Validation of a virtual driver assessment tool for older drivers

A Report from Research Funded by the NRMA – ACT Road Safety Trust

Prepared by Xiaolan Li, Ranmalee Eramudugolla and Kaarin J Anstey

Centre for Research on Ageing, Health and Wellbeing

Research School of Population Health, College of Medicine, Biology and Environment

+61 2 6125 8410

[kaarin.anstey@anu.edu.au](mailto:kaarin.anstey@anu.edu.au)

The Australian National University

Canberra ACT 0200 Australia

[**www.anu.edu.au**](http://www.anu.edu.au)

CRICOS Provider No. 00120C

Contents

[Executive Summary 2](#_Toc385405354)

[Preface 3](#_Toc385405355)

**Introduction**……………………………………………………………………………………….4

[Comparing performance on virtual against on-road driving environments 4](#_Toc385405356)

[Generalizability of previous validation studies 5](#_Toc385405357)

[Simulator sickness 5](#_Toc385405357)

[Variables and assessment criteria 5](#_Toc385405357)

[Matching the assessment criteria across virtual and on-road tests 7](#_Toc385405357)

[Objectives of the current study 7](#_Toc385405356)

**Study Aims**………………………………………………..……………………………………….8

**Methods**….…………………………………………………...…………………………………….9

[Materials 9](#_Toc385405356)

[Simulator Hardware and Software 9](#_Toc385405357)

[Virtual environment – development of the model 9](#_Toc385405357)

[Virtual driver assessment - scenarios 11](#_Toc385405357)

[Participants 12](#_Toc385405356)

[Motion sickness susceptibility screening 12](#_Toc385405357)

[Procedures 12](#_Toc385405356)

[Lab based sensory and cognitive assessment 13](#_Toc385405357)

[On-road Driving assessment 16](#_Toc385405357)

[Virtual Driving assessment 18](#_Toc385405357)

[Virtual Driving assessment scoring 18](#_Toc385405357)

**Results**.......…...…………………………………………………………………………………..20

[Sample characteristics 21](#_Toc385405356)

[Correlation between simulator score and on-road performance 22](#_Toc385405356)

[Impact of simulator sickness 24](#_Toc385405356)

[Adjusting for covariates 25](#_Toc385405356)

[Simulator score as a predictor of failing on-road test 25](#_Toc385405356)

[Aspects of simulator performance most associated with on-road safety 26](#_Toc385405356)

[Differences in cognitive demands between on-road and simulator 27](#_Toc385405356)

**Discussion**..………..……………………………………………………………………………..29

[Validity of the virtual assessment 29](#_Toc385405356)

[Factors contributing to performance on the virtual assessment 30](#_Toc385405356)

[Limitations 30](#_Toc385405356)

[Strengths 30](#_Toc385405356)

[Future directions 30](#_Toc385405356)

**References**.……………………..………………………………………………………………..31

# Acknowledgements

The development of the virtual driver assessment tool was funded by a project grant awarded to K. Anstey by the NRMA – ACT Road Safety Trust. The software for the driver assessment tool was developed by and purchased from Forum 8 Ltd. Participants and data for the on-road assessments and cognitive testing were drawn from an NHMRC funded study (#1045024) awarded to K. Anstey. The assessment tool was prepared with the assistance of Dr Stephen Isbel, Ms Jasmine Price, Ms Anita Byrnes (Forum 8) and Dr Joanna Brooks.

This work has been prepared exclusively by the Australian National University and is not endorsed or guaranteed by the Trust.

# Executive Summary

This study aims to develop a cost-effective computer-based driving assessment tool to validate against an on-road driving assessment.

The study also sought to determine: 1) relative validity of virtual test against on-road performance in healthy older adults, 2) sensitivity and specificity for identifying at-risk older drivers, 3) association with cognitive, sensory and motor skills. Future work will also examine 4) predictive validity against 24 month crash report data, and 5) the reliability and internal consistency of scenarios.

The virtual driving assessment was developed by creating a virtual environment modelled on the Curtin-Woden area of Canberra where the standardised on-road driving assessment is set. Five scenarios were developed based on areas of the on-road driving route that contained challenging intersections or were identified as most discriminative of at-risk drivers. Four of the five scenarios were accompanied by audio directions (‘driver instructed’), and the final scenario required self-navigation using road signs to reach a hospital. The scenarios were piloted on a small number of older adults, and a scoring protocol developed to match the criteria and scoring methods used in the standard on-road driving test.

A sample of 44 older adults who were already enrolled in a larger driving study volunteered to take part in the simulator study. All participants were screened as having low susceptibility to motion sickness, as simulator exposure can induce discomfort akin to motion sickness. Nevertheless, 5 of these participants withdrew after commencement, due to discomfort. The remaining 39 participants completed the virtual driving assessment. Analysis of the data indicated that performance on the simulator significantly predicted performance on the standard on-road assessment. The virtual assessment had a sensitivity of 67% and specificity of 79% in predicting failure on the on-road assessment. Lane position and speed control in the simulator were both correlated with on-road safety. Performance on the on-road test was also highly correlated with other measures linked to driving ability and older driver screening tests including: the Multi-D, the Useful Field of View test (UFOV), the DriveScore test, the OT-Dora Test, Trail Making Test, as well as participant age and weekly driving exposure. Correlation with cognitive assessments indicated that the virtual assessment relied particularly on participants’ spatial skills and learning.

In conclusion, a virtual driving assessment was developed that was closely matched to standard on-road driving assessment and scoring criteria. The assessment set-up is a low cost, easy to score assessment that is a valid predictor of on-road driving performance. Future work will establish its reliability with repeated testing, as well as its validity against long-term driving behaviour. The assessment can then be used as a safe and cost-effective outcome measure in studies of interventions or medication effects on driving.

# Preface

The research described in this report was approved by the Australian National University Human Research Ethics Committee, under protocol 2012/643

Main related research output as of December 2014:

Li, X., Eramudugolla, R., Anstey, K.J. (2014) Development of a virtual driver assessment tool for older drivers. Poster presented at the 2014 ARRB Conference, Sydney.

# Introduction

In Australia, road safety is an ongoing public concern and quality evidence-based approaches to transport intervention are increasingly sought by policy makers and law enforcement agencies [[1](#_ENREF_1)]. Research in population health, ageing, substance-abuse, mental health and rehabilitation can provide significant input into informing road safety interventions [[1-6](#_ENREF_1)]. This report details the development of a desktop simulator-based virtual driving assessment. The virtual driving test was validated against on-road performance in a sample of healthy older drivers. A well-validated virtual driving task will significantly enhance capacity for research and development into driver safety.

Relative to on-road testing, driving simulators provide a controlled and safe environment for measuring driving skills in individuals with health conditions that potentially compromise their ability to safely drive a motor vehicle. Although an on-road assessment provides the ultimate measure of driver ability and safety, it is usually associated with minimal experimental control [[7](#_ENREF_7)], precluding the investigation of specific variables and conditions that contribute to road crashes [[8](#_ENREF_8)]. Modern simulator technology and graphics are such that recent validation studies indicate simulators can be used as a reliable index of driving performance and that they can be used to estimate driving performance as measured by an on-road test [[4](#_ENREF_4), [7](#_ENREF_7), [9-12](#_ENREF_9)]. Importantly, many of these studies have used desktop simulators with one or multiple screens, and a fixed-base (i.e., no simulation of physical forces and feedback associated with a moving vehicle). This makes the technology accessible to a range of applied assessment settings, as well as research settings beyond dedicated driver research facilities.

The present study is focused on developing and validating a virtual driving test that can identify at-risk older drivers and can be used as an outcome measure in future research into older driver interventions, rehabilitation, and the effect of medical and physical factors on older driver behaviours.

**1.1 Comparing performance on virtual against on-road driving environments**

Despite the flexibility and control that is offered by simulator technology, and the promising results of recent validation studies, there are several important methodological issues that limit comparison between the two modes of assessment and limit generalizability of findings across contexts.

**1.1.1 Generalizability of previous validation studies**

Studies that have compared driving simulators against on-road driving have done so for a diverse range of applications. These include examining the validity of driving simulator performance to differentiate between groups with differing levels of driving experience [[4](#_ENREF_4)], in older drivers [[11](#_ENREF_11)], to evaluate in-vehicle information systems [[13](#_ENREF_13)], predict crash risk in individuals with neurological injury or disease [[7](#_ENREF_7), [10](#_ENREF_10)], to assess the safety of particular traffic intersections [[12](#_ENREF_12)]. Only one previous study has sought to validate a simulator test in older drivers [[11](#_ENREF_11)].

