos, J., Boychuck, Z., Sondergaard, D., Poulin, V., Menard, 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

Austroads. (2004). *Austroads research report: Model licence re-assessment procedure for older drivers: Stage 2 research (No. AP-R259/04)*. Sydney: Austroads.

Austroads, & National Transport Commission. (2016). *Assessing Fitness to Drive for Commercial and Private Vehicle Drivers*. Sydney, Australia: Austroads.

Baldock, M. R., Thompson, J. P., Dutschke, J., Kloeden, C. N., Lindsay, V. L., & Woolley, J. E. (2016). *Older Road Users: Emerging Trends Research Report AP-R530-16*. Sydney, Australia: Austroads Ltd.

Ball, K., Beard, B., Roenker, D. L., Miller, R., & Griggs, D. (1988). Age and visual search: Expanding the Useful Field of View. *Journal of the Optical Society A, 5*(12), 2210-2219.

Ball, 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*, 71-79.

Barco, P. P., Baum, C. M., Ott, B. R., Ice, S., Johnson, A., Wallendorf, M., & Carr, D. B. (2015). Driving errors in persons with dementia. *Journal of American Geriatric Society, 63*(7), 1373-1380. doi:10.1111/jgs.13508

Bass, J. C., & Chiles, C. (1990). Visual skill correlation with detection of solitary pulmonary nodules. *Investigative Radiology, 25*, 994-998.

Beck, M. R., Angelone, B. L., & Levin, D. T. (2004). Knowledge about the probability of change affects change detection performance. *Journal of Experminental Psychology Human Perception and Performance, 30*(4), 778-791. doi:10.1037/0096-

Beck, M. R., & Levin, D. T. (2003). The role of representational volatility in recognizing pre- and post change objects. *Perception & Psychophysics, 65*(3), 458-468.

Beck, M. R., Levin, D. T., & Angelone, B. (2007a). Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition, 16*(1), 31-51. doi:10.1016/j.concog.2006.01.003

Beck, M. R., Levin, D. T., & Angelone, B. L. (2007b). Metacognitive errors in change detection: Lab and life converge. *Consciousness and Cognition, 16*(1), 58-62. doi:10.1016/j.concog.2006.05.003

Beck, M. R., Martin, B. A., Smitherman, E., & Gaschen, L. (2013). Eyes-on training and radiological expertise: An examination of expertise development and its effects on visual working memory. *Human Factors and Ergonomics Society, 55*, 747-763.

Becker, M. W., & Rasmussen, I. P. (2008). Guidance of attention to objects and locations by long-term memory of natural scenes. *Journal Experimental Psychology Learning Memory and Cognition, 34*(6), 1325-1338. doi:10.1037/a0013650

Bédard, M., & Dickerson, A. E. (2014). Consensus statements for screening and assessment tools. *Occupational Therapy In Health Care, 28*(2), 127-131. doi:10.3109/07380577.2014.903017

Bédard, M., Leonard, E., McAuliffe, J., Weaver, B., Gibbons, C., & Dubois, S. (2006). Visual attention and older drivers: The contribution of inhibition of return to safe driving. *Experimental Aging Research, 32*(2), 119-135. doi:10.1080/03610730500511918

Bédard, M., Weaver, B., Darzins, 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

Blackmore, S. J., Breistaff, G., Nelson, K., & Trościanko, T. (1995). Is the richness of our visual world and illusion? Transsaccadic memory for complex scenes. *Perception, 24*(9), 1075-1081.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model* (2nd ed.). Mahwah New Jersey: Lawrence Erlbaum Associates.

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 & Prevention, 59*, 537-547. doi:http://dx.doi.org/10.1016/j.aap.2013.06.037

Braekhus, A., & Engedal, K. (2009). Mental impairment and driving licences for elderly people -a survey among Norwegian general practitioners. *Scandinavian Journal of Primary Health Care, 14*(4), 223-228. doi:10.3109/02813439608997089

Brink, J. M., & McDowd, J. M. (1999). Aging and selective attention: An issue of complexity or multiple mechanisms. *Journals of Gerontology: Psychological Sciences, 54*, 30-33.

Brown, F., Cheal, B., Cooper, M., & Joshua, N. (2013). *DriveSafe DriveAware Digital Development: General Practitioner Survey Findings*. Pearson Clinical Assessment Australia & New Zealand. Sydney, Australia.

Caird, J. K., Edwards, J. D., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors, 47*(2), 235-249.

Canadian Association of Occupational Therapists. (2009). CAOT Position Statement Occupational Therapy and Driver Rehabilitation (2009). Retrieved from http://www.caot.ca/pdfs/positionstate/driving.pdf

Carr, D. B., Schwartzberg, J. G., Manning, L., & Sempek, J. (2010). *Physician's Guide to Assessing and Counseling Older Drivers, 2nd edition*. Washington, DC: NHTSA.

Carsten, O. (2007). From driver models to modelling the driver: What do we really need to know about the driver? In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments* (pp. 105–120). London: Springer-Verlag.

Charlton, J., Sjaanie, K., Odell, M., Devlin, A., Langford, J., O'Hare, M., . . . Scully, M. (2010). *Influence of Chronic Illness on Crash Involvement of Motor Vehicle Drivers* (2nd ed.). Melbourne, Australia: Monash University Accident Research Centre.

Charlton, S. G., & Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(6), 456-471. doi:10.1016/j.trf.2011.04.010

Cheal, B., Bundy, A., Patomella, A., & Scanlan, J. N. (2018). Usability testing of touchscreen DriveSafe DriveAware with older adults: A cognitive fitness to drive screen. *Cogent Medicine, 5*, 1555785. doi:10.1080/2331205X.2018.1555785

Cheal, B., Bundy, A., Patomella, A., Scanlan, J. N., & Wilson, C. (2018). Converting the DriveSafe subtest of DriveSafe DriveAware for touchscreen administration. *Australian Occupational Therapy Journal, (in press)*. doi:10.111/1440-1630.12558

Cheal, B., & Kuang, H. (2015). *DriveSafe DriveAware for Touch Screen Administration Manual.* Sydney, Australia: Pearson Australia Group Pty Ltd.

Chihuri, S., Mielenz, T. J., DiMaggio, C. J., Betz, M. E., DiGuiseppi, C., Jones, V. C., & Li, G. (2016). Driving cessation and health outcomes in older adults. *Journal of American Geriatric Society, 64*(2), 332-341. doi:10.1111/jgs.13931

Classen, S., Alvarez, L., Ferreira, P., Chen, C., Meyer, C., & Nywening, A. (2016). Developing a Canadian-Specific Version of the Fitness-to-Drive Screening Measure©. *OTJR: Occupation, Participation and Health, 36*(2), 82-91. doi:10.1177/1539449216650462

Classen, S., Velozo, C., Winter, S., Bédard, M., & Wang, Y. (2015). Psychometrics of the Fitness-to-Drive Screening Measure. *OTJR: Occupation, Participation and Health, 35*(1), 42-52. doi:10.1177/1539449214561761

Classen, S., Velozo, C., Winter, S., Brumback, B., Bédard, M., Lutz, B., . . . Rogers, J. (2014). Fitness-to-drive Screening Measure (FTDS) User Manual. Retrieved from http://fitnesstodrive.phhp.ufl.edu/ca/

Classen, S., Wang, Y., Crizzle, A., Winter, S., & Lanford, D. (2013). Predicting older driver on-road performance by means of the useful field of view and trail making test part B. *American Journal of Occupational Therapy, 67*, 574-582. doi:10.5014/ajot.2013.008136

Classen, S., Wang, Y., Velozo, C., Bédard, M., Winter, S., & Lanford, D. (2013). Concurrent criterion validity of the safe driving behavior measure: A predictor of on-road driving outcomes. *Amercan Journal of Occupational Therapy, 67*(1), 108-116.

