TRAFFIC INJURY RESEARCH FOUNDATION

A META-ANALYSIS OF COGNITIVE SCREENING TOOLS FOR DRIVERS AGED 80 AND OVER

TIRF

The knowledge source for safe driving

THE TRAFFIC INJURY RESEARCH FOUNDATION

The mission of the Traffic Injury Research Foundation (TIRF) is to reduce traffic-related deaths and injuries. TIRF is a national, independent, charitable road safety research institute. Since its inception in 1964, TIRF has become internationally recognized for its accomplishments in a wide range of subject areas related to identifying the causes of road crashes and developing programs and policies to address them effectively.

Disclaimer

The Ontario Ministry of Transportation (MTO) has provided funding for the report: A Meta-Analysis of Cognitive Screening Tools for Drivers Aged 80 and Over. MTO does NOT warrant the accuracy, validity, completeness or currency of the report. Funding of the report is NOT to be construed as an endorsement of the contents of the report, the Traffic Injury Research Foundation or any other person or entity. Use of this report is completely at one's own discretion and risk.

Traffic Injury Research Foundation
171 Nepean Street, Suite 200
Ottawa, Ontario K2P 0B4
Ph: (613) 238-5235
Fax: (613) 238-5292
Email: tirf@tirf.ca
Website: www.tirf.ca

MAY 2014
Traffic Injury Research Foundation

ISBN: 978-1-926857-53-4

A META-ANALYSIS OF COGNITIVE SCREENING TOOLS FOR DRIVERS AGED 80 AND OVER

Prepared by:

Ward Vanlaar*, Anna McKiernan*, Heather McAteer*, Robyn Robertson*, Dan Mayhew*,

David Carr**, Steve Brown*, Erin Holmes*

*Traffic Injury Research Foundation

**Washington University at St. Louis

Traffic Injury Research Foundation

171 Nepean St. Suite 200

Ottawa, ON K2P 0B4

MAY 2014

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The Traffic Injury Research Foundation (TIRF) would like to gratefully acknowledge the assistance of the following individual who shared his expertise in the development of this report and who reviewed and commented on earlier drafts of this report. His insights and statistical expertise allowed us to develop a well-rounded and thoroughly reviewed meta-analysis.

Chengjie Xiong, Ph.D.

Associate Professor of Biostatistics

Department of Biostatistics

School of Medicine

Washington University

EXECUTIVE SUMMARY

Canada's population is aging and seniors represent the fastest growing population group in Canada. Based on population projections from Statistics Canada, Human Resources and Skills Development Canada reported that in 2011 there was an estimated 5 million Canadians over the age of 65 (2012). This population is on the rise and the number of seniors in Canada will reach 10.4 million by 2036. Using today's licencing rates in Canada, it can be expected that over 4.6 million Canadians aged 65 or older will hold valid driver licences after 2021. This number will rise to 6 million by 2031 (Robertson and Vanlaar 2008).

Today there is no single tool that can accurately identify an unfit driver with absolute certainty. In light of the expected increase in elderly drivers with cognitive impairments and dementia, efficient and effective screening for such impairment will become more important. Therefore, it is crucial to establish the most appropriate methods to assess cognitive impairment among elderly drivers.

The objective of this project is to gather and analyze information to inform the selection of a cognitive screening tool or tools that can be used during Group Education Sessions (GESs) in Ontario. Ultimately, this tool can be used during GESs to discriminate between elderly drivers who require more extensive testing (i.e., an on-road examination), from elderly drivers who are fit to continue driving without further examination.

In order to accomplish this objective, a meta-analysis of the literature on cognitive screening tools has been conducted. The first step in the meta-analysis is the review of literature, which included 446 articles related to the study topic. From these, particular articles were selected based on their relevance, i.e., whether they evaluated a cognitive tool. The number of tools selected from these evaluation studies totaled 42. These articles and their relevant tools were coded and entered into a database. Criteria were developed to ensure the tools would be applicable to the parameters of the GES:

- **Duration of tool** – a variable was created to select tools of different duration; we selected those tools that can be administered in less than 10 minutes;
- **Administration of tool** – a variable was created to select between tools that are not limited to individual (one-on-one) administration versus tools that can only be administered individually (e.g., they require the participant to recall words to the administrator or the test is timed); we selected those studies of tools that do not require individual administration and that can be administered in a group setting;

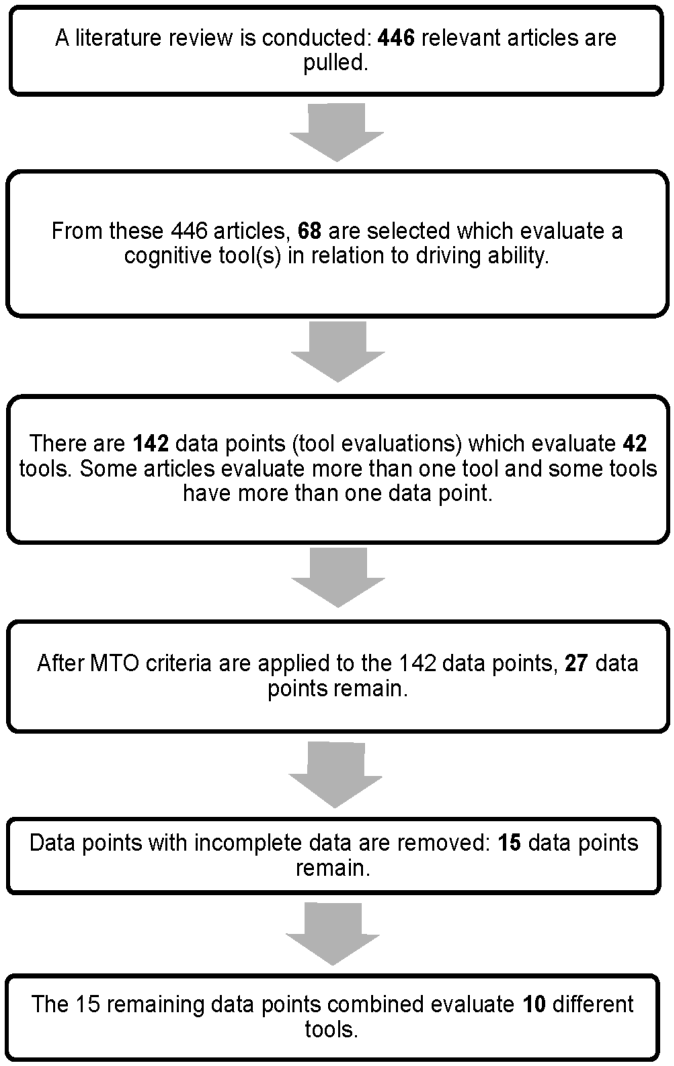
- > **Feasibility of administration** – a variable was created to distinguish between tools according to their ease of administration and scoring; we used it to select those studies of tools that are either easy to administer (i.e., simple instruction) or easy to score (i.e., minimal calculation needed to tally the score) while tools that were difficult to administer and difficult to score were not included;
- > **Computer** – a variable was created to distinguish between tools that only require pencil-and-paper versus tools that require an expensive piece of hardware like a computer or simulator; only tools that do not require an expensive piece of hardware, but that can be administered with pencil-and-paper were included; and,
- > **Expertise** – a variable was created to distinguish between tools that are easily administered without specialized training versus tools that do require such training; only tools that do not require specialized training were included.

Once applied, these criteria narrowed down the list to ten applicable tools:

1. Maze Task;
2. Rey-Osterrieth Complex Figure Test;
3. Clock Drawing Test;
4. Single Letter Cancellation Test;
5. Double Letter Cancellation Test;
6. Traffic Sign Recognition Test (TSRT);
7. Charron Test;
8. Visual Form Discrimination Test;
9. Wechsler Digit Symbol Substitution Test; and,
10. Eight Item Informant Questionnaire.

The articles containing these 10 tools were analyzed in order to obtain the data points for the meta-analysis.

Below is a flow chart describing each step in the selection process of data points:



The random-effects meta-analysis provided evidence suggesting that cognitive screening tools that meet MTO's logistical requirements for GES can be used to predict driving performance. A small to medium-sized, significant pooled effect of 1.940 was found, meaning that on average, when cognitive screening tools predict a driver is unsafe, there is a 94% greater chance that this driver will exhibit unsafe driving behaviour, rather than safe driving behaviour (or, alternatively, if the cognitive screening tools predict that a driver is safe, on average, it is 94% more likely that this driver will exhibit safe driving behaviour, rather than unsafe driving behaviour). Note that unsafe driving behaviour refers to unsafe performance during a road test, a simulator driving test, or crashing as evidenced by state-reported crashes (it follows that safe driving behaviour refers to safe performance during a road test, a simulator driving test, or not crashing).

Results from the meta-analysis can be used to rank order cognitive screening tools in an effort to identify better-performing tools. This exercise made clear that the results from this meta-analysis alone are not sufficient for the selection of one particular tool. It is recommended meta-analysis results are used in combination with other information such as sensitivity, specificity, and area under the Receiver Operator Characteristics (ROC) curve. However, such information is not available for all the tools included in this meta-analysis. Partial information was available in six studies. Using this partial information, and depending on one's preferences regarding concerns for the public versus the individual (based on information regarding sensitivity versus specificity), DSST or Clock Drawing can be considered as preferred cognitive screening tools.

In sum, the meta-analysis shows there is fairly robust evidence suggesting cognitive screening tools have predictive value of driving behaviour. The effect is significant and low to medium-sized, and it differs, not only across different ways of measuring (un)safe driving behaviour, but likely also across cognitive screening tools themselves. It was not possible, however, to study in more detail any effect of the screening tools due to our small sample size (15 data points to evaluate 10 different screening tools). The results of this meta-analysis can be used in combination with other sources of knowledge to select a cognitive screening tool.