* 1. **.2 Simulator sickness**

Simulator sickness refers to a range of symptoms that are brought about as a result of exposure to simulator features. These symptoms are similar to motion induced sickness, such as nausea, ocular discomfort, headache, dizziness and vomiting[[14](#_ENREF_14)]. As with motion sickness, susceptibility to simulator sickness varies across individuals, with an estimated 20% of individuals likely to report feelings of sickness during exposure to virtual environments [[15](#_ENREF_15)], with older adults more likely to experience simulator sickness [[16](#_ENREF_16)]. While there are screening questionnaires available for identifying and excluding participants who are likely to experience simulator sickness [[14](#_ENREF_14), [15](#_ENREF_15)], between 5-10% of young adults and 10-30% of older adults with *minimal* susceptibility are unable to complete simulator studies due to sickness symptoms [[16](#_ENREF_16)]. This issue limits the application of simulator technology, particularly for older drivers. It is possible that with improvements in technology and fidelity of virtual environments, simulator sickness may become less prevalent – however, currently this factor affects the comparability of virtual and real driving experience.

**1.1.3 Variables and assessment criteria**

While it is difficult to control for simulator sickness, and for differences in the sensory and motor feedback between the virtual and real environments, most validation studies have used very different types of variables, routes and criteria to assess performance in on-road compared to the virtual driving test. On-road driving assessments are typically open-circuit, involving suburban or university roads, with multiple skills being observed and rated by an experienced evaluator throughout the route. In contrast, many virtual driving assessments use short scenarios designed to test a single skill, and rely on computer measured variables with a degree of accuracy that cannot be replicated in on-road driver testing.

Differences in route and behavioural variables mean that new indices of performance need to be statistically derived to assist comparing on-road against simulator performance [[11](#_ENREF_11), [17](#_ENREF_17)]. For example, one study assessed 129 older participants in a simple driving simulator and validated performance against an on-road test conducted on a predefined route [[11](#_ENREF_11)]. In this study, two separate assessment criteria were used for the simulator test and the on-road driving test. The simulator test employed 10 short scenarios, each designed to test a specific aspect of driving – for example, the ability to regulate driving speed was assessed by driving on a 2km stretch of simulated road at the designated speed. Multi-tasking was assessed by having the participant drive along a simulated road and doing an arithmetic operation each time a billboard appeared on the side of the road. In contrast, the on-road assessment was conducted in the participant’s own car, and criteria included skills that were rated throughout the assessment – e.g., steering and breaking, as well as during specific sections of the route – e.g., regulating speed along a 500m section of the route. The most informative measures for the on-road and simulator assessment were derived using principal components analysis, and an index was created for each using an arithmetic combination of these measures. Although results showed a significant association between the simulator driving index and the on-road driving index (Table 1, [[11](#_ENREF_11)]), it is difficult to directly map or apply simulator performance to expected on-road performance, reduces ecological validity, and increases the chance that the association reflects correlation with an intermediate variable.

Table 1: Driving simulator assessments validated against on-road assessment

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Study | Simulator | Outcome Measure | Sample Age (years) | N | Measure of Association |
| [[7](#_ENREF_7)] | STISIM | On-Road | TBI sample 29yrs (SD=12) | 11 TBI  16 control | *r* =0.26 (*p*>0.10) |
| [[4](#_ENREF_4)] | STISIM | On-Road | 16 years | 60 | Kendall’s tau = 0.23, *p*<0.01 |
| [[11](#_ENREF_11)] | STISIM | On-Road | 72.9yrs(SD=7.1) | 129 | R2=0.657 |
| [[10](#_ENREF_10)] | STISIM | On-Road | HD sample 50yrs(SD=13) | 29 HD | *r*s=0.72 (*p*<0.001) |
| [[17](#_ENREF_17)] | Trainer F12PT-1L40 | On-Road | 72.1yrs(SD=5.4) | 49 | R2=0.50 (*p<*0.001) |

**1.1.4 Matching the assessment criteria across virtual and on-road tests**

Only three studies to our knowledge have validated virtual assessments against on-road assessment using relatively closely matched routes, variables and assessment criteria [[7](#_ENREF_7), [10](#_ENREF_10), [18](#_ENREF_18)], however two were used in very small samples for the purpose of identifying difficulties in patient populations – traumatic brain injury [[7](#_ENREF_7)] and Huntington’s Disease [[10](#_ENREF_10)]. One of these studies [[10](#_ENREF_10)] used a standard checklist used for on-road testing, the Test Ride for Investigating Practical fitness to drive (TRIP), and adapted this for simulator use. Variables required for TRIP criteria were matched as closely as possible for the on-road and simulator assessments. Both the on-road and simulator routes progressed gradually from low to high traffic density with gradual introduction of hazards. However, the simulator scenario included an additional task where participants were required to attend to unrelated symbols (pictures of horns or turn signals) throughout the virtual environment and respond to these while driving. The simulator-based TRIP measure was significantly associated with the on-road TRIP measure, and was able to classify at-risk drivers with a sensitivity of 62% and specificity of 94%. The third study [[18](#_ENREF_18)] used the low cost UC-win/Road driving simulator to model the same route as used in the on-road test, and used similar scoring procedures for both in a sample of young adults, however validity was examined only by analysing the difference between mean scores for the two tests.

**1.2 Objectives of the current study: validation of simulated driving assessment in older adults**

The present study aims to develop a virtual driving assessment modelled on a realistic driving assessment route, to assess older drivers. The virtual assessment will then be validated against performance on a standardised on-road assessment conducted by an Occupational Therapist with driving expertise. The study aims to match the environment, route and scoring criteria across both the on-road and virtual mediums, in order to produce a test that can be translated to other research projects.

The validation study will be conducted as a sub-study within a large, NHMRC funded (2012-2017) multi-site study on older drivers. This study already provides the research infrastructure required for recruitment, testing and follow-up of the sub-sample for the virtual assessment validation study.

# Study Aims

1. Development of the virtual driving assessment and task scenarios
2. Data collection from ACT older drivers on virtual assessments
3. Data analysis and comparison with on-road, off-road and follow-up data.

# Methods

## 2.1 Materials

**2.1.1 Simulator Hardware and Software**

As computer graphic capabilities has improved, simulation technology has become more accessible. However, driving simulators most frequently used in the literature (e.g., STISIM Drive®) remain very costly, require high-end hardware, and are limited in the flexibility available to researchers regarding the design of the environmental model and route. Forum 8 software UCWin/Road® was chosen due to its low cost and capacity to create an environmental model based on real terrain and geographic maps without the need for complex programming. UCWin/Road® has adapted civil engineering modelling software to include programmable driving scenarios and data recording. The software was run on a desktop PC (Dell XPS 8700, Core i7-4770 CPU, 3.5GHz, Quad Core 8), with nVidia GeForce GTX 780 3GB 2xDVI, HDMI graphics card; 3 x 17” LCD monitors and game-type (Playstation) 27cm steering and brake/accelerator pedals (Logitech 25®).

## E:\New Image.JPG

Figure 1. Set-up of the desktop simulator.

Logitech driving force steering wheel and pedals. Screens arranged to subtend 90˚ field of view.

**2.1.2 Virtual environment – development of the model**

The topographical mesh terrain corresponding to the Curtin-Woden valley area of Canberra (GIS), in which the on-road route is located, was imported for use in the Forum8 UC Win/Road software using the appropriate geographic coordinates of the area. A high resolution aerial map of the area (2012) was obtained from the ACT Government’s Office of the Surveyor-General Environment and Sustainable Development Directorate. This map was overlaid on the terrain as a guide to the road and built area locations (Figure 2). The road tools within UC Win/Road was used to build the roads and intersections using the aerial maps as a template Figure 3 and Figure 4.