Classen, S., Wen, P., Velozo, C., Bédard, M., Winter, S., Bumback, B., & Lanford, D. (2012). Rater reliability and rater effects of the safe driving behaviour measure. *The American Journal of Occupational Therapy, 66*, 69-77. doi:10.5014/

Classen, S., Wen, P. S., Velozo, C. A., Bedard, M., Winter, S. M., Brumback, B., & Lanford, D. N. (2012). Psychometrics of the Self-Report Safe Driving Behavior Measure for Older Adults. *American Journal of Occupational Therapy, 66*(2), 233-241. doi:10.5014/ajot.2012.001834

Classen, S., Winter, S., Velozo, C., Hannold, E., & Rogers, J. (2013). Stakeholder Recommendations to Refine the Fitness-to-Drive Screening Measure. *The Open Journal of Occupational Therapy, 1*(4). doi:10.15453/2168-6408.1054

Classen, S., Winter, S. M., Velozo, C., Bédard, M., Lanford, D., Brumback, L., & Lutz, B. (2010). Item development and validity testing for a self- and proxy report: The safe driving behavior measure. *American Journal of Occupational Therapy, 64*, 296-305.

Clay, O. J., Wadley, V., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, 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*, 724-731.

Clionsky, M., & Clionsky, E. (2014). Psychometric equivalence of a paper-based and computerized (iPad) version of the Memory Orientation Screening Test (MOST). *The Clinical Neuropsychologist, 28*(5), 747-755. doi:10.1080/13854046.2014.913686

Clough, B. A., & Casey, L. M. (2015). The smart therapist: A look to the future of smartphones and mHealth technologies in psychotherapy. *Professional Psychology: Research and Practice, 46*(3), 147-153. doi:10.1037/pro0000011

Cockburn, A., Ahlström, D., & Gutwin, C. (2012). Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies, 70*(3), 218-233. doi:10.1016/j.ijhcs.2011.11.002

Cook, D. J., Moradkhani, A., Douglas, K. S., Prinsen, S. K., Fischer, E. N., & Schroeder, D. R. (2014). Patient education self-management during surgical recovery: Combining mobile (iPad) and a content management system. *Telemedicine and e-Health, 20*(4), 312-317. doi:10.1089/tmj.2013.0219

Coxon, K., & Keay, L. (2015). Behind the wheel: Community consultation informs adaptation of safe-transport program for older drivers. *BMC Research Notes, 8*, 764. doi:10.1186/s13104-015-1745-0

Crundall, D. (2009). The deceleration detection flicker test: A measure of experience? *Ergonomics, 52*(6), 674-684. doi:10.1080/00140130802528337

Crundall, D., & Underwood, G. (1998). The effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics, 41*(4), 448-458.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception, 28*(9), 1075-1087.

Crundall, D., Van Loon, E., & Underwood, G. (2006). Attraction and distraction of attention with outdoor media. *Accident Analysis & Prevention, 38*, 671-677.

D'Ambrosio, L. A., Donorfio, L. K. M., Coughlin, J. F., Mohyde, M., & Meyer, J. (2008). Gender Differences in Self-Regulation Patterns and Attitudes Toward Driving Among Older Adults. *Journal of Women & Aging, 20*(3-4), 265-282. doi:10.1080/08952840801984758

Dalchow, J. L., Niewoehner, P. M., Henderson, R. R., & Carr, D. B. (2010). Test acceptability and confidence levels in older adults referred for fitness-to-drive evaluations. *American Journal of Occupational Therapy, 64*, 252-258.

Daly, J. M., Xu, Y., & Levy, B. T. (2015). iPad use in Iowa Research Network family physician offices. *Journal of Primary Care & Community Health, 6*(2), 137-141. doi:10.1177/2150131914558422

De Raedt, R., & Ponjaert-Kristoffersen, I. (2001). Short cognitive/neuropsychological test battery for first-tier fitness-to-drive assessment of older adults. *Clinical Neuropsychology, 15*(3), 329-336. doi:10.1076/clin.15.3.329.10277 de Winter, J. C. F., & Happee, R. (2012). Modelling driver behaviour: A rationale for multivariate statistics. *Theoretical Issues in Ergonomics Science, 13*(5), 528-545. doi:10.1080/1463922x.2010.546437

Delello, J. A., & McWhorter, R. R. (2016). Reducing the Digital Divide. *Journal of Applied Gerontology, 36*(1), 3-28. doi:10.1177/0733464815589985

Devos, H., Vandenberghe, W., Nieuwboer, A., Tant, M., Baten, G., & De Weerdt, W. (2007). Predictors of fitness to drive in people with Parkinson's disease. *Neurology, 69*(14), 1434-1441.

Di Stefano, M., & Macdonald, W. (2003). Assessment of older drivers: Relationships among on-road errors, medical conditions and test outcome. *Journal of Safety Research, 34*(4), 415-429. doi:10.1016/j.jsr.2003.09.010

Di Stefano, M., & Macdonald, W. (2010). Australian occupational therapy driver assessors' opinions on improving on-road driver assessment procedures. *American Journal of Occupational Therapy, 64*, 325-335.

Dickerson, A., E., Molnar, L. J., Bedard, M., Eby, D. W., Classen, S., & Polgar, J. (2017). Transportation and aging: An updated research agenda for advancing safe mobility. *Journal of Applied Gerontology, Oct*, 1-19. doi:10.1177/0733464817739154

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. (2014a). Driving with dementia: Evaluation, referral, and resources. *Occupational Therapy In Health Care, 28*(1), 62-76. doi:10.3109/07380577.2013.867091

Dickerson, A. E. (2014b). Screening and assessment tools for determining fitness to drive: A review of the literature for the pathways project. *Occupational Therapy In Health Care, 28*(2), 82-121. doi:10.3109/07380577.2014.904535

Dickerson, A. E., & Bédard, M. (2014). Decision tool for clients with medical issues: A framework for identifying driving risk and potential to return to driving. *Occupational Therapy In Health Care, 28*(2), 194-202. doi:10.3109/07380577.2014.903357

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*, 670-680. doi:10.5014/

Dickerson, A. E., Molnar, L. J., Eby, D. W., Adler, G., Bédard, M., Berg-Weger, M., . . . Trujillo, L. (2007). Transportation and aging: A research agenda for advancing safe mobility. *Gerontologist, 47*(5), 578-590.

Dickerson, A. E., Reistetter, T., Davis, E. S., & Monahan, M. (2011). Evaluating driving as a valued instrumental activity of daily living. *American Journal of Occupational Therapy, 65*(1), 64-75. doi:10.5014/ajot.2011.09052

Dickerson, A. E., Reistetter, T., & Gaudy, J. R. (2012). The perception of meaningfulness and performance of instrumental activities of daily living from the perspectives of the medically at-risk older adults and their caregivers. *Journal of Applied Gerontology, 32*(6), 749-764. doi:10.1177/0733464811432455

Divita, J., Obermayer, R., & Nugent, W. (2004). Verification of the change blindness phenomenon while managing critical events on a combat information display. *Human Factors, 46*(2), 205-218.