Ultimately, this selection will depend on choices that are more of a practical nature (e.g., Can the tool be efficiently and practically incorporated into the GES curriculum?) and political nature (e.g., What is the trade-off between concerns for the public versus concerns for the individual?). Based on this investigation, and the information at hand to date, we recommend that any of the ten tools included in this meta-analysis can be considered for use in a GES setting, acknowledging that not all of them are equally effective in predicting unsafe or safe driving behaviour according to the rank ordering results of this meta-analysis.

1. INTRODUCTION

1.1 Background

Canada's population is aging and seniors represent the fastest growing population group in Canada. Based on population projections from Statistics Canada, Human Resources and Skills Development Canada reported that in 2011 there was an estimated 5 million Canadians over the age of 65 (2012). This population is on the rise and the number of seniors in Canada will reach 10.4 million by 2036. Using today's licencing rates in Canada, it can be expected that over 4.6 million Canadians aged 65 or older will hold valid driver licences after 2021. This number will rise to 6 million by 2031 (Robertson and Vanlaar 2008).

With an increase in this age population on the roads, it is important to understand what physical, cognitive, and mental issues these drivers may face. In particular, older drivers are at an increased risk of suffering from late-life cognitive impairment and dementia (Carr et al. 2006; Carr and Ott 2010; Dobbs 2005). These diseases are concerning as they impair the ability to drive. In this respect, the number of drivers with dementia is on the rise: between 1997 and 2005, there were 210,000 people in Ontario diagnosed with dementia, 40,000 of which held active driver licences. More worrisome is that 9,000 of these drivers have been in car crashes, a third of which had taken psychotropic drugs (Rapoport et al. 2008). It has been predicted by a Queen's University study that the number of drivers in Ontario with dementia will double by 2028 and reach 100,000.

Today there is no single tool that can accurately identify an unfit driver with absolute certainty. In light of the expected increase in elderly drivers with cognitive impairments and dementia, efficient and effective screening for such impairment will become more important. Therefore, it is crucial to establish the most appropriate methods to assess cognitive impairment among elderly drivers.

Presently, the Ontario Ministry of Transportation (MTO) has a system in place to assess elderly drivers with the goal of differentiating drivers who may be at greater risk of causing a crash from those who are not. Three driving programs have been developed to meet this end. First, the Mandatory Medical Reporting Program requires that every person, 16 years of age or older, who suffers from any condition which could affect their driving is reported to the Ministry by their physician or optometrist. Police officers may do the same if they discover an unsafe driver through collision or other investigations. Second, according to the 70 and Older Collision Program, any driver aged 70 or older must pass a vision, knowledge, and G2 test (level two road test in the graduated licence program of Ontario) in order to maintain their licence if they are involved in a collision or convicted of a traffic offence related to a collision. Third, the 80 and

Above Licence Renewal Program consists of a two-year renewal cycle along with the use of a vision test, driver record review, and a Group Education Session (GES) for all drivers aged 80 and over. If necessary, a road test is included if the driver appears to pose a risk to road safety. The GESs are delivered at 200 approved locations across the province. Each session is approximately 90 minutes long and is designed to accommodate a maximum of 15 participants per session. The sessions educate the participants about important issues such as high-risk driving situations for their age group, physical challenges related to aging, and ways to reduce collision risk.

The challenge that MTO faces is the lack of a reliable way to assess whether a driver suffers from a cognitive impairment. Although obvious cognitive impairment may be easy to identify in-person, mild or discrete impairments are still detrimental to the senior's ability to drive and much more difficult to detect. In order to address this problem, it would be helpful to incorporate an assessment element into the current GES framework which will accurately indicate driver impairments. This screening tool must be easily understood and applied by non-medical professionals who interact with senior drivers during GESs. It must be inexpensive and easily administered to a group of 15 seniors and simple for drivers to complete. In order to identify tools which meet these criteria and can potentially be used in a GES for drivers aged 80 and over, a meta-analysis review of cognitive screening tools was conducted.

1.2 Objectives

The objective of this project is to gather and analyze information to inform the selection of a cognitive screening tool that can be used during the GES in Ontario. Ultimately, this tool can be used during GESs to discriminate between elderly drivers who require more extensive testing (i.e., an on-road examination), from elderly drivers who are fit to continue driving without further examination. Such a tool can help to make informed decisions about which senior drivers should be referred to take a road test. As Mathias and Lucas (2009) argue in their meta-analysis, "relicencing decisions should not be based on cognitive tests alone, but they do have the potential to identify individuals who require on-road testing before licencing decisions are made"(p.639). In order to accomplish this objective, a new meta-analysis of the literature on cognitive screening tools has been conducted with a focus on tools that are feasible and easily administered in the license renewal setting.

1.3 The structure of this report

This report summarizes the findings from the meta-analysis along with background information on the issue of elderly driver assessment. It includes potential solutions and future steps that can be incorporated into MTO's current procedures.

In the literature review section of this report (2. Literature Review), topics such as crash risk, crash type, exposure and diseases common to the older driver are explored. Included in this is the current debate surrounding the accuracy of the assessment process. Also, in section 3, entitled "Cognitive Screening Tools," an overview is provided of the tools that were included in this study and their applicability to the needs of the current GESs. A distinction is made between tools that were included in the meta-analysis and those that were excluded; the former are described in more detail.

The methodology section (4. Methods) describes the process that was undertaken in order to conduct this meta-analysis including the keyword search, extraction of appropriate articles, selection and rejection of articles, the coding of the studies, and the statistical analysis itself. Methodology is followed by the results section (5. Results). This results section does not only provide an overview of the results from a random-effects meta-analysis and regression, but also of diagnostics such as tests for publication bias and a solution for multiple outcomes in the same study. This results section is followed by a discussion of the results, as well as limitations (6. Discussion).

In the conclusion, section 7, a summary of results including the overall ranking of the cognitive tools included in the meta-analysis is presented.

2. LITERATURE REVIEW

2.1 Effects of aging on driving

There is considerable debate regarding whether aged-based screening is both relevant and beneficial to elderly drivers and all other road users. Aging is a complex process which is different for each individual. Furthermore, there are several different diseases, ailments, and conditions experienced by the elderly, both physically and mentally, which will affect their driving. These factors lead to the crash rate and crash type unique to the elderly driving population. Because of this, it is important to explore the most effective manner for evaluating elderly drivers. In this section, each of these topics is discussed further with respect to the issues of elderly drivers and age-based screening.

As previously discussed, the population, in Canada, as in most Western cultures, is aging and the elderly will represent a greater proportion of the population in the coming decades. The United States (U.S.) is facing a similar population boost: by 2030, 70 million people in the U.S. will be older than 65 (twice as many as today), which translates to a representation in the driving population of 25%, corresponding to a 14% increase (Wilson 2007).

As life expectancy increases considerably in our modern society, it is crucial to understand the ramifications these changes will have on road safety. In Canada, in 2006, seniors aged 65 and older accounted for the second largest proportion of road deaths at 16%, meaning 462 road fatalities. The same group experienced 15,545 or 7.8% of road injuries (Transport Canada 2007). Canadian seniors have the second highest motor vehicle death rate among licenced drivers, with an average of 15.7 deaths per 100,000 licenced drivers, compared to 9.6 deaths for drivers aged 45-54. In 2009, there were 494,756 Ontario licenced drivers aged 75 and older, 12,044 of which had collisions that year. This translates to 243 collisions per 10,000 elderly licenced drivers (ORSAR 2009).

There is a specific crash type that is most common for elderly drivers, which is the angled collision. The rate of angled collisions increases significantly after age 70, while the rate of all other crash types, including rear-end and head-on crashes actually decrease for this age group (Romoser and Fisher 2009). These crashes most commonly occur at intersections, especially when the elderly driver is required to make a left turn (Bryer 2000; Mayhew et al. 2006). This may be because elderly drivers have difficulty assessing space for entering traffic flow and experience failures in areas of perception and diagnosis. This means that when an elderly driver is evaluating how much space they need for a left turn across oncoming traffic, they are more likely to misjudge or miscalculate, compared to other age groups, causing a crash. Situations with an increased amount of stimuli (several different lanes of traffic flowing in different directions, pedestrians, traffic signals) cause difficulty for the elderly. In this respect, one study estimated that 19% of elderly driver crashes are due to their cognitive abilities being overwhelmed (Van Elslande and Fleury 2000).

Of course, not all elderly drivers suffer from impaired faculties behind the wheel, and often their gained experience makes them more careful and safer drivers than young drivers. For instance, when comparing the age groups of 20-40, 65-69, and 70+, Lerner (1993) found that all three groups needed the same amount of perception reaction time (two seconds) when coming to a stop at controlled intersections. This means that it does not take older drivers longer to react to traffic signals than younger drivers. Instead, intersection problems may be related to perception failures such as, dividing their attention between stimuli and comprehending traffic controls. In a similar study, Lerner et al. (1993) found that older drivers need longer gaps (spaces between vehicles) in order to perform a maneuver. For example, an older driver would estimate that a longer period of time is needed in order to switch into another lane, whereas a younger driver would estimate less time is needed for this maneuver. From these studies it can be concluded that older drivers do not necessarily require more time to initiate movement than younger drivers but instead demand longer intervals of time before being willing to attempt a maneuver on the road (Lerner 1993).

Notable declines that come with aging can impair an elderly driver, making common driving maneuvers that they have performed for decades, such as turning left in an intersection, much more challenging. For example, sight diminishes with age. The elderly driver needs eight times more light than a 20 year-old driver in order to see properly, making nighttime driving much more difficult (Lampman 2002). Similarly, loss of peripheral vision occurs with age, narrowing an elderly driver's field of view, and increasing their crash rate up to 30% (Voelker 1999). Older adults also have a diminished capacity for processing information, especially in their periphery due to sight loss, compared to younger drivers. Other weaknesses due to age are decline in short-term memory and speed of processing – skills which are needed to efficiently survey the road and appropriately respond to hazards. For instance, age-related cognitive decline might lead to decreases in secondary looks (i.e., checking more than once for traffic or a free space for merging) and situational awareness (i.e., noticing pedestrians or other road users surrounding the driver), especially during turns (Romoser and Fisher 2009).