Figure 2

*Figure 2.*

*Section of aerial map of Red Hill / Yarralumla.*

Figure 3

*Figure 3*

*A. Road layout for Yamba drive roundabout; B. Adjusting elevation to terrain; C. Intersection markings and other features obtained from Google Maps street view. D. Model of roundabout.*

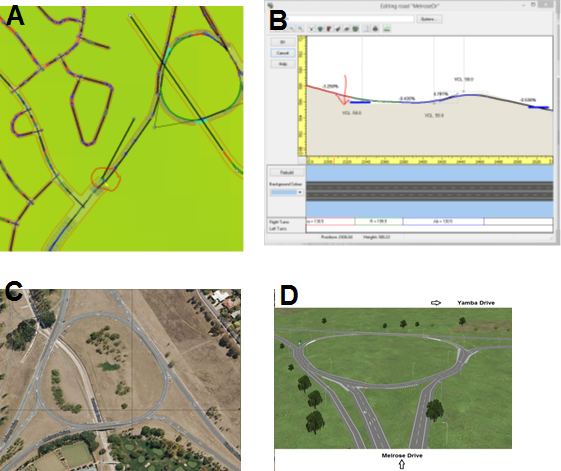
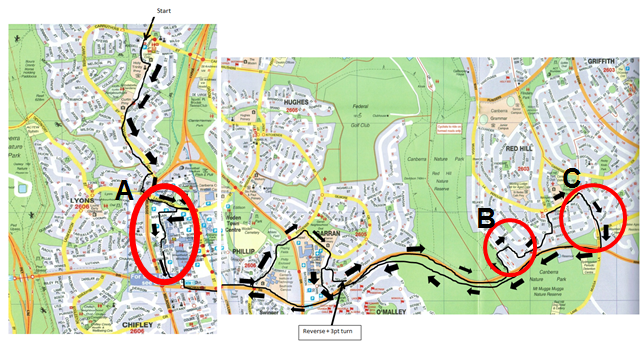


Figure 4

*Figure 4. Buildings and other features of the environment were matched, where possible, to simulate the actual locale.*

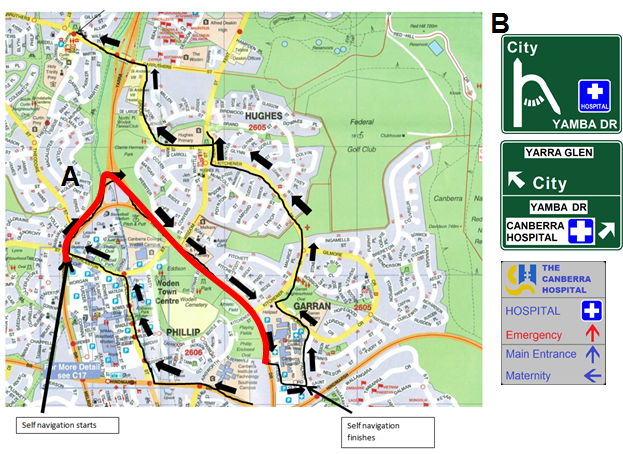
### 

### 2.1.3 Virtual Driver Assessment – Scenarios

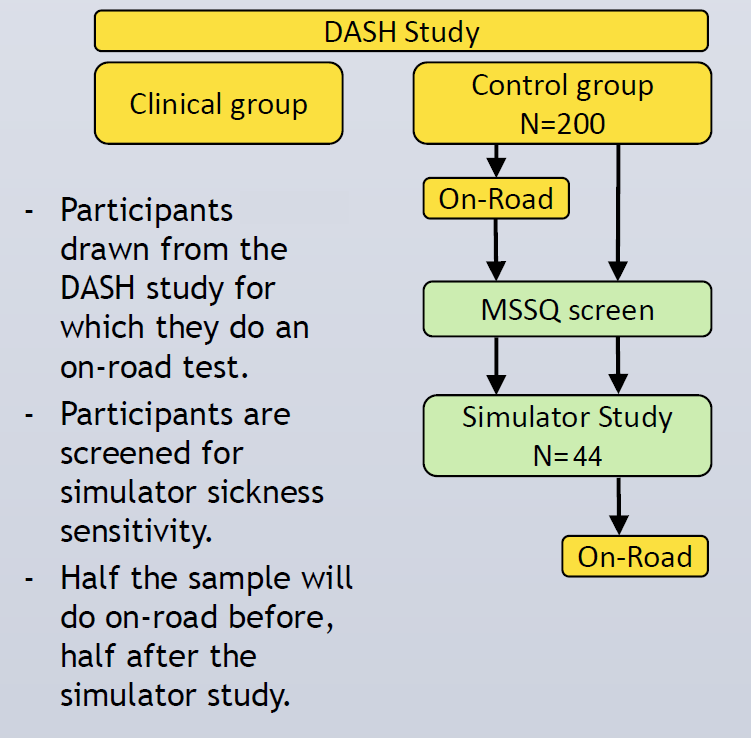
Five scenarios were developed based on consultation with the Occupational Therapist responsible for assessment for the on-road test. Four areas of the on-road route were identified as involving challenging intersections and conditions: 1.Instructor directed but ‘easy to miss’ right turning; 2. Shopping strip with multiple pedestrian crossings; 3. Crossing a Give-way and Stop sign controlled intersection with a two-way road; 4. Complex right turn at a bend (Figure 5 A-C).The fifth scenario was modelled on a section of the on-road route that required the participant to navigate to the nearest hospital using the available road signs and via a large and complex roundabout.

*Figure 5. Driver Instructed scenarios: A. Scenarios 1 and 2: Right turn and shopping strip. B. Scenario 3 Give-way and stop sign controlled intersection. C. Complex right turn.*

*Figure 6. Self-Navigation: A. section of the route for self-navigation. B. Three signs indicating route to hospital.*



## 2.2 Participants

Forty four community-dwelling older adult drivers in Canberra, ACT with a valid driving license volunteered to participate in the study. The sample age ranged from 65 to 88 years (Mean = 75.6, SD=5.87), with 79.5% of the sample being male drivers. Recruitment occurred between March 2014 and December 2014. Participants were recruited through a larger study on older drivers funded by the NHMRC (Driving, Ageing, Safety and Health – DASH) with a participant base of N=230 as at December 2014.

*Figure 7. Flow of participants through DASH and simulator study.*

### 2.2.1 Motion Sickness susceptibility Screening

Participants expressing interest in the study were initially given a screening questionnaire over the telephone. This questionnaire, the Motion Sickness Susceptibility Questionnaire (MSSQ) is a validated measure of simulator sickness susceptibility [[19](#_ENREF_19)]. The total MSSQ score was calculated based on the childhood and adulthood experiences of motion sickness in a variety of situations (e.g., car, bus, boat, ship, fun fair rides, plane, etc.) and a cut-off of MSSQ>60 was used to screen out individuals who may be susceptible to sickness in the simulator. The cut-off was based on the published norms and our pilot testing.

## 2.3 Procedures

### Data collected from general health and driving questionnaires

All participants were enrolled in the NHMRC funded DASH study, and consented to using the data from that study to validate the simulator assessment. As part of the DASH study, the participants completed a set of questionnaires examining their demographics, health and driving experiences. The data obtained for the current study included: date of birth, gender, level of education, years of driving, average kilometres driven in a week, average number of days driven, crashes in the past 12 months, crashes in the past 5 years. The questionnaires were mailed to participants upon enrolment in the study and were completed and checked for completion upon return by a researcher.

### 2.3.1 Lab based sensory and cognitive assessment

As part of the DASH project, all participants were assessed on a battery of cognitive and older driver screening tests within a laboratory setting. The test results used for the present study included a measure of global cognitive function called the Mini-Mental Status Exam (MMSE), the following tests of vision and cognition:

*Visual function:*Assessed using a battery of tests undertaken binocularly (with the exception of visual fields which will be undertaken monocularly and combined to derive an integrated field as used in our previous studies. The refractive correction used habitually for driving (and worn for the driving assessment described here) will be used in conjunction with the appropriate correcting lens for the test working distance.

* Static high contrast visual acuity will be measured using the Australian Vision Chart 5, which uses logMAR principles, at a working distance of 4.0 m (NVRI, Melbourne).
* Letter contrast sensitivity will determined using the Pelli-Robson chart under the recommended testing conditions.