Dixon, S., Bunker, T., & Chan, D. (2007). Outcome scores collected by touchscreen: Medical audit as it should be in the 21st century? *Annals of the Royal College of Surgeons of England, 89*(7), 689-691. doi:10.1308/003588407X205422

Dobbs, A. R. (2005). *The development of a scientifically based driving assessment and standized procedure for evaluating medically at-risk drivers.* Paper presented at the Proceedings of the Canadian Multidisciplinary Road Safety Conference XV, June 5-8, Fredericton, NB.

Dobbs, A. R. (2013). Accuracy of the DriveABLE cognitive assessment to determine cognitive fitness to drive. *Canadian Family Physician, 59*, 156-161.

Dobbs, A. R., Heller, R. B., & Schopflocher, D. (1998). A comparitive approach to identify unsafe older driver. *Accident Analysis & Prevention, 30*(3), 363-370.

Donorfio, L. K., D'Ambrosio, L. A., Coughlin, J. F., & Mohyde, M. (2009). To drive or not to drive, that isn't the question - the meaning of self-regulation among older drivers. *Journal of Safety Research, 40*(3), 221-226. doi:10.1016/j.jsr.2009.04.002

Dorrington, P., Wilkinson, C., Tasker, L., & Walters, A. (2016). User-centered design method for the design of assistive switch devices to improve user experience, accessibility, and independence. *Journal of Usability Studies, 11*(2), 66-82.

Downing Peck, A. (2011). App-solutely fabulous. Hundreds of new apps for iPad and tablets make MHealth a reality and a lifestyle choice. *Medical Economics, November 25, 2011*, 11-14.

Drachman, D. A. (1988). Who may drive? Who may not? Who shall decide? *Annals of Neurology, 24*, 787–788.

Eby, D. W., & Kostyniuk, L. P. (2004). *Crashes and Driver Distraction: A Review of Databases, Crash Scenarios, and Distracted-Driving Scenarios*. Michigan, USA: Transportation Research Institute, University of Michigan.

Edwards, J. D., Lunsman, M., Perkins, M., Rebok, G. W., & Roth, D. L. (2009). Driving cessation and health trajectories in older adults. *Journal of Gerontology, Series A: Biological Sciences, 64*(12), 1290-1295. doi:10.1093/gerona/glp114

Eonta, A. M., Christon, L. M., Hourigan, S. E., Ravindran, N., Vrana, S. R., & Southam-Gerow, M. A. (2011). Using everyday technology to enhance evidence-based treatments. *Professional Psychology: Research and Practice, 42*(6), 513-520. doi:10.1037/a0025825

Esser, P., Dent, S., Jones, C., Sheridan, B. J., Bradley, A., Wade, D. T., & Dawes, H. (2016). Utility of the MOCA as a cognitive predictor for fitness to drive. *J Neurol Neurosurg Psychiatry, 87*(5), 567-568. doi:10.1136/jnnp-2015-310921

Fernandez-Duque, D., & Thornton, I. M. (2000). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition, 7*(1-3), 323-344. doi:10.1080/135062800394838

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

Findlater, L., Froehlich, J. E., Fattal, K., Wobbrock, J. O., & Dastyar, T. (2013). *Age-Related Differences in Performance with Touchscreens Compared to Traditional Mouse Input*. Paper presented at the ACM Conference on Human Factors in Computing Systems (CHI) 2013, Paris, France.

Fisher, W. P. (2007). Rating scale instrument quality criteria. *Rasch Measurement Transactions, 21*, 1095.

Fletcher-Watson, S., Leekam, S. R., Connolly, B., Collis, J. M., Findlay, J. M., McConachie, H., & Rodgers, J. (2012). Attenuation of change blindness in children with autism spectrum disorders. *British Journal of Developmental Psychology, 30*(Pt 3), 446-458. doi:10.1111/j.2044-835X.2011.02054.x

Fletcher-Watson, S., Leekam, S. R., Findlay, J. M., & Stanton, E. C. (2008). Brief report: Young adults with autism spectrum disorder show normal attention to eye-gaze information-evidence from a new change blindness paradigm. *Journal of Autism and Developmental Disorders, 38*(9), 1785-1790. doi:10.1007/s10803-008-0548-8

Fletcher-Watson, S., Leekam, S. R., Turner, M. A., & Moxon, L. (2006). Do people with autistic spectrum disorder show normal selection for attention? Evidence from change blindness. *British Journal of Psychology, 97*(4), 537-554. doi:10.1348/000712606x114057

Folk, C. L., & Hoyer, W. L. (1992). Aging and shifts of visual spatial attention. *Psychology & Aging, 7*, 453-465.

Folk, C. L., & Lincourt, A. E. (1996). The effects of age on guided conjunction search. *Experimental Aging Research, 22*, 99-118.

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*, 189-198.

Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A. (2010). *Mini-Mental Stated Examination (MMSE-2) User's Manual* (2nd ed.). Lutz, Florida: Psychological Assessment Resources.

Foster, J., Behrmann, M., & Stuss, D. (1995). Aging and visual search: Generalized cognitive slowing or selective deficit in attention? *Aging Neuropsychology and Cognition, 2*(4), 279-299.

Fox, G. K., Bowden, S. C., & Smith, D. S. (1998). On-road assessment of driving competence after brain impairment: Review of current practice and recommendations for a standardized examination. *Archives Physical Medicine & Rehabilitation, 79*(10), 1288-1296.

Fricke, J., & Unsworth, C. (2001). Time use and importance of instrumental activities of daily living. *Australian Occupational Therapy Journal, 48*(3), 118-131.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention, 37*(3), 461-472. doi:10.1016/j.aap.2004.11.003

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

Gaspar, J. G., Neider, M. B., Simons, D. J., McArely, J. S., & Kramer, A. F. (2013). Change detection: Training and transfer. *PLoS One, 8*(6), e67781. doi:10.1371/ journal.pone.0067781

Gibbons, C., Smith, N., Middleton, R., Clack, J., Weaber, B., Dubois, S., & Beard, B. (2017). Using serial trichotomization with common cognitive tests to screen for fitness to drive. . *The American Journal of Occupational Therapy, 71*, 7102260010. doi:https://doi.org/10.5014/ajot.2017.019695

Gitlow, L. (2014). Technology use by older adults and barriers to using technology. *Physical & Occupational Therapy In Geriatrics, 32*(3), 271-280. doi:10.3109/02703181.2014.946640

Glaser, D., Jain, S., & Kortum, P. (2013). Benefits of a physician-facing tablet presentation of patient symptom data: Comparing paper and electronic formats. *BMC Medical Informatics & Decision Making, 13*(99), 1-8.

Glennan, S. (2005). Modeling mechanisms. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 443-464. doi:10.1016/j.shpsc.2005.03.011

Global Burden of Disease Study 2013 Collaborators. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the

Global Burden of Disease Study 2013. *The Lancet, 386*(9995), 743-800. doi:10.1016/s0140-6736(15)60692-4

Godwin, Z., Tan, J., Bockhold, J., Ma, J., & Tran, N. K. (2015). Development and evaluation of a novel smart device-based application for burn assessment and management. *Burns, 41*(4), 754-760. doi:10.1016/j.burns.2014.10.006

Gordon, N. P., & Hornbrook, M. C. (2016). Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: A database and survey study of seniors in a large health plan. *Journal of Medical Internet Research, 18*(3), e50. doi:10.2196/jmir.5105

Grandgenevre, P., Vaiva, G., Boloix, E., Bubrovszky, M., Schwan, R., & Laprevote, B. (2015). Dissociation of explicit and implicit responses during a change blindness task in schizophrenia. *Neuropsychologia, 71*, 11-17.