Regardless of the correlation between age and age-related impairments, aging is associated with increasing prevalence of various medical conditions which can have a negative effect on crash risk. Functional impairments experienced behind the wheel are the detrimental effects of cognitive impairments (the most common being dementia) and physical impairments such as visual impairment and arthritis (Lampman 2002). Driving impairment due to disease may be discreet in the beginning, being caused by conditions such as "mild cognitive impairment" in which individuals have greater memory problems compared to others of the same age, but whose memory problems do not significantly affect their functioning (Howe 2007). It is more difficult to detect drivers with such an affliction and it may not be crucial to do so in the early stages. Nevertheless, a significant number of those with mild cognitive impairment will develop dementia within five years (Gauthier et al. 2006). Dementia affects the brain and the ability to think, remember, and speak. It impairs a person's logical reasoning, memory, and their language skills. Anywhere from 3% to 19% of adults over the age of 65 have dementia and the disease increases their crash risk 2.5 times. Drivers with dementia are more likely to cause a crash due to carelessness or unsafe driving maneuvers such as improper turning or following too close (Cooper et al. 1993).

Of greater concern is that those demented drivers who have crashed continue to drive. Cooper et al. (1993) found that of 43 crash involved drivers who had dementia, 36 continued to drive after their initial crash. Of those who continued to drive, 14 had at least one more crash before ceasing to drive. Cognitive impairment that is not due to a traumatic event (such as a stroke) is often discovered by doctors in the later stages when it is more severe and therefore apparent. This means that drivers may still operate their vehicle for a fairly long time with impairment such as dementia (Cooper et al. 1993). This is concerning because the crash rate can increase on average 2-4 times compared to healthy controls (Carr and Ott 2010), further highlighting the need for an accurate screening tool to indicate these issues in drivers.

Impairments may also come from diseases or conditions that affect the driver in other domains, for instance, visual impairment such as cataracts, glaucoma, or macular degeneration (Lampman 2002). A study evaluating the impairments of elderly drivers injured in crashes found that those who had crashed were four times more likely to have glaucoma when compared to the control group which consisted of elderly drivers who had not had crashes (McGwin et al. 1997). Physical impairment may also result from conditions such as arthritis. Arthritis is when joints in the body begin to degenerate causing them to become red, swollen and painful. This pain and resulting loss of strength can cause the driver to become tired and distracted while driving and decreases the range of motion they have to properly survey the road (Lampman 2002).

In order to manage these medical issues, the elderly are prescribed various medications, and those that are sedating can negatively affect their ability to drive. Older drivers generally take more medications than younger drivers and are more susceptible to the side effects of these medications on their central nervous systems (Ray 1992). Drugs such as antihistamines, antidepressants, benzodiazepines, and opioids are associated with increased crash risk and driving impairment such as drowsiness, loss of coordination, and dizziness (Wang and Carr 2004).

2.2 Debates regarding elderly drivers

A person's age alone cannot predict their ability to drive. Nevertheless, due to a potentially increased crash risk, crash types, ailments of aging and common elderly diseases, properly screening elderly drivers for signs of physical or cognitive impairments appears to be a logical and valid approach. Despite its face validity, screening of senior drivers has caused considerable debate. Two of the main debates regarding elderly drivers and assessment are discussed here in more detail: 1) what are the ethical implications of removing someone's licence based on their potential for crashing? And, 2) do age-based assessments have positive benefits for road safety? The first debate highlights the difficulties faced by physicians in maintaining a balance between legal obligations and the morality of confidentiality and how licence revocation affects an elderly driver. The second debate explores the myths behind the perception of elderly driving and summarizes several studies claiming that age-based assessment does not necessarily lead to increased traffic safety.

Ethical implications. Physicians can use medical examinations as an opportunity to uncover any potential dangers the driver may pose on the road. In theory this makes sense, because a physician is highly trained and can fully determine the severity of impairment. Conversely, the patient-doctor relationship is one of trust and confidentiality, and reporting a potentially impaired patient to the licencing authority may breach such a relationship.

In reality, most physicians are actually uncertain about their legal obligation to report unsafe drivers to the appropriate authorities (Carr and Ott 2010). The process of reporting is different for each jurisdiction, some have mandatory reporting of drivers that pose a risk and others have voluntary reporting (Kelly et al. 1999). For instance, in Canada, seven provinces, including Ontario, have mandatory physician reporting of medically unfit drivers to their Ministry of Transportation and three provinces have discretionary reporting (Ginzberg and Minichiello 2011).

In order to report an unsafe driver, doctors need to involve the patient, obtain their consent, and complete the appropriate administrative steps. For instance, based on testimony from a Michigan doctor, a doctor cannot turn over an unsafe driver to the jurisdiction without obtaining the patient's driver licence number. In order to obtain the licence number, the doctor must ask the patient for their driver's licence, which prompts the patient to ask why the doctor needs their licence number. Once the doctor explains that they have to report their impairment to the jurisdiction resulting in their licence being revoked, the patient is much less likely to hand over their licence, preventing the doctor from filing the report. Administrative steps such as these make it next to impossible to report unsafe drivers to the jurisdiction and very few drivers are willing to consent to having their licence removed (Wilson 2007).

If a driver does heed the warnings of their healthcare professionals to cease driving, there can be a great benefit to road safety. A doctor's warning to unfit drivers is an actual preventative medical intervention intended to reduce the trauma of motor vehicle crashes. A recent study found a 45% decrease in the rate of annual crashes per 1,000 patients who were given a warning by their doctors not to drive (Redelmeier et al. 2012). This means that, if patients were willing to hand over their licences once doctors warn them they are no longer safe drivers, almost half of crashes among this population that are due to certain illnesses and conditions could be prevented.

However, some drivers do not voluntarily cease to drive because by doing so they give up a considerable amount of freedom. When told they are no longer able to drive, the elderly feel shocked, sad, frustrated, surprised, powerless, and devastated. Often they are frustrated that they are expected to continue their lifestyle without driving and most likely disagree with their doctor's recommendation (Byszewski et al. 2010). Without the ability to drive, it is common to become isolated and depressed. Licence removal is not only a loss of mobility but can also be experienced as a loss of identity, independence, and status (Shapiro

1999). Of the elderly population, 88% rely on their own personal vehicle for transportation. If the driver's licence is taken away, a significant number of people are left reliant on family or friends for transportation, soon causing the person to feel like a burden on the ones they love (Martinez 1995). Driving cessation has also been associated with a decrease in social integration (Mezuk and Rebok 2008), decreased out-of-home activities (Marottoli et al. 2000), an increase in depressive symptoms (Fonda et al. 2001) and an increased risk of nursing home placement (Freeman et al. 2006).

Age-based assessments. Instead of leaving the screening process solely up to health care professionals, it is also a possibility to screen for these impairments at the point of licence renewal for elderly drivers. At present, there is still debate regarding whether or not such screening is effective and results in increased road safety. The problem lies in the fact that is it impossible to estimate the individual risk of a driver. Instead, you can only assign a person to a group which is considered to have a higher level of risk based on statistical information about the crash likelihood of this group. Just because a driver is "elderly" and the driving group labelled "elderly" has a high crash risk does not mean this individual elderly driver will cause a crash (Meng 2010). A study conducted found that when "extrapolating from values in one large-scale prospective cohort study, the cognitive test that most strongly predicted future crashes would, if used as a screening tool, potentially prevent six crashes per 1000 people over 65 screened, but at the price of stopping the driving of 121 people who would not have had a crash" (Martin et al. 2011, p.1).

In this regard, the European Federation of Psychologists Association (EFPA) argues that age-based screening does not produce any benefits and that older drivers are generally the safest group of drivers on the road. Although they acknowledge that the elderly are more frail, and are more likely to be injured or killed in crashes (Evans 2001), they do not equate this to the elderly having a higher crash rate – it only appears as such due to their greater likelihood of injury. Crashes are much more likely to be reported to police if there is a severe injury or death. Also, because crash statistics come from police records, and because the elderly have a higher chance of being injured or killed in a crash, they have a higher number of crashes represented in these statistics. Even though younger drivers may have the same number of crashes as the elderly, these crashes are less likely to make it into police records because less of them involve injury or death. Due to this phenomenon, EFPA argues that the crash rates of the elderly are inflated (2010). Li et al. (2003) is in agreement, estimating that 60-95% of increased crash risk of the elderly is due to this "frailty bias."

The low mileage bias also argues that the elderly are unfairly targeted as a higher risk driving population. This bias stipulates that the elderly exhibit different driving patterns after retirement, which in turn exposes them to situations in which they are more likely to crash. After retirement, a driver will be driving less and most of their driving will be in urban areas, which have higher crash rates due to greater volume of road users and competing stimuli. This means that their crash rates appear higher in comparison to younger drivers who may do more driving and less high-risk driving (on the highway, for example). When comparing older drivers and middle-aged drivers with the same yearly mileage the age difference in crash rates disappears (Langford et al. 2006). This means that it is the quantitative and qualitative differences in driving exposure that increases crash rates, not the age of the driver.

Several studies show that age-based screening for fitness to drive does not produce any safety benefits. For instance, crash rates were compared across six Australian states: five with varying screening processes and one state without age-based screening. It was found that crash rates were the lowest for the state without age-based screening (Langford et al. 2004). A similar study conducted in Denmark found no increase in road safety after implementation of the age-based screening and actually found a detrimental effect. When comparing the crash rates of before and after mandatory age-based screening, no statistically significant difference in the number of older drivers involved in fatal crashes was found. There was a significant increase in unprotected older road users who were killed between the two periods of observation. This could mean that the age-based screening caused older drivers to stop using their vehicles and instead use unprotected, less safe modes of transportation, for example, walking (Siren and Meng 2012). However, based on a different interpretation of their findings, one could also argue that the true cause of an increase in fatalities is the lack of safe, alternative transportation modes for seniors, rather than the screening itself. This speaks to the importance of not implementing age-based screening in a vacuum, but rather as one aspect of a more holistic approach.