*Cognitive Function battery*

1. *Boston Naming Test*(BNT) measures the respondent’s ability to name objects. The BNT has been widely used as a neuropsychological assessment tool to measure confrontational word retrieval in individuals with language disturbance caused by stroke, Alzheimer's disease, or other dementing disorder.
2. *Controlled Oral Word Association Test* (FAS) is a measure of verbal fluency. Respondents are required to generate as many words as possible in 1 minute that start with a specific letter.
3. *The California Verbal Learning Test (CVLT)* is a well-established and validated neuropsychological test of verbal memory. The respondent is asked to learn a list of words over three trials, and later recall the same words after a delay of 20 minutes (Wagner et al 2012).
4. *Trails B* is a brief measure of attentional flexibility. The measure is sensitive to driving difficulties as well as dysexecutive symptoms.
5. *Digit Span Backwards* has been taken from WAIS-R and provides a brief measure of working memory span.
6. *Stroop colour word interference* is a brief, computerised and widely used measure of feedback error utilization / inhibition. The test is sensitive to detecting difficulties associated with normal aging and the early signs of Alzheimer’s Disease*.*
7. *The Game of Dice Task* is a brief, computerised test measuring decision-making under explicit risk (Brand, M. and Schiebener, 2012)
8. *Reading the Mind in the Eyes Test* measures theory of mind and social cognition and is available to complete online (Baron-Cohen, 2001) and can help identify isolated theory of mind deficits and riskfor frontotemporal dementia. This test is already part of our battery. Respondents are presented with 36 sets of eyes, and for each one, instructed to choose which one of four words best describes what the person in the picture is thinking or feeling.
9. *Benton Visual Retention Test* has been widely used in other major epidemiology studies. It is quick and easy to administer and measures perceptual-motor abilities. The Benton Visual Retention Test is an individually administered test for people aged from eight years to adulthood. The respondent is shown a series of designs, one at a time, and asked to draw each one as accurately as possible from memory. The test is untimed, and the results are professionally scored by form, shape, pattern, and arrangement on the paper.

The battery of tests also included a set of measures used to screen older drivers to identify at-risk individuals. These included the following:

*The Multi-D test battery:* This consists of three subtests assessing complex reaction time, sensitivity to visual motion, and postural balance.

1. A Colour Choice Reaction Time test was administered on a computer requiring hand and foot responses as well as inhibition of responses. The hand and foot switches are mapped to one of four locations on the screen in which an image of a red car appears at random. The participant must respond by pressing the appropriate switch as quickly as possible, and inhibiting this response if the car is blue in colour.
2. Dot motion test is a computer-based measure of central motion sensitivity using random dot stimuli presented a working distance of 3.2m. Participants are required to briefly view the random dot display and identify the direction of motion. The coherence of random dot motion is systematically reduced in order to identify the point at which an accuracy of 75% is reached.
3. Postural sway (displacement of the body at the level of the waist) was measured with a sway meter with the participant standing on a foam rubber mat (40cm x 40cm x 15cm thick) of medium density with their eyes closed.

*The Useful field of view (UFOV®)* test is PC-based and linked to a touch screen (17 in.). For the present study subtest two of the UFOV® was used. Participants attend to dual targets presented simultaneously on a screen: a white, schematic image of a car or truck is presented centrally with a second car figure presented randomly in one of eight locations along eight radial spokes (location from the upper vertical: 0°, 45°, 90°, 135°, 180°, 225°, 270° or 315°) at a peripheral eccentricity of 10°. Following presentation, a random noise mask is shown and participants indicate: (a) which vehicle was presented in the centre of the screen by pressing a picture of a car or truck; and (b) where the second car was located by pressing one of the eight possible locations onscreen.

*DriveAware/DriveSafe: University of Sydney Visual Slide Recognition Test* (VRST–USyd) will be administered according to standardized instructions. This test consists of 15 images of the same roundabout, projected on a screen to simulate the view through a windshield, in which the number and position of pedestrians and vehicles vary. Participants are asked to observe each image for 3 sec and, when the image has been removed from the screen, to report details about the position and direction of travel of each pedestrian and vehicle in the slide. The published raw score cut-off of 95/164 were used to classify safe versus unsafe drivers. This test is widely used by occupational therapists. In a recent paper, the authors of this instrument updated the test (now called DriveSafe) and included a new scale called DriveAware**.** Both were used in the baseline assessment.

*The 14-item Road Rules and Road Craft test.* This test comprises 14 questions about road safety.

*OT Drive Home Maze Test:* The Occupational Therapy-Drive Home Maze Test (OT-DHMT) is part of the OT-DORA Battery used in licensing recommendations for older and/or functionally impaired drivers. Unsworth, C.A. Pallant, J.P., Russell, K.J. & Odell, M. (2011).  OT–DORA Battery: Occupational Therapy Driver Off-Road Assessment Battery. Bethesda, MD: AOTA Press.

### 2.3.2 On-Road Driving Assessment

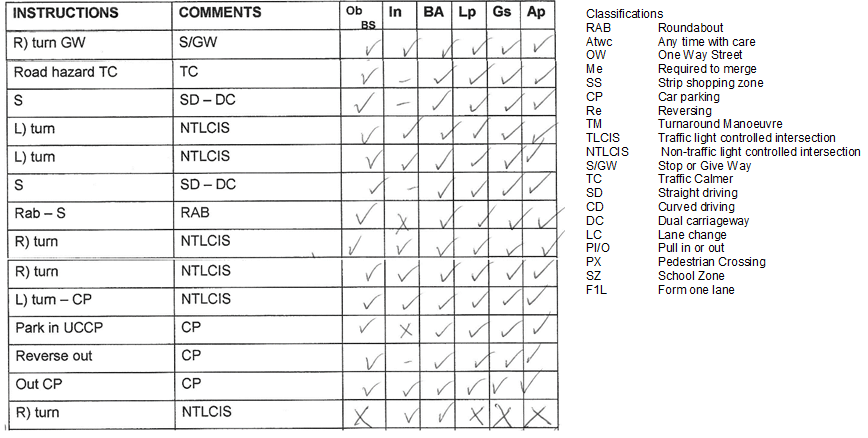
As part of the DASH study, participants underwent a 45-minute on-road driving assessment on a standard route around the Curtin-Woden region of the Canberra, in a dual-brake car with a qualified driving instructor and a driving Occupational Therapist (OT). The OT scored the participants’ driving skills using a standardised criteria that has been used in other driving studies. Sections of the route are classified according to the type of road, intersection or turning is involved and marked against the 6 criteria (observation (Ob), Indication (In), Brake/Acceleration (BA), Lane position (Lp), Gap selection (Gs) and Approach (Ap (Figure 8)). A single error is marked as a cross, and a tick is marked in the absence of an error. As indication is not required at every section of the route, this is not marked unless at intersections or where lane changes are involved. Scoring involves awarding a point of 1 for each section of the route with no errors, and a point of 0 for any section with one or more errors. The points are tallied to produce an overall score. The occupational therapist uses this and general behavioural observations to allocate a safety rating between 1 and 10.

Figure 8. Example of assessment record for on-road test using the standard scoring.

**On Road Driving Assessment Rating Scale**

1. The driver should cease driving. The Instructor took physical intervention and the assessment was terminated within the first 30 minutes of the drive. A crash would have occurred without intervention from the instructor. ***Would not pass a Canberra driving test.***
2. The driver should cease driving. The Instructor took physical intervention to prevent a crash or a near collision or to prevent hostility from other drivers. ***Would not pass a Canberra driving test.***
3. Very poor driving performance. Instructor gave verbal or physical intervention(s) to prevent an incident. The driver should consider their driving future. ***Would not pass a Canberra driving test.***
4. Poor. Verbal intervention from instructor. Requires excessive education and retraining to continue driving, **OR**, the driver committed one critical error requiring physical intervention to prevent an incident, in an otherwise good performance. ***Would not pass a Canberra driving test.***
5. Below average. Driver has the ability to learn and retrain to bring their driving up to standard that would allow them to pass a test. ***Many faults and bad habits and therefore may or may not pass a Canberra driving test.***
6. Average. Requires some retraining. Lacks some defensive skills. For example speeding above 10 klm/hr, not checking blind spots sufficiently, poor use of signals. ***May or may not pass a Canberra driving test.***
7. Average. Some poor habits and behaviours. Eg. Speeding between 5-10 klm/hr, some lack of use of signals and blind spots. ***Would pass a Canberra driving test.***
8. Above average. Few and minor habits. Eg. Speeding by less than 5klm/hr, forgetting an indicator. ***Would pass a Canberra driving test.***
9. Good. Very few faults or habits that affect good driving behaviour. ***Would pass a Canberra driving test.***
10. Excellent. No faults. ***Would pass a Canberra driving test.***

As noted above, drivers rating from 1 to 4 would not pass a Canberra Driving Test. Drivers rating 7 to 10 would pass the test. Drivers that rate 5 or 6 may or may not pass a driving test. This would depend on examiners personal opinion and luck of the test route.