Hargrave, D. D., Nupp, J. M., & Erickson, R. J. (2012). Two brief measures of executive function in the prediction of driving ability after acquired brain injury. *Neuropsychological Rehabilitation, 22*(4), 489-500. doi:10.1080/09602011.2012.662333

Hasher, L., & Zacks, R. (1998). Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning and Motivation* (Vol. 22, pp. 193-225). San Diego, C A: Academic Press.

Hayhoe, M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38*, 125-137.

Hegde, V. (2013). *Role of human factors / usability engineering in medical device design.* Paper presented at the Proceedings Annual Reliability and Maintainability Symposium (RMAS), Orlando, FL.

Heikoop, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2015). Psychological constructs in driving automation: a consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science, 17*(3), 284-303. doi:10.1080/1463922x.2015.1101507

Heitger, M. H., Anderson, T. J., Jones, R. D., Dalrymple-Alford, J. C., Frampton, C. M., & Ardagh, M. W. (2004). Eye movement and visuomotor arm movement deficits following mild closed head injury. *Brain, 127*(3), 575-590. doi:10.1093/brain/awh066

Hines, A., & Bundy, A. (2014). Predicting driving ability using DriveSafe and DriveAware in people with cognitive impairments: A replication study. *Australian Occupational Therapy Journal, 61*(4), 224-229. doi:10.1111/1440-1630.12112

Hoffman, L., McDowd, J. M., Atchley, P., & 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., 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

Hoggarth, P. (2013). Diagnosis of cognitive impairment and the assessment of driving safety in New Zealand: A survey of Canterbury GPs. *The New Zealand Medical Journal, 126*(1387), 87-97.

Hoggarth, P., Innes, C., Dalrymple-Alford, J., & Jones, R. (2013). Predicting on-road assessment pass and fail outcomes in older drivers with cognitive impairment using a battery of computerized sensory-motor and cognitive tests. *Journal of the American Geriatrics Society, 61*(12), 2192-2198.

Hollis, A. M., Duncanson, H., Kapust, L., Xi, P. M., & O'Connor, M. G. (2015). Validity of the Mini-Mental State Examination and the Montreal Cognitive Assessment in the prediction of driving test outcome. *Journal of American Geriatric Society, 63*(5), 988-992.

Howell, M., Hood, A. J., & Jayne, D. G. (2015). Use of a patient completed iPad questionnaire to improve pre-operative assessment. *Journal of Clinical Monitoring and Computing*. doi:10.1007/s10877-015-9818-0

Hughes, P. K., & Cole, B. L. (1986). What attracts attention when driving? *Ergonomics, 29*(3), 377-391. doi:10.1080/00140138608968272

Hussain, A., & Kutar, M. (2009). *Usability metric framework for mobile phone application*. Paper presented at the The 10th Annual PostGraduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting, Liverpool, UK.

Innes, C., Jones, R., Dalrymple-Alford, J., Hayes, S., Hollobon, S., Severinsen, J., . . . Anderson, T. (2007). Sensory-motor and cognitive tests predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences, 260*(188-198).

Innes, C. R., Jones, R. D., Anderson, T. J., Hollobon, S. G., & Dalrymple-Alford, J. C. (2009). Performance in normal subjects on a novel battery of driving-related sensory-motor and cognitive tests. *Behavior Research Methods, 41*(2), 284-294. doi:10.3758/BRM.41.2.284

International Organization for Standardization. (2010). ISO 9241-210: 2010 Ergonomics of Human System Interaction - Part 210: Human-centered Design for Interactive Systems. Genève, Switzerland: International Organization for Standardization.

Jang, R. W., Man-Son-Hing, M., Molnar, F. J., Hogan, D. B., Marshall, S. C., Auger, J., . . . Naglie, G. (2007). Family physicians' attitudes and practices regarding assessments of medical fitness to drive in older persons. *Journal of General Internal Medicine, 22*(4), 531-543. doi:10.1007/s11606-006-0043-x

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(5), 529-546. doi:10.1002/wcs.130

Jones, B. T., Jones, B. C., Smith, H., & Copley, N. (2002). Memory for centrally attended changing objects in an incidental real-world detection paradigm. *British Journal of Psychology, 93*, 289-302.

Jones, K., Rouse-Watson, S., Beveridge, A., Sims, J., & Schattner, A. (2012). GP perspectives of assessing older and functionally impaired patients. *Australian Family Physician, 41*(4), 235-239.

Justiss, M. D., Mann, W. C., Stav, W., & Velozo, C. (2006). Development of a behind- the-wheel driving performance assessment for older adults. *Topics in Geriatric Rehabilitation, 22*(2), 121-128.

Kahvedzic, A., McFadden, R., Cummins, G., Carr, D., & O'Neill, D. (2015). Impact of new guidelines and educational program on awareness of medical fitness to drive among general practitioners in Ireland. *Traffic Injury Prevention, 16*(6), 593-598. doi:10.1080/15389588.2014.979408

Kay, L. (2008). *Predicting fitness to drive in people with cognitive impairments: An occupational therapy perspective.* (Doctor of Philosophy), University of Sydney.

Kay, L., & Bundy, A. (2009). *DriveSafe and DriveAware an Assessment Tool for Driver Rehabilitation*. Sydney, Australia: PsychCorp.

Kay, L., Bundy, A., & Clemson, L. (2003). Awareness of driving ability in senior drivers with neurological conditions. *American Journal of Occupational Therapy, 63*, 146-150.

Kay, L., Bundy, A., & Clemson, L. (2008). Predicting fitness to drive using the Visual Recognition Slide Test (USyd). *American Journal of Occupational Therapy, 62*, 187-197.

Kay, L., Bundy, A., & Clemson, L. (2009a). Predicting fitness to drive in people with cognitive impairments by using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation, 90*(9), 1514-1522. doi:10.1016/j.apmr.2009.03.011

Kay, L., Bundy, A., & Clemson, L. (2009b). Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation, 31*(13), 1074-1082. doi:10.1080/09638280802509553

Kay, L., Bundy, A., Clemson, L., Cheal, B., & Glendenning, T. (2012). Contribution of off-road tests to predicting on-road performance: A critical review of tests. *Australian Occupational Therapy Journal, 59*(1), 89-97. doi:10.1111/j.1440-1630.2011.00989.x

Kay, L., Bundy, A., Clemson, L., & Jolly, N. (2008). Validity and reliability of the on-road driving assessment with senior drivers. *Accident Analysis & Prevention, 40*(2), 751-759. doi:10.1016/j.aap.2007.09.012

Keskinen, E. (2014). Education for older drivers in the future. *IATSS Research, 38*(1), 14-21. doi:10.1016/j.iatssr.2014.03.003

Keskinen, E., Peräaho, M., Laapotti, S., Hernetkoski, K., & Katila, A. (2010, 22-24.8.2010). *The fifth level to the GDE-model—5soc and 5pro.* Paper presented at the 3rd Joined NORBIT and the 5th Japanese–Nordic Conference in Traffic and Transportation Psychology, University of Turku, Finland.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data*. (Technical Report No. DOT HS 810 594). Springfield, Virginia: National Highway Traffic Safety Administration.

Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., & Ifukube, T. (2011). *Elderly User Evaluation of Mobile Touchscreen Interactions.* Paper presented at the INTERACT 2011, Lisbon, Portugal.

Kontos, E., Blake, K. D., Chou, W. Y., & Prestin, A. (2014). Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. *Journal of Medical Internet Research, 16*(7), e172. doi:10.2196/jmir.3117

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., & Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *Amercan Journal of Occupational Therapy, 60*, 428-434.

Korner-Bitensky, N., & Sofer, S. (2009). The DriveABLE Competence Screen as a predictor of on-road driving in a clinical sample. *Australian Occupational Therapy Journal, 56*(3), 200-205. doi:10.1111/j.1440-1630.2008.00749.x

Korner-Bitensky, N., Sofer, S., Gelinas, I., & Mazer, B. (1998). Evaluating driving potential in persons with stroke: A survey of occupational therapy practices. *American Journal of Occupational Therapy, Volume 52*(10), 916-919.

Korner-Bitensky, N., Toal-Sullivan, D., & von Zweck, C. (2007a). Driving and older adults: A focus on assessment by occupational therapists. *Occupational Therapy Now, 9*(5), 10-12.

Korner-Bitensky, N., Toal-Sullivan, D., & von Zweck, C. (2007b). Driving and older adults: Towards a national occupational therapy strategy for screening. *Occupational Therapy Now, 9*(4), 3-5.

Kostyniuk, L. P., & Shope, J. T. (2003). Driving and alternatives: Older drivers in Michigan. *Journal of Safety Research, 34*(4), 407-414. doi:10.1016/j.jsr.2003.09.001

Koustanaï, A., Boloix, E., Van Elslande, P., & Bastien, C. (2008). Statistical analysis of "looked-but-failed-to-see" accidents: Highlighting the involvement of two distinct mechanisms. *Accident Analysis & Prevention, 40*(2), 461-469. doi:10.1016/j.aap.2007.08.001

Koustanaï, A., Van Elslande, P., & Bastien, C. (2012). Use of change blindness to measure different abilities to detect relevant changes in natural driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 15*(3), 233-242. doi:10.1016/j.trf.2011.12.012

Kunemund, H., & Tanschus, N. M. (2014). The technology acceptance puzzle. Results of a representative survey in Lower Saxony. *Zeitschrift für Gerontologie und Geriatrie, 47*(8), 641-647. doi:10.1007/s00391-014-0830-7

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., Braitman, K., Charlton, J., Eberhard, J., O'Neill, D., Staplin, L., & Stutts, J. (2008). TRB Workshop 2007: Licensing authorities' options for managing older driver safety--practical advice from the researchers. *Traffic Injury Prevention, 9*(4), 278-281. doi:10.1080/15389580801895210

Lavi, I., Malki, G., & Kornowski, R. (2014). Medical iPad use in the cardiac catheterization and echo laboratories. *International Journal of Cardiovascular Imaging, 30*(2), 245-252. doi:10.1007/s10554-013-0352-1

Laycock, K. M. (2011). Driver assessment: Uncertainties inherent in current methods. *British Columbia Medical Journal, 53*(2), 74-78.

Leng, F. Y., Yeo, D., George, S., & Barr, C. (2014). Comparison of iPad applications with traditional activities using person-centred care approach: Impact on well-being for persons with dementia. *Dementia, 13*(2), 265-273. doi:10.1177/1471301213494514

Levasseur, M., Audet, T., Gélinas, I., Bédard, M., Langlais, M., Therrien, F., . . . D'Amours, M. (2014). Development and validation of an awareness tool for safe and responsible driving (OSCAR). *Journal of Scientific Research & Reports, 3*(18), 2422-2433.

Levin, D. T., Simons, D. J., Angelone, B. L., & Chabris, C. F. (2002). Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology, 93*(3), 289-302.

Lew, H. L., Poole, J. H., Lee, E. H., Jaffe, D. L., Huang, H. C., & Brodd, E. (2005). Predictive validity of driving simulator assessment following traumatic brain injury: A preliminary study. *Brain Injury, 19*, 177-188.

Linacre, J. M. (2003). Data Variance: Explained, Modeled and Empirical. *Rasch Measruement Transactions.* Retrieved from http://www.rasch.org/rmt/rmt173g.htm

Linacre, J. M. (2014). *Winsteps Rasch Measurement Computer Program User's Guide*. Beaverton, Oregon: Winsteps.com.

Lincoln, N. B., & Fanthome, Y. (1994). Reliability of the Stroke Drivers Screening Assessment. *Clinical Rehabilitation, 8*, 157-160.

Lincoln, N. B., Radford, K. A., & Nouri, F. M. (2012). *Stroke Drivers' Screening Assessment Revised Manual 2012*. Retrieved from Nottingham, UK:

Loth, E., Carlos Gomez, J., & Happe, F. (2008). Detecting changes in naturalistic scenes: Contextual inconsistency does not influence spontaneous attention in high-functioning people with autism spectrum disorder. *Autism Research, 1*(3), 179-188. doi:10.1002/aur.19

Loureiro, B., & Rodrigues, R. (2014). *Design guidelines and design recommendations of multi-touch interfaces for elders*. Paper presented at the ACHI 2014: The Seventh International Conference on Advances in Computer-Human Interactions, Barcelona, Spain.

Lundberg, C., Caneman, G., Samuelsson, S. M., Hakamies-Blomqvist, L., & Almqvist, O. (2003). The assessment of fitness to drive after stroke: The Nordic stroke driver screening assessment. *Scandinavian Journal of Psychology, 44*, 23-30.

Lundqvist, A., Gerdle, B., & Ronnberg, J. (2000). Neuropsychological aspects of driving ability among older drivers. *Applied Cognitive Psychology, 14*, 135-148.

Lutzenberger, M., & Albayrak, S. (2013). *Can you simulate traffic psychology? An analysis.* Paper presented at the Winter Simulation Conference, Washington, DC.

Lützenburger, M., & Albayrak, S. (2013). *Can you simulate traffic psychology? An analysis.* Paper presented at the 2013 Winter Simulation Conference, Washingon, DC.

Maccari, L., Casagrande, M., Martella, D., Anolfo, M., Rosa, C., Fuentes, L. J., & Pasini, A. (2013). Change blindness in children with ADHD: A selective impairment in visual search? *Journal of Attention Disorders, 17*(7), 620-627. doi:10.1177/1087054711433294

Macknik, S., Fisher, B., & Bridgeman, B. (1991). Flicker distorts visual space constancy *Vision Research, 31*, 2057-2064.

Marottoli, R., & Richardson, E. (1998). Confidence in, and self-rating of, driving ability among older drivers. *Accident Analysis & Prevention, 30*(3), 331-336.

Marrington, S. A., Horswill, M. S., & Wood, J. M. (2008). The effect of simulated cataracts on drivers' hazard perception ability. *Optometry and Vision Science, 85*(12), 1121–1127.