While research has shown that there are biases that likely cause the crash risk of senior drivers to be overestimated, there may also be biases at work that cause this risk to be underestimated. Langford et al. (2004) explore this by making an important point with respect to the type of data used to analyze crash statistics. If a study's results are based on per-population or per-driver crash rates, it does not take into account differences in driving activity. There may be many elderly drivers that hold valid driver licences but rarely, if ever, drive. The number of valid licences held by drivers who do not drive would likely be higher in jurisdictions that do not have mandatory age-based assessments. Without the assessment, there is no need for the elderly driver to submit their licence, allowing them to maintain the status of a valid driver. If crash statistics are calculated using valid licences (per-driver), it is not unlikely that those jurisdictions without assessment processes will have more licenced drivers due to all the inactive drivers, and this would result in lower crash rate calculations (Langford et al. 2004).

It is clear that, to date, the research community has not yet reached a consensus as to whether age-based screening increases traffic safety in our society or not. In the absence of such a consensus, it may be prudent to err on the side of caution and adopt some form of screening in combination with other measures, such as ensuring access to safe, alternative modes of transportation. This makes the need for an effective and efficient screening tool pressing.

2.3 Results from other reviews

Several studies have analyzed the validity of the elderly driver assessment process, some of which specifically explore the literature surrounding cognitive assessment tools. Mathias and Lucas (2009) conducted a meta-analysis of cognitive predictors of unsafe driving in older drivers. The authors emphasize that although the on-road driving assessment is considered the "gold-standard" for assessing driving competence, on-road is not always feasible on a large scale as a result of financial constraints and time constraints. Instead, cognitive tests should be valued for the additional useful information they can provide, such as the identification or indication of the presence of cognitive deficits that cannot be ascertained from an on-road examination alone.

That being said, there is currently no agreement about which cognitive abilities should be assessed with respect to driving performance and which tests should be used to assess these abilities (Anstey et al. 2005). Several different assessments have been evaluated for consistency in road performance prediction but findings and conclusions vary. For instance, there is research available showing that a number of commonly used cognitive tests, such as Trail Making and the Rey Complex Figure Task, are associated with on-road driving ability (De Raedt and Ponjaert-Kristoffersen 2000; Schanke and Sundet 2000; Whelihan et al. 2005). As well, Ball et al. (2005) developed and evaluated a battery of tests which was designed to assess functionality in elderly drivers with the purpose of detecting their crash risk. Drivers completed the battery at the time of their licence renewal and again five years later. The crash data between the two assessments was evaluated in order to measure the predictability of the cognitive measures for at-fault crashes of the participants. Results showed that the measures were effective predictors of crash involvement. The Useful Field of View Test (UFVT), Trail Making B, Delayed Recall, and the Motor Free Visual Perception Test (MFVPT) were found to be associated with increased crash risk (Ball et al. 2005). Similarly, the Mathias and Lucas (2009) meta-analysis found that the best predictors of on-road driving is the Ergovision test and the UFVT, as well as tests such as Paper Folding and Dot Counting. Similarly, Trail Making Tests A and B, Stroop Test, and the Automated Psychological Test were good predictors of driving problems (Mathias and Lucas 2009).

Issues regarding cognitive assessment were brought up in the review conducted by Reger et al. (2004). They warn that caution must be applied when neuropsychological testing forms the basis of driving assessment. With respect to which area should be tested that will most effectively predict driving behaviour, the authors suggest that tests of visuospatial skill may be most helpful in identifying at-risk drivers. They found that only visuospatial skills and attention-concentration assessments correlate with driving. The benefits of cognitive assessments that do not include an on-road aspect are the ability of the examiner to control any important variables and to standardize the procedure. An on-road test has several different variables that will alter the experience for each driver. For instance, the amount of traffic is different depending on the time of day, as is the weather. The examiner cannot control these variables in order to ensure each test taker is getting a standardized experience. Cognitive assessments allow for a "purer measurement of skill" important to driving (Reger et al. 2004).

More research is needed in order to develop the "gold standard" of cognitive assessments. Ideally, the assessments would target specific skills which best reflect driving ability, resulting in prediction of crash risk. Until this has been accomplished in the field, researchers, policymakers, and licencing agencies can only assess drivers to the best of their ability within the confines of their environment and resource range.

3. COGNITIVE SCREENING TOOLS

The following is a description of the cognitive screening tools which were analyzed for the meta-analysis. The first list of tools (described in greater detail in the following section), contains the tools that were included and evaluated for the meta-analysis. The tools following these descriptions are the ones that were excluded from the analysis because they did not meet the necessary criteria; they are described more briefly. Table 1 outlines the inclusion criteria that were applied when selecting tools to include in the meta-analysis. Inclusion criteria will be discussed in greater detail in Methods. Table 2 illustrates how each tool was selected with respect to each criterion. A similar table can be found in Appendix C which outlines how each excluded tool does not meet the criteria.

Each tool description provides an overview of the tool itself, how it is conducted and its composition, followed by the scoring that will be undertaken by the administrator, the duration of the tool (how long it takes to administer and complete), the technological requirements of the tool (computer, other apparatuses), and the level of expertise required to administer the tool. These subsections mirror the criteria that were used to code each tool. Note that where such information is not provided in relation to a tool, it is because this information was not available.

It warrants mentioning that the tools discussed in this project are those for which evaluative studies were available. It is acknowledged that there may be other tools available which could have been considered for inclusion, but if an evaluation study for these tools was not available in our literature review, they were not included in the descriptions of either the included tools or the excluded tools.

Table 1: Inclusion Criteria

Criteria	Description	Questions considered when coding
Duration	The duration of the tool cannot exceed 10 minutes	How long does it take to administer and complete the test?
Administration	The tool must be administrable to a group of about 15 participants	Is it possible to administer this test to a group or is it meant only for one-on-one administration?
Feasibility	The tool must be simple to instruct and easy to score (i.e., calculation of score is not complex)	How is the test administered (i.e., what instructions are given to participants)? How is it scored (i.e., complex calculations)?
Computer	The tool should not require a computer or any technology to administer	Does this tool need a computer for administration? What instruments are needed to administer this tool?
Expertise	The tool should not require any special expertise or training to administer and score	Is there any special expertise or training required in order to administer and score this tool?

Finally, all tools discussed below, included and excluded, have strictly speaking been designed for individual administration, and they typically are administered individually. None of these tools are administered in a group setting, but this does not necessarily mean they are not applicable in a GES setting.

3.1 Cognitive screening tools included in the meta-analysis

Rey-Osterrieth Complex Figure Test

Overview: The Rey-Osterrieth Complex Figure Test is a neuropsychological assessment designed to evaluate visual perception and long-term visual memory capacities. The participants are presented with a standard complex figure containing a variety of visual elements. Participants are instructed to copy the figure. Then, after two or three minutes, participants may also be instructed to reproduce the figure from memory. This second drawing is referred to as the 'immediate' drawing. Other versions of this test ask the participant to reproduce the drawing a third time, after a 20-30 minute time delay. These versions were not included as a time delay of this length is impractical in a GES setting.

Scoring: An 18-point scoring system is used, corresponding to the 18 different elements within the complex figure. To score either the copied figure or memory-reproduction, administrators consider whether each of the 18 elements has been copied or reproduced accurately. A point is awarded if the participant has correctly drawn an element in the figure; a point is withheld if the element is not drawn correctly. More specifically, administrators have the option of giving two points for each element if it is correct and placed properly; one point if the element is correct but placed poorly; one point if the element is distorted or incomplete, but recognizable, and placed properly; one-half point if the element is distorted or incomplete, but recognizable, and placed poorly; or zero points if the element is absent or not recognizable. Examples of the phrasing of elements include "cross attached to lower center," "four parallel lines within quadrant two, upper left," and "vertical midline."

Duration: There is no set time limit for either the copy or immediate recall portion of the tests. However, the copy and immediate versions of the test typically require fewer than ten minutes to complete.

Administration: No computer is required for the completion of the Rey-Osterrieth Complex Figure test. Since performance on the test is based on the accuracy of the reproduction of the figure, the test may be given to several people simultaneously.

Level of expertise: No expertise is required for the administration of the test. However, some training is required with respect to the scoring methods. In order to limit the influence of the subjectivity inherent to these types of tests (e.g., the Clock Drawing Test has similar subjective elements), administrators should have a clear idea of what counts as a correctly drawn element versus an incorrectly drawn element (Canham et al. 2005). In order to achieve this inter-rater reliability, an "answer key" can be distributed to all administrators. This key would have examples of what each portion of the image looks like and some examples of the correct way of drawing these portions. Administrators would give points for each portion of the image drawn correctly which could be added up for a total score.

Single Letter Cancellation Test - Error

Overview: The Single Letter Cancellation Test evaluates the presence and severity of visual scanning deficits. The test consists of a regular-sized piece of paper with six lines of 52 letters. The participant is instructed to cross out the letter H whenever it appears in the rows (the letter H is presented 104 times).

Duration: No time limit is given for the test, however participants typically require less than five minutes to complete the task.

Scoring: There are several ways to score the test. Most easily, the score is calculated by counting the number of omissions (i.e., 'H's that were not crossed out). Scoring may also include counting the number of letters crossed out that were not appropriate (i.e., not 'H'). Therefore, a lower score corresponds to better performance on the test.

Administration: Administration of the Single Letter Cancellation Test is straightforward and does not require the use of a computer or any other technical equipment. Participants would be given the appropriate sheet for the test and the administrator would have to explain the test instructions. This test can be easily administered to several individuals simultaneously

Level of expertise: No expertise is required for the administration or the scoring of the single letter cancellation test (Canadian Stroke Network). Instructions provided to complete the test are straightforward; participants simply need to know which letter to cross out, and scoring consists of counting only, there are no calculations

Double Letter Cancellation Test - Error

Overview: The Double Letter Cancellation Test is different from the Single Letter Cancellation Test insofar as the former requires participants to cross out two letters instead of just one. The test assesses visual scanning abilities. The participant is presented with a regular sized sheet of paper containing six lines of 52 letters. He or she is then instructed to cross out the letters 'C' and 'E' whenever either one appears ('C' and 'E' are presented 105 times combined).