### 2.3.3 Virtual Driver Assessment

The participant was seated directly in front of the simulator set-up (Figure 1) positioned 50cm from the middle screen. The participant was familiarised with the controls, and advised that the virtual car had automatic transmission. As there was no functional indicator stalk on the Logitech steering set-up – a makeshift stalk was attached to the side of the steering which. Participants were asked to press the stalk when indication was required. With the participant’s consent a video recorder was set up to focus on the participant’s head and to capture observation and indicator stalk usage.

Based on pilot testing, two arrangements of the screens (wide view 100º, and narrow view 90º) were chosen as options for participants in terms of visual comfort. After choosing a comfortable screen angle, the researcher administered a brief baseline simulator sickness questionnaire. This questionnaire was designed to monitor symptoms during exposure to the simulator (see Appendix A). Four practice driving scenarios were then administered in the following order – 1) straight driving and stopping at a red light; 2) straight driving followed by a left turn via a slip lane at a traffic light controlled intersection; 3) straight driving followed by a right turn at a traffic light controlled intersection; 4) straight driving and straight through a roundabout. These practice scenarios were designed to introduce the controls and the physics of driving within the virtual environment. The simulator sickness questionnaire was re-administered following the practice scenarios. Participants were permitted to repeat the practice scenarios as often as required in order to feel confident driving in the virtual environment.

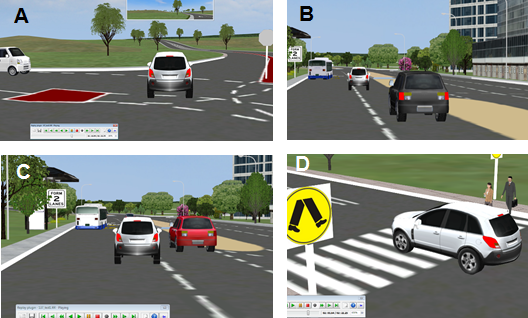
The test scenarios were administered in the same order to all participants, except for cases where the software crashed and required re-starting or moving on to a different scenario before returning to the problematic scenario. In general, all five scenarios were equally likely to crash during testing – and about 20% of participants were affected by this issue. The general order of scenarios were: 1) Driver Instructed – Launceston Street to Corinna Street via Worgan street; 2) Driver Instructed – Corinna Street Shopping strip; 3) Driver Instructed – Dalrymple street and intersection; 4) Driver Instructed – Tamar street and intersection with stop and give-way signs; 5) Self-Navigation – to hospital.

The simulator sickness questionnaire was re-administered at two further time points between the test scenarios to yield a total of four measures: baseline and 3 repeats during exposure.

The researcher (simulator assessor) sat next to the participant with their chair positioned behind and to the left as if in the rear passenger seat – where the driving assessor is typically located during on-road assessments. The simulator assessor recorded the participant’s behaviour according to standard criteria. Each scenario was saved for review using the play back plug-in.

### 2.3.4 Virtual Driving Assessment Scoring

Each scenario was scored in a similar way to the on-road driving test using the criteria for Observation, Indication, Brake/Acceleration, Lane Position, Gap Selection and Approach. The researcher was trained by the Occupational Therapist to use the same scoring criteria as is used in the on-road assessment. Each participant’s simulator performance was scored by viewing the computer recorded replay, and using the simulator assessors (Xiaolan Li) observations during the assessment. A separate researcher (Ranmalee Eramudugolla) reviewed and scored the data for 44 participants and was blind to participant identity, cognitive or on-road results.



*Figure 9. Screen captured examples of errors during virtual assessment (participant = white hatchback). A. Gap selection error at a non-traffic light controlled intersection in scenario 3 (NTLCIS); B and C. Gap selection errors during lane change in Scenario 1. D.* Failure to brake at a pedestrian crossing.

For the purposes of the present study, the video recording of observation and indicator use was not reviewed. However, errors in observation and indication were marked where the simulator assessor had recorded this. Scoring involved marking any error in lane positioning, breaking/acceleration, gap selection, approach, observation or indication at specific sections of the route (Figure 9). For example, drifting in and out of the lane during a curved driving section of a scenario (CD) would be scored as one error in lane positioning. Similarly, if a right turn requiring preparation with a lane change results in a near miss, a gap selection error would be marked during the preceding straight driving section (SD) and the ensuing right turn at the non-traffic light controlled intersection (Right Turn, NTLCIS) would be marked as an error in approach. The five scenarios were divided into the same sections as in the corresponding sections of the on-road score sheet. Below is an example of the scoring sheet for the virtual assessment (Scenario 1).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **ORT date** | **Sim date** |  |  |  |  |  |  |
| **ID** |  | **Scenario 1 - Launceston** | | | | | |
| **Repeats** |  | **Ob** | **Ind** | **Bra/Acc** | **Lane pos** | **Gap sel** | **App** |
| S lights | TLCI |  |  |  |  |  |  |
| S | SD-DC |  |  |  |  |  |  |
| R Turn | NTLCIS |  |  |  |  |  |  |
| S | SD-SL |  |  |  |  |  |  |
| R Turn | NTLCIS |  |  |  |  |  |  |
| Road Haz | TC-PX |  |  |  |  |  |  |
| Ped Xing | PX |  |  |  |  |  |  |
| L Turn | S/GW |  |  |  |  |  |  |
| Ped Xing | PX |  |  |  |  |  |  |
| SD | SD-SL |  |  |  |  |  |  |
| Right Bend | CD |  |  |  |  |  |  |
| L Turn | S/GW |  |  |  |  |  |  |

A total of 48 points could be awarded for the whole assessment, with each scenario having a maximum score of either 12 or 6.

|  |  |
| --- | --- |
| **Scenario** | **Total possible score** |
| 1 | 12 |
| 2 | 12 |
| 3 | 6 |
| 4 | 6 |
| 5 (Self-Nav) | 12 |
| **Total** | **48** |

## Results

### 3.1 Sample characteristics

Participants enrolled in the simulator study were slightly younger than those enrolled only in the main DASH study (N=187: those with entered data), and had a slightly better on-road driving performance although this was not significantly different (Simulator: 6.14(1.32) vs DASH = 5.89(1.67), p>0.10). The simulator participants’ performance on the global cognitive measure (MMSE), memory language and visuospatial measures, and the UFOV measure were also not different from that of DASH-only participants (all p>0.10). However, simulator participants performed better on the other driver screening batteries than did DASH-only participants – where their predicted fail rate was lower on the Multi-D (simulator=17.32% (23.6), vs DASH=31.1% (28.8), *p*<0.01), and their score was higher on the DriveSafe measure (simulator=91.3(13.1) vs DASH = 83.3 (15.1), *p*<0.01). Five participants (11.3%) withdrew from the study due to simulator sickness symptoms. These participants were not different from those who completed the study in terms of age (withdrawn=76.8yrs (SD=7.29), completed=75.4yrs (SD=5.7), *p*>0.10), on-road performance (withdrawn=6.0(SD=0.71), completed=6.15 (SD=1.4), *p*>0.10), or motion sickness susceptibility (withdrawn=9.9 (SD=10.5), completed=10.6 (SD=12.2), *p*>0.10), but were more likely to be female (withdrawn=60%, completed=15%, *p*<0.05).