Marshall, S. C. (2008). The role of reduced fitness to drive due to medical impairments in explaining crashes involving older drivers. *Traffic Injury Prevention, 9*(4), 291-298. doi:10.1080/15389580801895244

Marshall, S. C., Demmings, E. M., Woolnough, A., Salim, D., & Man-Son-Hing, M. (2012). Determining fitness to drive in older persons: A survey of medical and surgical specialists. *Canadian Geriatrics Journal, 15*(4), 101-119. doi:10.5770/cgj.15.30

Marshall, S. C., & Man-Son-Hing, M. (2011). Multiple chronic medical conditions and associated driving risk: A systematic review. *Traffic Injury Prevention, 12*(2), 142-148. doi:10.1080/15389588.2010.551225

Marshall, S. C., Molnar, F. J., Man-Son-Hing, M., Blair, R., Brosseau, L., Finestone, H., . . . Wilson, K. G. (2007). Predictors of driving ability following stroke: A systematic review. *Topics in Stroke Rehabilitation, 14*(1), 98-114.

Martens, M. H., & Fox, M. R. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(6), 476-492. doi:10.1016/j.trf.2007.05.003

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

Matthew, A. G., Currie, K. L., Irvine, J., Ritvo, P., Santa Mina, D., Jamnicky, L., . . . Trachtenberg, J. (2007). Serial personal digital assistant data capture of health-related quality of life: A randomized controlled trial in a prostate cancer clinic. *Health and Quality of Life Outcomes, 5*(1), 38. doi:10.1186/1477-7525-5-38

McCarley, J. S., Vais, M., J., Pringe, H., Kramer, A. F., Irwin, D., & Strayer, D. L. (2004). Conversation disrupts change detection in complex traffic scenes. *Human Factors, 46*(3), 424-436.

McCarthy, J. (2015). How Many Health Apps Actually Matter? *Empowering Health.* Retrieved from http://www.healthcareitnews.com/news/how-many-health-apps-actually-matter

McCurdie, T., Taneva, S., Casselman, M., Yeung, M., McDaniel, C., Ho, W., & Cafazzo, J. (2012). mHealth consumer apps the case for user-centered design. *Horizons, Fall 2012*, 49-54.

Medhizadah, S., Classen, S., & Johnson, A. (2018). Conducting the 32-item fitness-to-drive screening measure. *OTJR: Occupation, Participation and Health, 38*(2), 89-95.

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), *Human Behaviour and Traffic Safety* (pp. 485-520). New York: Plenum Press.

Miglioretti, D. L., Gard, C. C., Carney, P. A., Onega, T. L., Buist, D. S., Sickles, E. A., . . . Elmore, J. G. (2009). When radiologists perform best: The learning curve in screening mammogram interpretation. *Radiology, 253*(3), 632-640.

Molnar, F. J., Byszewski, A. M., Marshall, S. C., & Man-Son-Hing, M. (2005). In-office evaluation of medical fitness to drive Practical approaches for assessing older people. *Canadian Family Physician, 51*, 372-379.

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 American Geriatric Society, 54*(12), 1809-1824. doi:10.1111/j.1532-5415.2006.00967.x

Molnar, L. J., & Eby, D. W. (2008). The relationship between self-regulation and driving-related abilities in older drivers: An exploratory study. *Traffic Injury Prevention, 9*(4), 314-319. doi:10.1080/15389580801895319

Morgan, C. M., Winder, S. M., Classen, S., McCarthy, D., & Awadzi, K. D. (2009). Literature review on older adult gender difference for driving self-regulation and cessation. *Topics in Geriatric Rehabilitation, 25*(2), 99-117.

Mourant, R. R., & Rockwell, R. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors, 14*, 325-335.

Moussa, Y., Mahdanian, A. A., Yu, C., Segal, M., Looper, K. J., Vahia, I. V., & Rej, S. (2017). Mobile health technology in late-life mental illness: A focused literature review. *American Journal of Geriatric Psychiatry, 25*(8), 865-872. doi:10.1016/j.jagp.2017.04.003

Myers, A., Ball, K., Kalina, T., Roth, D. L., & Goode, K. (2000). Relation of Useful Field of View and other screening tests to on-road driving performance. *Perceptual and Motor Skills, 91*, 279-290.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, A., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699.

National Health Service. (2011). Digital First The Delivery Choice for England's Population. Retrieved from http://transformuk.com/wp-content/uploads/2013/03/DigitalbyDefault_v20120626.pdf

Nayebi, F., Desharnais, J., & Abran, A. (2012). *The state of the art of mobile application usability evaluation.* Paper presented at the 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Montreal, Quebec.

Nielsen, & IAB Australia. (2015). The Mobile Story Nielsen Mobile Rating. Retrieved from https://www.iabaustralia.com.au/research-and-resources/research-resources/item/12-research-and-resource/2013-3rd-mobile-ratings-report-september-2015

Nouri, F. (1988). Fitness to drive and the general practitioner. *International Disability Studies, 10*(3), 101-103. doi:10.3109/09638288809164122

Nouri, F. M., & Lincoln, N. B. (1992). Validation of a cognitive assessment predicting driving performance after stroke. *Clinical Rehabilitation, 6*, 275-281.

O'Donnell, J. M., Morgan, M. K., & Manuguerra, M. (2018). Patient functional outcomes and quality of life after microsurgical clipping for unruptured intracranial aneurysm. A prospective cohort study. *Journal of Neurosurgery, Published online March 2, 2018*. doi:10.3171/2017.8.JNS171576

O'Regan, J., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7*(1-3), 191-211. doi:10.1080/135062800394766

O'Regan, J. K., Rensink, R. A., & Clark, J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398*, 34.

Occupational Therapy Australia. (2015). *Position Paper: Occupational Therapy in Driver Assessment and Rehabilitation*. Victoria, Australia: Occupational Therapy Australia.

Omer, S., Dolan, C., Dimitrov, B. D., Langan, C., & McCarthy, G. (2014). General practitioners' opinions and attitudes towards medical assessment of fitness to drive of older adults in Ireland. *Australasian Journal on Ageing, 33*(3), E33-36. doi:10.1111/ajag.12045

Onoda, K., Hamano, T., Nabika, Y., Aoyama, A., Takayoshi, H., Nakagawa, T., . . . Yamaguchi, S. (2013). Validation of a new mass screening tool for cognitive impairment: Cognitive Assessment for Dementia, iPad version. *Clinical Interventions in Aging, 8*, 353-360. doi:10.2147/CIA.S42342

Østensen, E., Gjevjon, E., Øderud, T., & Moen, A. (2017). Introducing technology for thriving in residential long-term Care. *Journal of Nursing Scholarship, 49*(1), 44-53. doi:10.1111/jnu.12268

Papa, M., Boccardi, V., Prestano, R., Angellotti, E., Desiderio, M., Marano, L., . . . Paolisso, G. (2014). Comorbidities and crash involvement among younger and older drivers. *PLoS One, 9*(4), e94564. doi:10.1371/journal.pone.0094564

Patomella, A., & Bundy, A. (2015). P-Drive: Implementing an assessment of on-road driving in clinical settings and investigating Its internal and predictive validity. *American Journal of Occupational Therapy, 69*(4), 6904290010p6904290011-6904290018. doi:10.5014/ajot.2015.015131

Patomella, A., Tham, K., Johansson, K., & Kottorp, A. (2010). P-Drive on-road: Internal scale validity and reliability of an assessment of on-road driving performance in people with neurological disorders. *Scandinavian Journal of Occupational Therapy, 17*(1), 86-93. doi:10.3109/11038120903071776

Patomella, A., Tham, K., & Kottorp, A. (2006). P-Drive: Assessment of driving performance after stroke. *Journal of Rehabilitation Medicine 2006, 38*, 273-279. doi:10.1080/16501970600632594

Pearson, P., & Schaefer, E. (2005). Toupee or not toupee? The role of instructional set, centrality, and relevance in change blindness. *Visual Cognition, 12*(8), 1528-1543. doi:10.1080/13506280444000814

Peraaho, M., Keskinen, E., & Hatakka, M. (2003). *Driver competence in a heirarchical perspective; Implications for driver education*. Turku, Finland: University of Turku,.