Duration: No time limit is given for the test, however participants typically require less than five minutes to complete the test.

Scoring: There are several ways to score the test. Most easily, the score is calculated by counting the number of omissions (i.e., 'C's and 'E's that are not crossed out). Scoring may also include counting the number of letters crossed out that were not appropriate (i.e., not 'C' or 'E'). Therefore, a lower score corresponds to better performance on the test.

Administration: Administration of the Double Letter Cancellation Test is straightforward and does not require a computer or any other special piece of equipment. Like the Single Letter Cancellation Test, the Double Letter Cancellation Test can be administered easily to several individuals simultaneously.

Level of expertise: No expertise or special training is required to administer and score the Double Letter Cancellation Test (Canadian Stroke Network). As with the Single Letter Cancellation Test, scoring the test requires only counting and addition, there are no complex calculations.

Traffic Sign Recognition Test (TSRT)

Overview: The Traffic Sign Recognition Test is thought to test the cognitive capacities involved in the driving task, notably memory. This test also has the added benefit of being very easy to incorporate into the licence renewal process. The specifics of how the Traffic Sign Recognition Test is administered – including whether the test is conducted one-on-one and whether completion time is a factor in scoring – vary substantially across licencing districts. Some Traffic Sign Recognition Tests are conducted individually and the participant answers orally, while other versions consist in a multiple choice test. Despite these differences, the fundamentals of the test are the same: the participant is asked to explain the meaning of a number of traffic signs. Failure to recognize a sign correctly counts against the participant's overall score and may indicate the need for an on-road test.

Scoring: The scoring of the Traffic Sign Recognition Test is normally based on the number of errors. To illustrate, the Traffic Sign Recognition Test from North Carolina cited in Stutts et al. (1996) asks the participant to identify twelve signs. If the participant fails to recognize three or more signs accurately, he or she is considered to have failed. There is no consensus on whether or how completion time should be factored into the overall score of the Traffic Sign Recognition Test.

Duration: The duration of the test will vary with differences in the number of questions asked.

Administration: The test is easy to administer, as it involves only a pencil, paper, and an answer key. This test is currently administered in a variety of forms as part of the licencing process, so no added training would be necessary to include it as part of a cognitive screen.

Level of expertise: This test requires no specific expertise, and can be administered by licence examiners (Stutts et al. 1996).

Maze Task (time)

Overview: The Maze Task is a short, pencil-and-paper screening tool designed to test attention, visuoconstructional skills, planning, and foresight. Participants are presented with a maze and instructed to complete the maze as quickly as possible without running into dead ends or crossing the solid lines. In order to ensure that the task is clear, participants may be given the option to complete a demonstration maze first, which is not scored.

Scoring: The Maze Task is scored by measuring the time required for the participant to complete the maze. If the participant requires over 60 seconds to complete the maze then he or she is considered to be cognitively unfit to drive. Errors like reaching dead-ends or crossing solid lines count only insofar as they may extend the amount of time required to properly complete the maze. A maze is considered properly completed if, as per the instruction, a route has been traced from the entrance of the maze to its exit. As such, it does not matter how many dead-ends the participant reaches as long as a clear, barrier-free route from entrance to exit is discernible. For example, if a participant submits a maze wherein the route from entrance to exit was only made possible via a crossed solid line, then he or she would not be considered as having passed, as the maze was not completed properly. Conversely, if a participant hits several dead-ends, but each time notices this and self-corrects by tracing the route back, and ultimately succeeds in finding the exit, then, given that the maze is completed in under 61 seconds, he or she will be considered as having passed. The difference is that in the latter case, despite the dead-ends reached by the participant, self-correction allowed for a barrier-free route from entrance to exit; while in the former case, a barrier-free route from entrance to exit was not achieved, as the route traced by the participant violated one of the instructions given at the outset, namely, that no solid lines should be crossed.

Duration: The Maze Task takes approximately one minute to complete and score.

Administration: The Maze Task does not require a computer. Administration of the screening tool is simple and requires little instruction. There are specific instructions that administrators are encouraged to use when presenting the Maze Task to participants: the administrator presents the maze in the correct orientation in front of the subject. He or she then instructs the participants to find the route, from the start to the exit of the maze. The administrator points out to the participants the location of the entrance and exit of the maze. He or she then instructs the test-takers to avoid running into dead-ends or crossing any solid lines. The administrator shows the participants examples of dead-ends and solid lines to ensure that these conditions are understood (Snellgrove 2006).

The administrator could time the class as a whole. If a cut-off of 61 seconds is used, and completion time is the only variable included in the scoring method, then there should not be any difficulties in administering the task to a group. The administrator could then go over the submitted maze and evaluate the number of errors (dead ends or crossing of solid lines) for each participant.

Level of expertise: The level of expertise required to administer and score the Maze Task is minimal. No special training is required, and scoring benchmarks are clearly delineated and easily interpreted (Snellgrove 2010).

Charron Test - Error

Overview: The Charron Test is designed to evaluate visual attention processing. The test consists of a piece of paper containing 19 pairs of objects and 37 pairs of numbers. The participant is instructed to place a checkmark next to all the non-identical pairs, as quickly as possible.

Scoring: Performance on the Charron Test is determined by counting the total number of errors. An error is defined as a non-identical pair that is not checked, or an identical pair that is checked.

Duration: The Charron Test does not have an official time limit. Certain methods of scoring of the Charron Test take into account the amount of time in seconds required for the participant to complete the test, however the 'error' version does not require this piece of information. Typically, participants finish the test in approximately five minutes.

Administration: Administration of the Charron Test is straightforward and does not require the use of computers or other technical devices. Since the score on this version of the test is determined solely by the number of errors and omissions, the Charron Test could be administered to several participants simultaneously.

Level of expertise: No expertise is required to administer or score the Charron Test (Mazer et al. 1998).

Wechsler Digit Symbol Substitution Test (DSST)

Overview: The Wechsler Digit Symbol Substitution Test is designed to measure working memory and psychomotor performance. This test, largely unmodified since its development, is one of several intelligence tests developed by David Weschler in 1939. The participant is presented with a list of numbers 1-9, with a unique symbol under each number. Organized into rows underneath are several dozen numbers overtop blank boxes. The task of the participant is to insert as many of the unique symbols associated with each of the numbers in the rows as possible within a 90-second timeframe. The participant is instructed to fill in the boxes in the order that they appear, and not to skip ahead.

Scoring: When the time limit passes, the test administrator counts the number of correctly filled-in boxes. This number represents the score. High scores are associated with normal to very mild cognitive impairment, while low scores indicate more severe cognitive impairment. Any boxes correctly filled in after two or more consecutive blank boxes are not counted in the score. According to Lafont et al. (2009), setting the benchmark for cognitive impairment at 29 yields the best compromise of sensitivity (91.7%) and specificity (81.2%).

Duration: The test can be completed and scored in less than five minutes.

Administration: The administration of the test is straightforward and does not require a computer. The test can be administered to several individuals simultaneously.

Level of expertise: The level of expertise required to administer the DSST is minimal (Lafont et al. 2010; Bettcher et al. 2013) because it is simple to instruct ("fill in the correct number with the corresponding symbol") and easy to score with simple addition being the only necessary calculation.

Visual Form Discrimination Test

Overview: The Visual Form Discrimination Test is a brief, multiple choice test designed to measure the ability to make fine visual distinctions. The test consists of two sample questions, and 16 test questions. Each question has the following format: a simple design (referred to as the 'target') is positioned above five answer options (referred to as 'stimuli'). One of the four stimuli matches the target exactly, and this is the correct answer. The remaining three stimuli contain minor variations, including rotation of peripheral elements or distortions of a major shape. The participant must indicate, for each of the 16 test items, which stimulus corresponds to each target.

Scoring: The scoring is based on a points system, with a higher score indicating better performance on the test. Two points are awarded for a correct match, one point is awarded for an incorrect match involving an error of a peripheral shape, and no points are awarded for an incorrect match involving a major shape error. The highest possible total score is 32 points.

Duration: There is no official time limit for the test. Typically, healthy older adults require less than thirty seconds to complete each question. Thus, a test that lasts longer than eight minutes may indicate some level of cognitive impairment.

Administration: Administration of the visual form discrimination test is straightforward and requires no computer or other technical devices. Since performance on the test is determined by the answers to the multiple choice questions, the Visual Form Discrimination Test can be administered to a group.

Level of expertise: No special expertise is required in order to administer or score the Visual Form Discrimination Test (Caplan and Scultheis 1998). An answer key is provided to administrators to allow for simple scoring.

Clock Drawing Test

Overview: The Clock Drawing Test measures cognitive capacity and executive functioning. Particularly, the test investigates visuoconstructional skills, working memory, and planning. The participant is given a blank piece of paper with a large circle drawn on it. The administrator indicates the top of the page to the participant. He or she is then instructed to draw numbers on the circle to make it look like the face of a clock, and then to draw the hands of the clock to read "ten after eleven." Participants are given an unlimited amount of time to produce the clock.

Scoring: There are a number of valid scoring methods for the test. The score is determined by several factors including the order and spacing of the numbers; whether the numbers are inside the clock face or not; the presence of clock hands and accuracy of the time drawn; and the presence of any markings unrelated to the task of drawing the clock (e.g., words, names, extra circles, underlines, or pictures). Since the clock-drawing test has two overall challenges – drawing the clock and drawing the time – scoring tends to reflect the participant's performance on both of these aspects individually.

An example of a simple scoring system from Shau-Haim et al. (1996) suggests awarding one point for correct aspects: approximate drawing of the clock face, numbers in sequence, correct spatial arrangement between numbers, presence of clock hands, hands showing the approximate time, and hands showing the exact time. According to this scoring system, a perfect score would be 6.