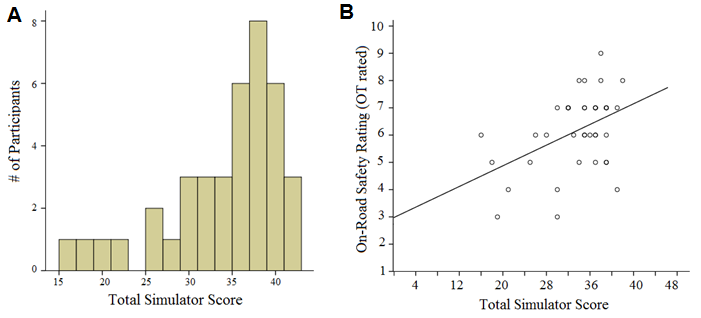
Table 1: Participant characteristics – Driving history and demographics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Mean | SD | Range | | Correlation with  On Road OT Safety Rating |
|  |  |  | **Min** | **Max** |  |
| Age (years) | 75.55 | 5.87 | 65 | 88 | **-0.352\*** |
| Gender (male%) | 79.50 |  |  |  | -0.182 |
| Total years of education | 15.50 | 4.67 | 6 | 26 | -0.032 |
| On road OT safety rating (1-10) | 6.14 | 1.32 | 3 | 9 | - |
| Years of driving experience | 55.80 | 6.94 | 32 | 68 | -0.086 |
| Average days per week driven | 5.34 | 1.59 | 2 | 7 | 0.106 |
| Average kilometres per week driven | 161.88 | 93.34 | 20 | 332 | **0.365\*** |
| Crashes over past 12 months | 0.10 | 0.30 | 0 | 1 | 0.009 |
| Crashes over past 5 years | 0.44 | 0.71 | 0 | 3 | -0.165 |
| Motion sickness susceptibility score | 10.56 | 11.89 | 0 | 42 | 0.176 |

\* statistically significant correlation at p<0.05

Participants in the simulator study had a generally high level of education (Table 1), although this ranged between 6 and 26 years. Their average distance travelled per week was 161.9 (93.3) km, and most reported having had no crashes with an average of 0.44(0.71) crashes in the preceding 5 years (ranging from 0-3). Performance on the on-road test, was significantly correlated with age, and distance driven per week – with older drivers and those with less driving exposure having poorer performance on the on-road test (Table 1).

### 3.2 Correlation between Simulator Score and on-road performance

The total scores for simulator performance ranged between 16 and 42 (out of 48), with a mean of 33.51 (SD=6.52). The distribution of scores was negatively skewed (Figure 10), with most participants scoring between 35 and 40 and long ‘tail’ representing those with lower scores. In order to examine whether simulator performance was related to real driving performance, bivariate correlations between the Total simulator score, the On-Road Test score, and scores on commonly used older driver screening measures was examined. Demographic measures were also included in these analyses.

*Figure 10 A Distribution of Scores for Simulator Performance; B Relationship between Simulator performance and on-road test performance.*

Bivariate correlations (Table 2, Figure 10B) revealed a significant positive relationship between Total Simulator Score and the On-Road Test Safety Rating (*r* = 0.445, *p*<0.01). The Simulator Score was also significantly associated with other driver characteristics also correlated with on-road performance, including the self-reported distance travelled per week (*r* = 0.407, *p*<0.05), where participants who reported less driving exposure each week also performed less well on the simulator scenarios (Table 2). Performance on the simulator was also significantly correlated with performance on off-road measures typically used to screen older drivers. This included the Trail Making Test (*r* = -0.349, *p*<0.05), the DriveSafe Test (*r* = 0.566, *p*<0.01) the OT DORA Maze Test (*r* = -0.334, *p*<0.05), the MultiD Test (*r* = -0.534, *p*<0.01) and the Useful-Field of View (UFOV) test (*r* = -0.466, *p*<0.01). The relationship with these measures suggest that poorer performance on the simulator scenarios was associated with slower ability to switch attention between simultaneous tasks (Trail Making Test), longer time to navigate through a Maze, poorer attention and recall of objects in a traffic scene (DriveSafe Test) and slower processing of

Table 2: Correlations between simulator and on-road test and other driving measures

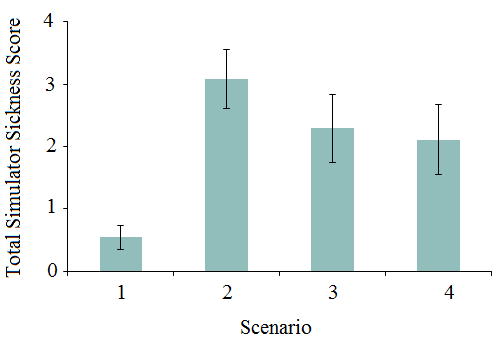
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Range | | Correlations (*r*) | |
|  | **Mean** | **SD** | **Min** | **Max** | **Simulator Score** | **On-Road** |
| Age (years) | 75.55 | 5.87 | 65 | 88 | -0.235 | **-0.352\*** |
| Gender (male%) | 79.50% |  |  |  | -0.288 | -0.182 |
| Total years of education | 15.50 | 4.67 | 6 | 26 | 0.133 | -0.032 |
| On road OT safety rating (1-10) | 6.14 | 1.32 | 3 | 9 | **0.445\*\*** | - |
| Years of driving experience | 55.80 | 6.94 | 32 | 68 | -0.023 | -0.086 |
| Average days per week driven | 5.34 | 1.59 | 2 | 7 | 0.183 | 0.106 |
| Average kilometres per week | 161.88 | 93.34 | 20 | 332 | **0.407\*** | **0.365\*** |
| Crashes over past year | 0.10 | 0.30 | 0 | 1 | 0.034 | 0.009 |
| Crashes over past 5 years | 0.44 | 0.71 | 0 | 3 | -0.113 | -0.165 |
| Visual Acuity - Pelli Robson test | 1.65 | 0.06 | 1.45 | 1.85 | 0.187 | -0.103 |
| Visual Acuity - high contrast | 0.00 | 0.13 | -0.22 | 0.44 | -0.301 | **-0.361\*** |
| Trail Making Test (secs) | 111.71 | 63.55 | 35 | 319 | **-0.349\*** | **-0.414\*\*** |
| Multi-D Score (% likely fail rate) | 17.32 | 23.62 | 0.36 | 95.70 | **-0.534\*\*** | **-0.440\*\*** |
| DriveSafe Score | 91.27 | 13.12 | 53 | 121 | **0.566\*\*** | **0.330\*** |
| DriveSafe Intersection Test | 6.25 | 1.37 | 2 | 8 | **0.718\*\*** | **0.367\*** |
| 14 Item Road Law Test | 30.64 | 4.16 | 21 | 37 | 0.104 | **0.349\*** |
| OT DORA MazeTest (secs) | 24.68 | 12.86 | 10 | 71 | **-0.334\*** | -0.250 |
| UFOV Score (17-500ms) | 129.73 | 126.58 | 17 | 500 | **-0.466\*\*** | **-0.322\*** |

\* *p*<0.05, \*\* *p*<0.01

visual information (UFOV). The MultiD score represents a predicted on-road fail rate (%) based on the participant’s complex reaction time, sensitivity to visual motion and postural balance. Those with a higher predicted fail rate tended to have poorer performance on the simulator. Performance on the simulator scenarios was not related to participant age, gender or level of education.

**3.3 The impact of simulator sickness**

Simulator related variables such as sensitivity to motion sickness, level of sickness symptoms during testing, amount of practice required on the simulator, and familiarity with the on-road test prior to undertaking the simulator test are likely to impact on simulator performance. Mean simulator sickness scores taken during testing indicated that on average, participants reported increased levels of sickness prior to scenario 2, and that this reduced slightly over the course of the next few scenarios (Figure 11). Bivariate correlations with this and other variables were examined.

****

*Figure 11. Mean simulator sickness score as a function of testing point (scenario). Error bars represent +/-1 standard error of the mean.*

Simulator performance was found to be significantly correlated with sickness levels reported during the first scenario (*r* = -0.359, *p*<0.05), such that higher levels of sickness symptoms was associated with poorer simulator performance. None of the other variables were significantly related to simulator performance.

Table 3: Correlation between simulator score, simulator sickness, days since ORT and practice sessions

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Range | | Correlations (*r*) | |
|  | **Mean** | **SD** | **Min** | **Max** | **Simulator Score** | **On-Road** |
| Days Since On-Road Test | 61.62 | 77.15 | -64 | 348 | 0.212 | 0.164 |
| Motion sickness susceptibility | 10.64 | 12.18 | 0 | 41.58 | 0.143 | 0.158 |
| Sickness score 1 (SSQ1) | 0.54 | 1.19 | 0 | 6 | **-0.359\*** | 0.012 |
| Sickness score 2 (SSQ2) | 3.08 | 2.97 | 0 | 10 | -0.039 | -0.067 |
| Sickness score 3 (SSQ3) | 2.29 | 3.21 | 0 | 12 | 0.080 | 0.203 |
| Sickness score 4 (SSQ4) | 2.11 | 3.31 | 0 | 12 | 0.095 | 0.092 |
| # Practice sessions required | 1.37 | 0.63 | 1 | 3 | -0.160 | -0.132 |

\* *p*<0.05, \*\* *p*<0.01.