Pfaffli, M., Thali, M. J., & Eggert, S. (2012). Physicians' knowledge and continuing medical education regarding fitness to drive: A questionnaire-based survey in Southeast Switzerland. *International Journal of Legal Medicine, 126*(3), 357-362. doi:10.1007/s00414-011-0637-7

Piersma, D., Fuermaier, A. B., de Waard, D., Davidse, R. J., de Groot, J., Doumen, M. J., . . . Tucha, O. (2016). Prediction of fitness to drive in patients with Alzheimer's dementia. *PLoS One, 11*(2), e0149566. doi:10.1371/journal.pone.0149566

Preusser, D. F., Williams, A. F., Ferguson, S. A., Ulmer, R. G., & Weinstein, H. B. (1998). Fatal crash risk for older drivers at intersections. *Accident Analysis & Prevention, 30*(2), 151-159.

Price Waterhouse Coopers, & Department of Health. (2013). *A Review of the Potential Benefits from the Better Use of Information and Technology in Health and Social Care Final Report*. Riverside, London: PricewaterhouseCoopers LLP.

Pringle, H., Irwin, D., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin & Review, 8*(1), 89-95.

Rabbitt, P. M. A. (1981). Cognitive psychology needs models for changes in performance with old age. In J. B. Long & A. D. Baddeley (Eds.), *Attention and Performance IX*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rakotonirainy, A., Steinhardt, D., Delhomme, P., Darvell, M., & Schramm, A. (2012). Older drivers' crashes in Queensland, Australia. *Accident Analysis & Prevention, 48*, 423-429. doi:10.1016/j.aap.2012.02.016

Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention, 26*(6), 733-750.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. (2001). Visual span in expert chess players: evidence from eye movements. *Psychological Science, 12*, 48-55.

Rensink, R. A. (2000). The Dynamic Representation of Scenes. *Visual Cognition, 7*(1-3), 17-42. doi:10.1080/135062800394667

Rensink, R. A. (2002). Change detection. *Annual Review Psychology, 53*, 245-277.

Rensink, R. A. (2005). Change Blindness. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 76-81). San Diego, CA: Elsevier.

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.

Richard, C. M., Wright, R. D., Ee, C., & Prime, S. L. (2002). Effect of a concurrent auditory task on visual search performance in a driving-related image-flicker task. *Human Factors, 44*(1), 108-119.

Rizzo, M., Sparks, J., McEvoy, S., Viamonte, S., Kellison, I., & Vecera, S. P. (2009). Change blindness, aging, and cognition. *Journal of Clinical and Experimental Neuropsychology, 31*(2), 245-256. doi:10.1080/13803390802279668

Rogers, W. A., & Mitzner, T. L. (2016). Envisioning the future for older adults: Autonomy, health, well-being, and social connectedness with technology support. *Futures*. doi:10.1016/j.futures.2016.07.002

Ryan, J. M., Corry, J. R., Attewell, R., & Smithson, M. J. (2002). A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper. *Quality of Life Research, 11*, 19-26.

Scailfa, C. T., & Joffe, K. A. (1997). Age differences in features and conjunction search: Implications for theories of visual search and generalized slowing. *Aging, Neuropsychology, and Cognition, 7*(1-3), 277-246.

Schmidt-Daffy, M. (2013). Fear and anxiety while driving: Differential impact of task demands, speed and motivation. *Transportation Research Part F: Traffic Psychology and Behaviour, 16*, 14-28. doi:10.1016/j.trf.2012.07.002

Schneider, N., Wilkes, J., Grandt, M., & Schlick, C. (2008). *Investigation of Input Devices for the Age-differentiated Design of Human-Computer Interaction.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting - 2008, New York.

Schömig, N., Metz, B., & Krüger, H. (2011). Anticipatory and control processes in the interaction with secondary tasks while driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(6), 525-538. doi:10.1016/j.trf.2011.06.006

Selander, H., Johansson, K., Lundberg, C., & Falkmer, T. (2010). The Nordic Stroke Driver Screening Assessment as predictor for the outcome of an on-road test. *Scandinavian Journal of Occupational Therapy, 17*(1), 10-17. doi:10.3109/11038120802714898

Selander, H., Lee, H. C., Johansson, K., & Falkmer, T. (2011). Older drivers: On-road and off-road test results. *Accident Analysis & Prevention, 43*(4), 1348-1354. doi:10.1016/j.aap.2011.02.007

Shechtman, O., Awadzi, K. D., Classen, S., Landford, D. N., & Yongsung, J. (2010). Validity and critical driving errors of on-road assessment for older drivers. *American Journal of Occupational Therapy, 64*, 242-251.

Shinar, D., & Oppenheim, I. (2011). Review of models of driver behaviour and development of a unified driver behaviour model for driving in safety critical situations. In C. Cacciabue, M. Hjälmdahl, A. Luedtke, & C. Riccioli (Eds.), *Human Modelling in Assisted Transportation* (pp. 215-223). Italia srl: Springer-Verlag.

Shinoda, H., Hayhoe, M., & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research, 41*, 3535–3545.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*, 261-267.

Sims, J., Rouse-Watson, P., Schattner, A., Beveridge, A., & Jones, K. (2012). To drive or not to drive: Assessment dilemmas for GPs. *International Journal of Family Medicine, 2012*, 1-6. doi:10.1155/2012/417512

Smilek, D., Eastwood, J. D., Reynolds, M. G., & Kingstone, A. (2007). Metacognitive errors in change detection: Missing the gap between lab and life. *Conscious Cognition, 16*(1), 52-57. doi:10.1016/j.concog.2006.04.001

Smilek, D., Eastwood, J. D., Reynolds, M. G., & Kingstone, A. (2008). Metacognition and change detection: Do lab and life really converge? *Conscious Cognition, 17*(3), 1056-1061. doi:10.1016/j.concog.2007.08.003

Smith, H., & Milne, E. (2009). Reduced change blindness suggests enhanced attention to detail in individuals with autism. *Journal of Child Psychology and Psychiatry, 50*(3), 300-306. doi:10.1111/j.1469-7610.2008.01957.x

Statista. (2017). Global Market Share Held by Tablet Vendors from 2nd Quarter 2011 to 3rd Quarter 2016. Retrieved from https://www.statista.com/statistics/276635/market-share-held-by-tablet-vendors/

Summala, H., Lamble, D., & Laakso, M. (1998). Driving experience and perception of the lead car's braking when looking at in-car targets. *Accident Analysis & Prevention, 30*(4), 401-407.

Summala, H., Nieminen, T., & Punto, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors, 38*(3), 422-451.

Summerfield, J., Lepsien, J., Gitelman, D., Mesulam, M., & Nobre, A. (2006). Orienting attention based on long-term memory experience. *Neuron, 49*, 905-916.

Summers, G. F. (1970). Introduction. In G. F. Summers (Ed.), *Attitude Measurement* (pp. 1-21). Chicago, IL: Rand McNally.