Other scoring systems add points for errors made. For example, Shulman et al. (1993) allocates a score of 1 to a perfectly drawn clock; 2 for a clock displaying minor visuospatial errors (e.g., mildly impaired spacing, lines drawn outside the circle); 3 for an inaccurate representation of the time when visuospatial orientation is perfect or shows only minor errors; 4 for moderate visuospatial errors that are severe enough to make the depiction of the correct time impossible; 5 for severe visuospatial errors that make drawing the time impossible (e.g., very poor spacing, numbers omitted, numbers that continue past 12, severe dysgraphia[1]); and 6 if the drawing bears no resemblance to a clock, or if the participant made no attempt to draw one. In contrast to the previous example, this scoring system counts a 1 as a perfect score.

<u>Duration</u>: There is no set time limit for the test, however completion normally requires less than two minutes.

<u>Administration</u>: A computer is not necessary for the Clock Drawing Test. Since there is no time limit and since scores are calculated based on the number of errors recorded, the administrator is not required to pay individual attention to the participant. This test can therefore, be administered to several individuals simultaneously.

<u>Level of expertise</u>: The administrator must be familiar with the scoring system used in his or her location, and the specific criteria at the foundation of that system. More specifically, it must be clear to test administrators what counts as, for example, “moderate spacing errors” versus “severe spacing errors.” To this extent, some level of expertise is required in order to score the Clock Drawing Test. However, no special training is required for the administration of the test (Shulman et al. 1993; Shua-Haim et al. 1996.)

Eight Item Informant Interview

<u>Overview:</u> The purpose of the Eight Item Informant Interview is to detect the presence and severity of dementia. The interview consists of eight questions, presented to the person responding to the questions to complete him or herself. While it is preferable for an informant (e.g., family member or caregiver) as opposed to the participant to complete the questionnaire, the interview may be administered to the participant directly.

<u>Scoring</u>: The questionnaire is scored by adding up the number of questions answered in the affirmative. Hence, a higher score indicates a greater likelihood that the participant is cognitively impaired to some degree. A score of zero or one denotes normal cognition, while a score greater than two indicates the likely presence of impairment.

[1] *Dysgraphia* is a deficiency in the ability to write primarily in terms of handwriting.

Duration: There is no official time limit for completion of the questionnaire, however taking too long to complete it may indicate some level of cognitive impairment.

Administration: Administration of the interview is straightforward and requires no computer or other technical equipment. The questionnaire could easily be administered to several individuals simultaneously.

Level of expertise: No expertise is necessary for the administration or scoring of the Eight Item Informant Interview (Alzheimer's Association 2012).

Table 2: Tools Included in the Meta-Analysis

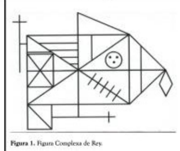
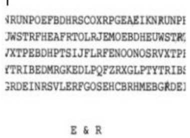
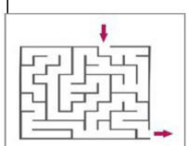
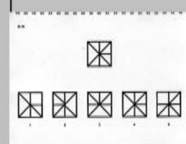
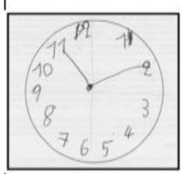
ID #	Tool	Image	Duration	Skills Tested	Administration	Feasibility	Computer	Expertise	Scoring Method
2C, 2D, 69A	Rey-Osterrieth Complex Figure Test		Around 10 minutes to administer; 5 minutes to score	Visual perception; long-term memory visual capacities	Pencil and paper: Re-creating a drawing	Individual completion in a group	No computer is required	No expertise required	Test is scored using a points key
38E, 108G, 108H,	Single Letter Cancellation Test (Error)		Under 5 minutes to administer; under 3 minutes to score	Visual scanning skills	Pencil and paper: Filling in letters	Individual completion in a group	No computer is required	No expertise required	Test is scored by counting number of omissions
108I, 108J	Double Letter Cancellation Test (Error)		Under 5 minutes to administer; under 3 minutes to score	Visual scanning skills	Pencil and paper: Filling in letters	Individual completion in a group	No computer is required	No expertise required	Test is scored by counting number of omissions
107, 113D	Traffic Sign Recognition Test (TSRT)		Depends on # of questions; under 3 minutes to score	Memory; knowledge of the meaning of road signs	Pencil and paper: Multiple choice	Individual completion in a group	No computer is required	No expertise required	Test is scored using an answer key
15, 25, 33, 33A, 38C, 38D, 78A	Maze Task (Snellgrove)		1 minute to administer; under 1 minute to score	Attention; visuoconstructi onal skills; planning, foresight	Pencil and paper: Completing a maze	Individual completion in a group	No computer required, one stop watch used to time entire class	No expertise required	Test is scored based on if completed and errors
108A, 108 B	Charron Test Error	No image available	Around 5 minutes to administer; under 3 minutes to score	Visual attention processing	Pencil and paper: Identification of non-matching pairs	Individual completion in a group	No computer is required	No expertise required	Test is scored by counting the number of errors

Table 2: Tools Included in the Meta-Analysis

ID #	Tool	Image	Duration	Skills Tested	Administration	Feasibility	Computer	Expertise	Scoring Method
32C, 62, 117E	Wechsler Digit Symbol Substitution Test		Under 5 minutes to administer; under 3 minutes to score	Working memory; psychomotor performance	Pencil and paper: Symbol matching	Individual completion in a group	No computer is required	No expertise required	Test is scored by counting the number of errors or blanks
32D, 38F, 38G, 69D	Visual Form Discrimination Test		30 seconds to 8 minutes to administer; under 2 minutes to score	Ability to make fine visual distinctions	Pencil and paper: Multiple choice	Individual completion in a group	No computer is required	No expertise required	Test is scored using an answer key
8B, 78B, 96, 103A	Clock Drawing Test		Under 2 minutes to administer; under 5 minutes to score	Executive functioning	Pencil and paper: Drawing	Individual completion in a group	No computer is required	No expertise required	Test is scored using a points key
78E	8 Item Informant Questionnaire	No image available	Under 10 minutes to administer; under 3 minutes to score	Presence and severity of dementia	Pencil and paper: Questionnaire	Individual completion in a group	No computer is required	No expertise required	Test scored using number of affirmative answers

3.2 Cognitive screening tools not included in the meta-analysis

The following are tools that were reviewed for the meta-analysis but excluded because they did not meet MTO requirements. These criteria are described further in the Methods section.

Useful Field of View Test (UFOV)

Overview: The UFOV is an online evaluation of "useful field of view" which measures the brain's ability to take in and react to information at a single glance. The test includes three subtests in which measures are taken of an applicant's processing speed, processing speed during a divided attention task, and processing speed during a selected attention task.

Scoring: The three subtests are scored by a range of 0-500. These scores are combined to calculate five categories of risk ranging from Category Level 1 (lowest risk) to Category Level 5 (very high risk).

Duration: The test takes 15 minutes to complete.

Administration: Computers are required for participants to complete the test. They are also required for the administrators to score the test results.

Level of expertise: Training for those administering the test is included in the software (UFOV User's Guide 2002).

Assessment of Driving-Related Skills (ADReS)

Overview: The ADReS is a tool used for a driving specific evaluation of older drivers using a battery of tests. Elements tested are a participant's vision, cognition, and motor function. A participant's vision is evaluated for visual acuity and visual fields. Cognitive skills that are tested are 1) memory, 2) visual perception, visual processing, and visuospatial skills, 3) selective and divided attention, and 4) executive skills. Also, motor function, muscle strength and endurance, range of motion of the extremities, trunk, and neck, and proprioception[2] are appraised. Tests such as the Snellen E Chart (visual), Trail Making B, Clock Drawing Test (cognitive), Range of Motion, and Rapid Pace Walk (physical) are all included.

Scoring: Each test in the ADReS has its own method of scoring. Physical tests are scored based on participant's muscle strength, their balance, or visual ability. Cognitive tests are timed or scored based on task performance. Results from all tests are summarized on the ADReS score sheet.

Duration: The test takes 15 minutes to complete.

Administration: A computer is not required. What is needed is a Snellen Chart, tape to mark distances on the floor, a stopwatch, and a pencil. There are two paper-and-pencil tests in ADReS, one of which requires a pre-printed form. The administrator works directly with the participant in order to assess their abilities in each area. This involves activities such as the administrator sitting across from the participant and holding

[2] Proprioception is the perception of movement and spacial orientation.

up a certain number of fingers, while the participant closes one eye and has to inform the administrator of how many fingers they are holding up. Another example is the Rapid Pace Walk where the participant is instructed to walk a 10-foot path, turn around, and walk back to the starting point as quickly as possible while the administrator times them with a stop watch.

Level of expertise: A driver rehabilitation specialist is required to administer the test (Mevius 2013).

Performance Analysis of Driving Ability (P-Drive)

Overview: The P-Drive is a list of 27 items which assess the quality of a participant's driving performance, commonly in a driving simulator. When used, P-Drive allows an administrator to record and analyze all relevant driver and vehicle data objectively. It is used to improve the participant´s driving ability for all kinds of target groups. The test is sorted into four subgroups: manoeuvring, orientating, obeying and responding to traffic rules, and paying attention.

Scoring: Each item is scored on a scale from 1 to 4, where 4 equals 'competent driving ability,' 3 equals 'questionable driving ability,' 2 equals 'problematic driving ability,' and 1 equals 'incompetent driving ability.'

Duration: The test takes 40-60 minutes to complete.

Administration: A computer and a driving simulator are needed. The administrator will supervise the participant while they take a pre-determined course presented in a driving simulator. The administrator assesses the participant's performance in the simulator using the 27 items in the P-Drive.

Level of expertise: An occupational therapist with a history of performing driver evaluations is needed to conduct the test and evaluate a participant's performance (Selander et al. 2011).

Nordic Stroke Drive Screening Assessment (NorSDSA)

Overview: The NorSDSA is a set of cognitive tests originally developed to evaluate fitness-to-drive in stroke clients. More recently, NorSDSA has been used for non-stroke clients who suffer from cognitive deficits or dementia.