**3.4 Adjusting for covariates**

In order to examine the relationship between simulator performance and on-road performance, a generalized linear model was developed with total simulator score as a predictor of the on-road safety score, adjusting for simulator sickness at point 1. This model revealed that the total simulator score was still significantly associated with the on-road performance (B=0.114, SE=0.03, *p*<0.001), with an estimated odds ratio (OR) of 1.12 (95%CI = 1.05-1.19). This result indicates that for an increase in 1 point on the simulator score, there is an increase of 0.11 points on the On-Road Safety Rating score. Although simulator sickness was highest at point 2, this was not correlated with the Total simulator score (Table 3). Nevertheless, to investigate whether peak simulator sickness modified the relationship between simulator and on-road performance, the model was re-run with sickness at point 2 as a covariate. Although attenuated, the association between total simulator score and on-road score remained significant (B=0.071, SE=0.03, *p*<0.05; OR=1.07). A similar result was obtained using a model that included participant age and gender as covariates (B=0.095, SE=0.03, *p*<0.01; OR=1.10).

**3.5 Simulator Score as a predictor of failing on-road test**

The on-road safety rating scale (see page 17) ranges from 1 to 10, with scores below 6 representing fail and scores of 6 and above as pass. The scale was recoded as a binary variable (giving 11 fail and 28 pass), and logistic regression was used to examine the accuracy with which the simulator score predicts pass/fail on the on-road test. The results showed that the simulator score significantly predicted the binary on-road test result (B=0.16(0.09), *p*<0.05) with OR = 1.17(95%CI=1.02-1.34) indicating that a 1 point decrease in the Simulator score was associated with a 17% increase in the likelihood of failing the on-road test. The model classification rate corresponded to a sensitivity of 67% and specificity of 79% and a negative predictive value of 92.6% - indicating that if a participant performed well on the simulator assessment, there is a high degree of confidence that the individual will pass an on-road assessment. The sensitivity this virtual driving assessment, while not high, is better than that reported in previous studies that have used matched scoring systems across on-road and virtual mediums [[10](#_ENREF_10)]. Given that the virtual assessment was scored using only the simulator replay plug-in (lane position, brake/acceleration, gap selection and approach) and not the video play back of the participant’s head and eye movements (observation), it is possible that with the more complete scoring protocol, the sensitivity and specificity may improve.

**3.6 Aspects of simulator performance most associated with on-road safety**

A portion of participants performed very poorly on the self-navigation component of the simulator assessment. These participants typically became disoriented, followed the incorrect route and most also failed to complete the route after repeated exposure to the scenario. Becoming lost in the self-navigation scenario was significantly correlated with lower scores on the On-Road safety rating (*r*=0.359, *p*=0.025). The score obtained in the

Table 4: Correlation between scenario type, lane position, brake/acceleration, and on-road performance and simulator sickness.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Correlations (*r*) | | | | | |
|  | **On-Road** | **SSQ1** | **SSQ2** | **SSQ3** | **SSQ4** |  |
| Driver Instructed scenarios score | 0.307 | **-0.569\*\*** | 0.033 | 0.097 | 0.175 |  |
| DI scenarios: brake/acceleration | 0.274 | **-0.520\*\*** | -0.059 | 0.024 | 0.052 |  |
| DI scenarios: lane position | 0.239 | **-0.430\*\*** | 0.131 | 0.157 | 0.271 |  |
| Self-Navigation scenario score | **0.417\*** | -0.063 | -0.030 | 0.041 | 0.053 |  |
| SN scenario: brake/acceleration | **0.373\*** | -0.071 | 0.033 | 0.060 | 0.072 |  |
| SN scenario: lane position | **0.440\*** | -0.051 | -0.095 | 0.018 | 0.030 |  |
| SN scenario: getting lost | **0.359\*** | 0.075 | 0.060 | 0.137 | 0.132 |  |
| Lane Position (overall) | **0.429\*\*** | -0.289 | 0.015 | 0.105 | 0.180 |  |
| Brake/Acceleration (overall) | **0.391\*** | **-0.361\*** | -0.016 | 0.052 | 0.076 |  |

\* *p*<0.05, \*\* *p*<0.01. Note: SSQ refers to Simulator Sickness Questionnaire.

Self-Navigation scenario was also correlated with on-road performance (*r*=0.417, *p*=0.01). In contrast, the total score obtained in the Driver Instructed scenarios was not significantly correlated with on-road performance (*r*=0.307, *p*=0.065). Performance on the Driver Instructed scenarios was also associated with simulator sickness at point 1, whereas performance on the Self-Navigation scenario was not associated with simulator sickness (Table 4). This may be explained by the pattern of simulator sickness over the course of the experiment, and the fact that all Driver Instructed scenarios were presented earlier in the experiment relative to the Self-Navigation, and were hence more likely affected by simulator sickness.

Of the variables used in scoring simulator performance, brake/accelerator use and lane positioning were the most common errors, whereas the opportunity for errors in gap selection and approach were largely contingent on the context and route (e.g., only at intersections). Total score for brake/accelerator use was significantly correlated with on-road safety rating (*r*=0.391, *p*=0.017), as was the total score for lane positioning during (*r*=0.429, *p*=0.008). Brake/accelerator use was also correlated with simulator sickness levels, whereas this was not significant for lane position (Table 4).

In summary, the aspects of simulator performance that were most associated with on-road safety was lane positioning, brake/accelerator use, as well as self-navigation and becoming lost within the self-navigation scenario. However, it is likely that higher levels of simulator sickness in the early part of the experiment would have affected performance on the driver instructed scenario – potentially attenuating its relationship to on-road performance.

**3.7 Differences in cognitive demands between on-road and virtual driving assessments**

The correlation between the Total simulator score and cognitive measures of memory, attention, executive function, processing speed, and visuospatial function (see Table 5). Both on-road and simulator performance were associated with trail making task, MultiD, DriveSafe, and UFOV. Compared to on-road performance, simulator performance was associated more measures of visuospatial abilities including navigation in the Maze test and figure copying in the Benton Test. Simulator performance was also associated with verbal learning with repetition. In contrast, the on-road test was associated with global cognition (MMSE), knowledge of road law, decision making and visual acuity. Neither mode of driving assessment was associated with measures of language and verbal expression, working memory (i.e., Digits Backward), longer term memory and recall (delayed recall) or verbal response inhibition (Stroop colour word test).The results suggest that while both driving assessments were correlated with measures of complex attention (i.e., task switching) and learning, participants may be relying more on learning and

Table 5: Correlations between Cognitive Function and simulator and on-road tests

|  |  |  |
| --- | --- | --- |
|  | On-road safety rating | Total simulator score |
| Mini Mental Status Exam (Global cognition) | .339\* | -0.001 |
| Tests of Language and verbal expression | | |
| Letter fluency (COWAT) | -0.106 | 0.107 |
| Boston Naming Test | 0.18 | 0.307 |
| Tests of Verbal learning and memory | | |
| California Verbal Learning Test (trial 1) | 0.222 | .330\* |
| California Verbal Learning Test (trial 2) | 0.222 | .330\* |
| California Verbal Learning Test (trial 3) | .326\* | .402\* |
| California Verbal Learning Test (delayed recall) | 0.261 | 0.308 |
| Tests of Executive function§ | | |
| Trail Making Test B (sec) | -.414\*\* | -.349\* |
| Digits Span Backwards | -0.012 | 0.179 |
| Stroop Colour Word Test (interference score) | -0.27 | -0.173 |
| Game of Dice Task (frequency of safe choices) | .305\* | 0.172 |
| Reading the Mind in the Eyes Test | 0.225 | 0.317 |
| Tests of Visuospatial ability | | |
| Benton Copy Test | 0.241 | .388\* |
| Maze Test (sec) | -0.25 | -.334\* |
| Tests of Visual acuity | | |
| Letter Contrast (both eyes) | -0.103 | 0.187 |
| Visual Acuity (both eyes) | -.361\* | -0.301 |
| Tests of Road Knowledge | | |
| 14 Item Road Law Test | .349\* | 0.104 |

\* *p*<0.05, \*\* *p*<0.01. § Executive function refers to a complex set of skills including: task switching, working memory, response inhibition, decision making, planning and organisation.

visuospatial skills to do well on the simulator, whereas on-road driving also requires general cognition, knowledge of road rules, vision and decision making.

## Discussion

The study successfully developed a virtual driving assessment that was matched to a standardised on-road route, in terms of environment, route complexity, tasks required, and the scoring method and variables. The simulator was moderately well tolerated by older adults, with 11.3% of screened participants experiencing simulator sickness and withdrawing from the study. Scoring for the present report used only the data obtained using the replay plug-in for the simulator software (lane position, brake/acceleration, gap selection, and approach) but did not include data recorded by the video monitoring (observation).