Tennant, A., & Pallant, J. F. (2006). Unidimensionality matters! (A tale of two smiths?). *Rasch Measruement Transactions, 20*, 1048-1051.

Transportation Research Board. (2016). *Transportation Research Circular. Taxonomy and Terms for Stakeholders in Senior Mobility (Number E-C211)*. Retrieved from http://onlinepubs.trb.org/Onlinepubs/circulars/ec211.pdf

Trzepacz, P. T., Hochstetler, H., Wang, S., Walker, B., Saykin, A. J., & Alzheimer's Disease Neuroimaging, I. (2015). Relationship between the Montreal Cognitive Assessment and Mini-mental State Examination for assessment of mild cognitive impairment in older adults. *BMC Geriatrics, 15*(107). doi:10.1186/s12877-015-0103-3

Tsai, H., Shillair, R., Cotten, S. R., Winstead, V., & Yost, E. (2015). Getting grandma online: Are tablets the answer for increasing digital inclusion for older adults. *Educational Gerontology, 41*(10), 695–709. doi:10.1080/03601277.2015.1048165]

Turkan, B. N., Amado, S., Ercan, E. S., & Percinel, I. (2016). Comparison of change detection performance and visual search patterns among children with/without ADHD: Evidence from eye movements. *Research in Development & Disability, 49-50*, 205-215.

Turney, B. W., & Reynard, J. M. (2014). Obtaining patient feedback in an outpatient lithotripsy service is facilitated by use of a touch-screen tablet (iPad) survey. *Urolithiasis, 42*(4), 317-321. doi:10.1007/s00240-014-0662-3

Upton, D., Upton, P., Jones, T., Jutlla, K., & Brooker, D. (2011). *Evaluation of the Impact of the Use of Touchscreen Technology with People with Dementia*. Retrieved from Worcester, UK: http://79.170.44.96/lifestorynetwork.org.uk/wp-content/uploads/downloads/2012/11/evaluation-of-the-impact-of-the-use-of-touchscreen-technology-with-people-with-dementia-.pdf

Veirk, E., & Kiesel, A. (2008). Change detection: Evidence for information accumulation in flicker pardigms. *Acta Psychologica, 127*, 309-323.

Velichkovsky, B. M., Dornhoefer, S. M., Kopf, M., Helmert, J., & Joos, M. (2002). Change detection and occlusion modes in road-traffic scenarios. *Transportation Research Part F, 5*, 99–109.

Vrkljan, B. H., McGrath, C., & Letts, L. (2011). Assessment tools for evaluating fitness to drive: A critical appraisal of evidence. *Canadian Journal of Occupational Therapy, 78*(2), 80-96. doi:10.2182/cjot.2011.78.2.3

Vrkljan, B. H., Myers, A., Crizzle, A., Banchard, R., & Marshall, S. C. (2013). Evaluating medically at-risk drivers: A survey of assessment practices in Canada. *Canadian Journal of Occupational Therapy, 80*(5), 295-303. doi:DOI: 10.1177/0008417413511788

Wallis, G., & Bulthoff, H. (2000). What's scene and not seen: Influences of movement and task upon what we see. *Visual Cognition, 7*(1-3), 175-190. doi:10.1080/135062800394757

Watkins, I., & Xie, B. (2015). Older adults' perceptions of using iPads for improving fruit and vegetable intake: An exploratory study. *Care Management Journals, 16*(1), 2-13. doi:10.1891/1521-0987.16.1.2

Weaver, B., & Bédard, M. (2012). Assessing fitness-to-drive: Practical tips on choosing the right screening tools for your practice. *Canadian Geriatrics Society Journal of CME, 2*, 5-9.

Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novice comparison of change detection in football images. *Visual Cognition (Special Issue: Change Blindness and Visual Memory), 2000*(7), 163-173.

Wetton, M. A., Horswill, M. S., Hatherly, C., Wood, J. M., Pachana, N. A., & Anstey, K. J. (2010). The development and validation of two complementary measures of drivers' hazard perception ability. *Accident Analysis & Prevention, 42*(4), 1232-1239. doi:10.1016/j.aap.2010.01.017

Wheatley, C. J., & Di Stefano, M. (2008). Individualized assessment of driving fitness for older individuals with health, disability, and age-related concerns. *Traffic Injury Prevention, 9*(4), 320-327. doi:10.1080/15389580801895269

White, C. B., & Caird, J. K. (2010). The blind date: The effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident Analysis & Prevention, 42*(6), 1822-1830. doi:10.1016/j.aap.2010.05.003

White, J. H., Miller, B., Magin, P., Attia, J., Sturm, J., & Pollack, M. (2012). Access and participation in the community: A prospective qualitative study of driving post-stroke. *Disability Rehabilitation, 34*(10), 831-838. doi:10.3109/09638288.2011.623754

Wiarda, N. R., McMinn, M. R., Peterson, M. A., & Gregor, J. A. (2014). Use of technology for note taking and therapeutic alliance. *Psychotherapy, 51*(3), 443-446. doi:10.1037/a0035075

Wilson, A. S., Kitas, G. S., Carruthers, D. M., Reay, C., Skan, J., & Harris, S. (2002). Computerized information-gathering in specialist rheumatology clinics: An initial evaluation of an electronic version of the Short Form 36. *Rheumatology, 41*, 268-273.

Wilson, L. R., & Kirby, N. H. (2008). Individual differences in South Australian general practitioners' knowledge, procedures and opinions of the assessment of older drivers. *Australasian Journal on Ageing, 27*(3), 121-125. doi:10.1111/j.1741-6612.2008.00304.x

Wood, J. M., Anstey, K. J., Lacherez, P. F., Kerr, G. K., Mallon, K., & Lord, S. R. (2009). The on-road difficulties of older drivers and their relationship with self-reported motor vehicle crashes. *Journal of American Geriatric Society, 57*(11), 2062-2069. doi:10.1111/j.1532-5415.2009.02498.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 & Prevention, 50*, 1161-1168. doi:10.1016/j.aap.2012.09.009

Woolnough, A., Salim, D., Marshall, S. C., Weegar, K., Porter, M. M., Rapoport, M. J., . . . Vrkljan, B. (2013). Determining the validity of the AMA guide: A historical cohort analysis of the assessment of driving related skills and crash rate among older drivers. *Accident Analysis & Prevention, 61*, 311-316. doi:10.1016/j.aap.2013.03.020

Wright, M., Green, A., & Baker, S. (2000). Limitations for change detection in multiple Gabor targets. *Visual Cognition, 7*, 237-252.

Yack, C., & Camic, P. M. (2017). Touchscreen interventions and the well-being of people with dementia and caregivers: A systematic review. *International Psychogeriatrics, 29*(8), 1261-1280.

Yale, S., Hansotia, P., Knapp, D., & Ehrfurth, J. (2003). Neurologic conditions: Assessing medical fitness to drive. *Clinical Medicine & Research, 1*(3), 177-188.

Zhang, Y., & Kaber, D. B. (2013). An empirical assessment of driver motivation and emotional states in perceived safety margins under varied driving conditions. *Ergonomics, 56*(2), 256-267. doi:10.1080/00140139.2012.739208

Zhao, N., Chen, W., Xuan, Y., Mehler, B., Reimer, B., & Fu, X. (2014). Drivers' and non-drivers' performance in a change detection task with static driving scenes: Is there a benefit of experience? *Ergonomics, 57*(7), 998-1007. doi:10.1080/00140139.2014.909952