Scoring: Four subtests providing six subscores: Dot Cancellation which measures in seconds to completion (maximum of 15 minutes) and number of errors, Directions which yields a maximum of 32 points, Compass which also yields a maximum of 32 points, and Road Sign Recognition which is scored 0-12 after three to five minutes. Higher scores on Directions, Compass, and Road Sign Recognition are representative of a better score on the test.

Duration: The test takes 28 minutes to complete (Selander et al. 2011).

Administration: Individual administration is necessary for the NorSDSA because the administrator is required to time the participant. Completion time is used for scoring.

Level of expertise: The NorSDSA is relatively simple to administer, with calculation of timing being the main requirement, but scoring may be more time consuming, because more than one score is weighted and combined to develop the overall score.

Visual Selective Attention Test (VSAT)

Overview: The VSAT serves to quickly measure attentional processes that are commonly disrupted in acute and chronic brain damage or disease. Participants respond to 24 visual stimuli presented on a computer screen, such as numbers or letters, using the computer keyboard. These stimuli are shown for two seconds and then the participant is given another two seconds to respond to what he or she witnessed.

Scoring: A scoring manual is provided with the software. The test takes 15 minutes to score.

Duration: The test takes six minutes to complete.

Administration: A computer is required to use VSAT (Trenerry et al. 2012). This test would require individual administration.

Level of expertise: Low because the test and scoring is done on the computer.

DriveABLE Cognitive Driving Assessment (DCAT)

Overview: The DCAT is a computer presented scientifically validated driving assessment tool for evaluating medically at-risk drivers. Memory, attention, reaction time, and judgment are assessed using a touchscreen and push button response.

Scoring: Scoring and report generation are automated.

Duration: The test takes 30-60 minutes to complete.

Administration: A computer is needed for this test.

Level of expertise: The test is administered by healthcare professionals and certified evaluators that have been specially trained by DriveABLE (Mevius 2013).

Drive Aware

Overview: The Drive Aware test is comprised of eight questions that measure awareness of a participant's driving ability. For example, retrospective self-evaluation of driving performance, right of way rules, and participant reaction to vehicles or pedestrians that "appear out of nowhere."

Scoring: A participant's response for each question is scored on a defined scale from 1 (very aware) to 3 (very unaware). Using information provided in referrals and the participant's performance on clinical tests, clinicians award scores by using a scale from 1 (participant has poor performance or knowledge) to 3 (participant has good performance or knowledge). For each item, a discrepancy score is calculated by subtracting the clinician's score from the participant's score. This results in a total discrepancy score. The participant's awareness of driving ability is rated according to the following criteria: discrepancy score of 1 or less (intact awareness), discrepancy score of 2 to 4 (partial awareness), and discrepancy score of 5 to 10 (absent awareness).

Duration: Unknown.

Administration: Individual administration is needed because both the participant and clinician must complete the questions and the clinician is required to calculate the score. Scoring would require a high degree of subjectivity in so far as the clinician awards scores based on information provided by referrals.

Level of expertise: A clinician is needed to administer this test (Kay et al. 2009).

Drive Safe

Overview: Drive Safe measures a participant's awareness of the driving environment. Thirteen images that simulate the view through a windshield of the same roundabout are projected on a screen. The number and positioning of pedestrians and vehicles vary from one image to the next. Participants observe each image for 3 seconds, then once the image has been removed from the screen, report details about the position and direction of travel of each pedestrian and vehicle in that image. The images vary in complexity, requiring participants to report from 4 to 16 elements.

Scoring: The participants complete three practice images to ensure that they understand the instructions. Performance is recorded as a score out of 140. Verbal responses or hand gestures are recorded by the evaluator. The test is administered according to the standard instructions included with the standard scoring sheet.

Duration: The test takes 20 minutes to complete.

Administration: A computer or projector is required.

Level of expertise: The test is administered by a clinician.

Visual Recognition Slide Test (VRST-Usyd)

Overview: The VRST-Usyd tests a participant's awareness of the driving environment. He or she is shown 15 images of the same roundabout, through a windshield, where the number of pedestrians and vehicles vary in each image. Participants have 3 seconds to observe the images. The images are then removed from the screen and the participants must report details about the position and direction of travel of each pedestrian and vehicle in the slide.

Scoring: The participants complete the practice images to ensure that they understand the instructions. Performance is recorded as a score out of 164. Verbal responses are preferred if participants can provide them. The test is scored using the standard scoring sheet.

Duration: The test takes 20 minutes to complete.

Administration: A computer or projector is required.

Level of expertise: The test is administered by a clinician (Kay et al. 2008).

Gross Impairment Screening Battery (GRIMPS)

Overview: The GRIMPS assesses performance-based physical and cognitive abilities likely to affect driving performance. It includes both physical measures (e.g., rapid walk, arm reach, foot-tap test, vision tests) and cognitive measures (e.g., cued and delayed recall, symbol scan, Trail Making, MVPT).

Scoring: Each portion of the battery is scored differently.

Duration: The test takes 11 minutes to complete.

Administration: In order for tests and results to be sufficiently executed and recorded, the administrator must work individually with the participant. For example, the cued recall portion of the exam requires the administrator to sit across from the participant at a table, recite three words and ask the participant to repeat those same words in the same order. The administrator then documents if the task was completed successfully.

Level of expertise: The test is administered by motor vehicle administration staff (Ball et al. 2006).

Safe Driving Behaviour Measure (SDBM)

Overview: The SDBM is a 68-item self-report or proxy measure to assess safe driving behaviours.

Scoring: The measure score represents the reported level of difficulty for the items given the participant's ability level. Difficulty with the driving task is rated via a 5-point adjectival scale ranging from 1 (cannot do) to 5 (not difficult).

Duration: The test takes over 30 minutes to complete.

Administration: The questionnaire can be self-administered or completed by a family member or caregiver.

Level of expertise: The test is administered by driving evaluators (Classen et al. 2012).

Mini-Mental State Examination (MMSE)

Overview: The MMSE is a 30-point questionnaire used to screen adults for psychiatric, neurological, and general medical conditions. The exam tests the participant's skill in areas such as time and space orientation, attention and calculation, recall, and language. For example, the participant is asked to draw two pentagons interlocking with one another.

Scoring: Scores range from 0 to 30, with scores of 25 or higher considered normal. Scores less than 10 indicate severe impairment, while scores between 10 and 19 indicate moderate dementia. People with mild Alzheimer's disease tend to score in the 19 to 24 range.

Duration: The test takes 10 minutes to complete.

Administration: This exam requires the participant to verbally recall words for the administrator, meaning the administrator must be present with the individual in order to administer the exam.

Level of expertise: This test is administered by an occupational therapist (Mevius 2013).

Driving Health Inventory

Overview: The Driving Health Inventory is a battery of computer-based performance tasks that include measures of visual acuity, leg strength and stamina, head and neck flexibility, short-term and working memory, visualization of missing information, visual search with divided attention, and visual information processing speed. It is administered as a first tier screening for fitness to drive.

Scoring: Performance feedback indicating level of impairment is produced by the computer.

Duration: Unknown.

Administration: A computer is required.

Level of expertise: The test is administered by an occupational therapist (Mevius 2013).

Addenbrooke's Cognition Examination Revised (ACE-R)

Overview: The ACE-R is a brief cognitive test that was developed to detect mild cognitive impairment and dementia. It includes tests of visuospatial capabilities, fluency, and memory. The structure of the test requires the administrator to lead the participant through several different exercises. For example, participants must verbally calculate subtracting 7 from 100, then 7 from 93, then 7 from 86 and so on. Or, to test their language, they have to write down the phrase "Happy Birthday."

Scoring: The maximum score is 100 and the proportion of the score is divided as follows: orientation (10), attention (8), memory (35), verbal fluency (14), language (28), and visuospatial ability (5).

Duration: The test takes 15-20 minutes to complete (Ferreira et al. 2012).

Administration: The test relies heavily on the interaction between the administrator and the participant. The administrator guides the participant through the various activities and is responsible for recording their responses and results.

Level of expertise: The test is commonly administered by a healthcare professional.

Sensory Motor Cognitive Tests (SMC Tests)

Overview: The SMC Tests are a battery of computerized sensory-motor and cognitive tests used to quantify dysfunction in both areas. The sensory-motor tests comprise tests of visuoperception, ballistic movements, and visuomotor tracking. The cognitive tests include newly developed measures of complex attention, visual search, decision-making, impulse control, planning, and divided attention. For example, the "Footbrake Reaction Time" tests the reaction and movement time a participant needs to respond to a signal. The "Complex Attention" test assesses the participant's ability to sustain attention over an extended period of time.

Scoring: Scoring is computed by the simulator which records elements such as reaction time and maintenance of attention.

Duration: The test takes 45 minutes to complete.

Administration: A driving simulator is required (Innes et al. 2007).

Level of expertise: The administrator of these tests would need working knowledge of the examinations used by the computer and the computer simulator.

Screening Tool for the Identification of Cognitively Impaired Medically At-Risk Drivers: A Modification of the DemTect (SIMARD-MD)

Overview: The SIMARD-MD is a screening tool that makes use of select tests from the longer DemTect. The SIMARD consists of four pencil-and-paper tasks designed to challenge the cognitive skills that have been found to be predictive of driver performance: two tasks that measure the ability to remember words, one task involving converting numbers to words, and one task where the participant names objects that can be purchased at a supermarket within a one-minute time frame.

Scoring: Each task is scored differently, using a point system for accuracy. The scores are interpreted using the following key: a score of 30 or less indicates a high probability of failing a driving test, a score of 31-70 indicates that a referral for a driving assessment is appropriate, and a score greater than 70 indicate a low probability of failing a driving test.

Duration: The test takes seven minutes to complete.

Administration: This test requires the participant to name objects out loud, meaning the administrator needs to be present with each participant in order to listen to and record their responses (Dobbs and Schopflocher 2010).