**4.1 Validity of the virtual assessment**

Data from 39 participants indicated that performance on the simulator route was significantly associated with their safety rating on the on-road assessment, indicating that for an increase in 1 point on the simulator score (range 0 to 48), there was an increase of 0.11 points on the On-Road Safety Rating score (range 1 to 10). In terms of predicting whether a participant is likely to fail an on-road test, a 1 point decrease in the simulator score was associated with a 17% increase in the likelihood of failing the on-road test. Although the simulator assessment had a sensitivity of 67% and specificity of 79% and a negative predictive value of 92.6%, this is comparable to other simulator assessments reported in the literature [11].

Older adults’ performance on the virtual driving assessment was also significantly correlated with their performance on a range of driver screening batteries and other measures known to be linked to driving ability in older persons – including the Multi-D, the Useful Field of View test (UFOV), the DriveScore test, the OT-Dora Test, Trail Making Test, as well as participant age and weekly driving exposure. These results attest to the validity of the virtual driving assessment as a measure of driving ability in older adults, at least in relation to an on-road assessment on a standard route. Very few participants reported crashes in the previous 5 years, therefore our results did not find a significant association between self-reported crash rate and either virtual or on-road driving performance.

**4.2 Factors contributing to performance on the virtual assessment**

Previous studies have found that the ability to maintain lane position in a simulator is a sensitive predictor of impaired driving under conditions of distraction [[20](#_ENREF_20)], sleep deprivation and alcohol intake [[21](#_ENREF_21)]. Here, we found lane position as well as speed control in the virtual scenarios were significantly correlated with on-road driving safety in older drivers. Poor performance on the self-navigation component of the virtual assessment was also more strongly associated with on-road safety than performance on the driver-instructed components. This difference may reflect greater impact of simulator sickness on the driver-instructed scenarios than on the self-navigation scenario. However, it is also possible that the visual demands of searching for road signs during self-navigation would increase errors and hence its sensitivity to driving impairments. Previous work has shown this to be the case in young drivers [[22](#_ENREF_22)]. Future studies in which the order of driver instructed and self-navigation components is counterbalanced will help identify the relative contribution of sickness effects and self-navigation on performance.

**4.3 Limitations**

The virtual driving assessment has several limitations that are common to all simulated driving tests. This includes the fact that approximately 20% of older adults are unable to undergo the test due to susceptibility to motion sickness, and a further 11.3% experience simulator sickness during testing and cannot complete the testing. Although some participants may continue to feel subtle levels of discomfort during testing which likely impacted on their performance, analyses revealed that statistically adjusting for reported sickness levels did not abolish the association between simulator performance and on-road performance.

Although the virtual environment and scenarios were built to closely match the real on-road route, and generated a notable degree of familiarity in participants when doing the assessment, the lack of realistic steering wheel, indicator stalk, low frame rate and low fidelity steering and acceleration dynamics reduced the tests face validity.

* 1. **Strengths**

The present study has multiple strengths in terms of the design and outcomes. The virtual driving assessment developed here is the first, to our knowledge, to match the virtual model and scoring procedure to an on-road assessment, and to validate this in a healthy sample of older adults. The main strengths include being:

* + Predictive of on-road performance even in a small sample size, and after correcting for simulator sickness symptoms.
  + Sensitivity and specificity of the test in terms of identifying older adults likely to fail an on-road test, while not high, was comparable with other similar studies in the literature.
  + Stronger association with on-road performance than currently used off-road older driver screening tests.
  + Closely matched route and environment to the on-road driving assessment, as well as matched assessment criteria and method of scoring.
  + Low cost set-up with a scoring method that can be easily used and intuitive to clinicians.

Given the positive results reported here, it is likely that the predictive power of the virtual assessment will improve in future studies with minor modifications such as: inclusion of measures of observation and indication, increasing the sample size, and using a more realistic steering/pedal set-up. Despite the sample having better than average driving ability than the study from which they were recruited, the simulator performance was sufficiently varied and was sensitive to participants with poorer skills.

**4.5 Future directions**

Future studies will expand the sample size and include all variables in the scoring criteria to assess the validity of the virtual driving assessment. Future work will also examine the simulator’s test-retest reliability and interrater reliability in order to identify its utility as an outcome measure in studies of older driver training and intervention, and also its long-term predictive validity by examining its association with crash rates over time. Once the virtual assessment is identified as a valid and reliable measure of driving capacity in older adults, further studies can use it to examine the impact of medical conditions, medication and substance use under safe conditions.

We expect that the virtual driving assessment developed here will contribute to building capacity in two emerging areas of driving safety research within the Centre for Research on Ageing, Health and Wellbeing (CRAHW): (1) driver training interventions for older adults; (2) ‘brain training’ to maintain older drivers’ capacity to drive safely, and (3) substance and medication use and driving safety in older adults.

## References

1. Australian Transport Council., *National Road Safety Strategy*. 2011-2020.

2. Anstey, K.J., et al., *Different cognitive profiles for single compared with recurrent fallers without dementia.* Neuropsychology, 2009. **23**(4): p. 500-8.

3. Horswill, M.S., et al., *Older drivers' insight into their hazard perception ability.* Accid Anal Prev, 2011. **43**(6): p. 2121-7.

4. Mayhew, D.R., et al., *On-road and simulated driving: Concurrent and discriminant validation.* Journal of Safety Research, 2011. **42**: p. 267-275.

5. Anstey, K.J. and G.A. Smith, *Associations of biomarkers, cognition and self-reports of sensory function with self-reported driving behaviour and confidence.* Gerontology, 2003. **49**: p. 196.

6. Anstey, K.J., et al., *Cognitive, sensory and physical factors enabling driving safety in older adults.* Clin Psychol Rev, 2005. **25**(1): p. 45-65.

7. Lew, H.L., et al., *Predictive validity of driving-simulator assessments following traumatic brain injury: a preliminary study.* Brain Injury, 2005. **19**: p. 177-188.

8. Belanger, A., S. Gagnon, and S. Yamin, *Capturing the serial nature of older drivers’ responses towards challenging events: A simulator study.* Accident Analysis and Prevention, 2010. **809-817**.

9. Bedard, M., et al., *Predicting driving performance in older adults: We are not there yet!* Traffic Injury Prevention, 2008. **9**: p. 936-341.

10. Devos, H., et al., *Validating of driving simulation to assess on-road performance in Huntington Disease.* Proceedings of the Seventh Internation Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 2013: p. 241-247.

11. Lee, A.H., D. Cameron, and A.H. Lee, *Assessing the driving performance of older adult drivers: on-road versus simulated driving.* Accident Analysis and Prevention, 2003. **35**: p. 797-803.

12. Yan, X., et al., *Validating a driving simulator using surrogate safety measures.* Accident Analysis and Prevention, 2008. **40**: p. 274-288.

13. Wang, Y., et al., *The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing.* Ergonomics, 2010. **53**: p. 404-420.

14. Kennedy, R.S., et al., *Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness.* The International Journal of Aviation Psychology, 1993. **3**: p. 203-220.

15. Golding, J.F., *Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness.* Brain Research Bulletin, 1998. **47**: p. 507-516.

16. Brooks, J.O., et al., *Simulator sickness during driving simulation studies.* Accident Analysis and Prevention, 2010. **42**: p. 788-796.

17. Casutt, G., et al., *The relation between performance in on-road driving, cognitive screening and driving simulator in older healthy drivers.* Transport Research Part F, 2014. **22**: p. 232-244.

18. Meuleners, L. and M. Fraser, *A validation study of driving errors using a driving simulator.* Transportation Research Part F: Traffic Psychology and Behaviour, 2015. **29**: p. 14-21.

19. Golding, J.F., *Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness.* Brain Research Bulletin, 1998. **47**(5): p. 507-516.

20. Reed, M.P. and P.A. Green, *Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task.* Ergonomics, 1999. **42**(8): p. 1015-1037.

21. Arnedt, J.T., et al., *How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task?* Accident Analysis and Prevention, 2001. **337-344**.

22. Engstrom, J., E. Johansson, and J. Ostlund, *Effects of visual and cognitive load in real and simulated motorway driving.* Transportation Research Part F: Traffic Psychology and Behaviour, 2005. **8**: p. 97-120.