Level of expertise: The test is often administered by a clinician.

Trail Making A

Overview: The Trail Making Test A measures visuospatial capabilities, planning, and motor function. The screen consists of 25 circles distributed over a sheet of paper. Each circle contains a number from 1 to 25. Participants are instructed to connect the numbers in ascending order as quickly as they can, without lifting the pencil from the sheet of paper. Performance on the test is based on the time it takes to complete the task.

Scoring: The administrator calculates the amount of time the participant requires to complete the trail. If an error is made (e.g., a number is skipped), the administrator points it out immediately and allows the participant to correct it. Errors like this have no effect on the overall score except that the time spent correcting the error is included in the completion time.

Duration: The test takes under five minutes to complete.

Administration: The administrator must make use of some timing device, and must sit individually with the participant in order for the timing to be precise and to point out errors immediately (Strauss et al. 2006).

Level of expertise: The level of expertise required for administering Trail Making A is low because it only requires the administrator to use a stop watch.

Trail Making B

Overview: The Trail Making Test B consists of 25 circles distributed across a sheet of paper. The circles contain either a number (1-12) or a letter (A-K). Participants are instructed to connect the circles in an ascending pattern, alternating between numbers and letters (i.e., 1-A-2-B-3-C, etc.), as quickly as possible without lifting the pen or pencil from the paper. Performance on the task is based primarily on the time it takes for the participant to complete the trail.

Scoring: The administrator calculates the time required for the participant to complete the trail. If an error is made, the administrator points it out immediately and the participant corrects the error. Errors like this have no effect on the overall score except to the extent that the time required to correct the error is included in the total completion time.

Duration: The test takes under five minutes to complete.

Administration: The administrator must make use of some timing device, and must sit individually with the participant in order for the timing to be precise and to point out errors immediately (Strauss et al. 2006).

Level of expertise: The level of expertise required for administering Trail Making B is low because it only requires the administrator to use a stop watch.

Colour Trails 1

Overview: The Colour Trails Test 1 consists of a sheet of paper with 25 encircled numbers, from 1 to 25. The circles are also coloured: odd numbers are pink and even numbers are yellow. The participant is instructed to connect the circles in ascending order as quickly as possible without lifting the pencil or pen from the paper. Performance on the test is based primarily on the time it takes to complete the task.

Scoring: The Colour Trails Test 1 is scored according to the length of time required for the participant to complete the trail, which is recorded for each individual participant by the administrator.

Duration: The test takes five minutes to complete.

Administration: This test requires the administrator to time the length needed to complete the test, meaning it is administered individually (Dugbartey et al 2000, Konstantopoulos et al. 2013).

Level of expertise: The level of expertise required for administering the Colour Trails Test 1 is low because it only requires the administrator to use a stop watch.

Colour Trails 2

Overview: The Colour Trails Test 2 consists of 50 circles, each one containing a number from 1 to 25. Each number is repeated one time, once in a yellow circle, and once in a pink circle. The participant is instructed to connect the numbers 1 to 25, alternating between the pink and yellow circles, as quickly as possible without lifting the pen or pencil from the page.

Scoring: Performance on the Colour Trails Test 2 is determined by the length of time to complete the test along with an "interference indicator," a function of the number of errors, near-misses, or prompts. The calculation of this indicator makes use of the time score for Part 1 of the Colour Trails Test, suggesting that the two tests cannot be conducted separately while retaining the same sensitivity and specificity of the Colour Trails Test 2.

Duration: The test takes five minutes to complete.

Administration: This test requires the administrator to time the length needed to complete the test, meaning it is administered individually.

Level of Expertise: The level of expertise required for administering the Colour Trails Test 2 is low because it only requires the administrator to use a stop watch. However, scoring the test may require special training. Specifically, administrator would have to be taught how to calculate the interference indicator and the effect of errors on the overall score (Dugbartey et al. 2000, Strauss et al. 2006).

Frostig Visual-Motor Worksheet

Overview: The Frostig Visual-Motor Worksheet is a paper and pencil test in which participants copy a pattern using a grid of dots; then they are scored and timed.

Scoring: Participant's score is calculated by combining the time the participant takes for completion, their errors of omission and their errors of commission.

Duration: The test takes one to five minutes to complete.

Administration: Because each participant is timed, this test would require individual administration.

Level of expertise: The test is administered by physicians and other therapists (Johnson 2007).

Dot Cancellation Test

Overview: The Dot Cancellation Test is a measure of attention and concentration and is a subtest of the Stroke Driver Screening Assessment. Subjects are instructed to cross out all groups of four dots on an A4 piece of paper.

Scoring: The number of uncrossed groups of four dots, groups of dots other than four that are crossed, and the time spent for completion are tallied for the participant's score.

Duration: The test takes 15 minutes to complete.

Administration: The test would require individual administration due to the timing component (administrator would need to measure the length of time the participant needed for the test). As well, this test would be time consuming to score, requiring the administrator to differentiate between the crossed and uncrossed dots.

Level of expertise: The test is commonly administered by a psychologist (Akinwuntan et al. 2005).

Hazard perception

Overview: Participants must recognize potential traffic hazards in video clips of traffic scenes filmed from the driver's point-of-view. Participants must press the touchscreen when they identify a potential hazard.

Scoring: The software records the participant's response time, which is then used to calculate the hazard perception response latency.

Duration: Unknown.

Administration: A computer is needed for this assessment (Wood et al. 2013). Scoring is done by the computer.

Level of expertise: The level of expertise required is low because the test is administered and scored by the computer.

Hazard change detection

Overview: This test measures the participants' ability to detect the presence of hazards independent of other factors. Pairs of still images of traffic scenes are displayed on a computer screen using the flicker paradigm. Each pair contains an original and an altered image and participants are required to press the screen where the image is altered.

Scoring: The score is calculated by the number of pairs the participant correctly differentiates.

Duration: Unknown.

Administration: A computer is needed for this task (Wood et al. 2013).

Level of expertise: Scoring is calculated mainly by the computer which means that the level of expertise required to administer the test is low.

Gestalt Completion

Overview: Drawings are presented to the participant which are composed of black blotches representing parts of the object being portrayed. The participant then identifies to the administrator what the drawing portrays. Participants must identify 15 pictures during a timed period.

Scoring: Unknown.

Duration: The test takes three minutes to complete (Dumitrescu 2005).

Administration: The administrator must work individually with the participant so that they can record their portrayal of the pictures and time the exercise.

Level of expertise: Unknown.

Snowy Pictures

Overview: This task involves identifying an object that is obscured by irregular shaped patterns.

Duration: The test takes three minutes for 15 pictures (Wood et al. 2013).

(Limited information is available for this test. It is similar to Gestalt Completion)

AARP Reaction Time

Overview: This test is based on the reaction time test presented in the AARP Older Driver Skill Assessment and Resource Guide, which is similar to Trail Making A. The test consists of a photo of a driving scene onto which 14 numbers are overlaid; the subject must touch the numbers in order.

Scoring: Scoring is based on the total time the participant needs to locate and touch all 14 numbers.

Duration: The test takes under ten minutes to complete.

Administration: This examination has a timing component, meaning the administrator needs to time the participant accurately using a stop watch. The administrator must also supervise the examination to ensure the numbers are being touched in correct order.

Level of expertise: Minimal expertise is needed for administration because it only requires basic supervision in order to ensure numbers are touched in correct order and using a stop watch. (Stutts et al. 1996).

Motor-Free Visual Perception (MFVP)

Overview: The MFVP assesses an individual's visual perceptual ability without any motor involvement needed and tests spatial relationships, visual discrimination, figure-ground, visual closure, and visual memory.

Scoring: The MFVP yields a single raw score that represents the participant's overall visual perceptual ability. The raw score for this test is determined by subtracting the number of errors made from the number of the last item attempted.

Duration: The test takes 25 minutes to complete

Administration: This test is administered individually by the administrator presenting the participant with a stimulus page for five seconds, removing this page and offering the participant the options page (the "answer" page) in which the participant selects their desired answer which is recorded by the administrator.

Level of expertise: The test can be administered by occupational therapists, psychologists, optometrists, and other health professionals (McCane 2006).

Stroop Test

Overview: The Stroop Test consists of one page with colour words printed in black ink, a colour page with 'Xs' printed in colour, and a colour-word page with words from the first page printed in colours from the second page (the colour and the word do not match). The participant goes down each sheet reading words or naming the ink colours as quickly as possible within a time limit.

Scoring: Scores are based on the number of items completed on each of the three stimulus sheets. An interference score, which is useful in determining the individual's cognitive flexibility, creativity, and reaction to cognitive pressures can also be calculated.

Duration: The test takes five minutes to complete.

Administration: The test requires recall (i.e., reading out loud the words or colours), meaning the administrator must work with the participant individually.

Level of expertise: Rehabilitation Specialists commonly administer this test (Golden 2004).

Behind the Wheel Assessment (BTW)

Overview: The BTW is an on-road examination for older adults. It involves a fixed route with a gradual progression of driving difficulty.

Scoring: Behavioural errors are used to score driving maneuvers. Each maneuver is scored on a scale of 0-3 and based on behaviour errors for the given maneuver. The driver is given an outcome rating of safe, safe with restrictions, unsafe remediable, and unsafe not remediable.

Duration: The test takes approximately one hour to complete.

Administration: This on-road examination requires the administrator to be in the vehicle with the driver.

Level of expertise: The assessment is performed by a driving rehabilitation specialist (Justiss et al. 2006).

Probable Risk Checklist

Overview: The Probably Risk Checklist estimates the risk of medically impaired drivers. The checklist is designed to record driver information which relates to their potential risk on the road. Objective risk factors such as the driver's age, their driving exposure, and any previous crashes or convictions are recorded as well as subjective risk factors such as the driver's overall attitude, character, and estimated reliability.

Scoring: For each criterion on the checklist, the administrator indicates the presence or absence of a risk factor by marking 'yes,' 'no,' or 'unknown.' The administrator then tallies the answers and suggests an overall estimate of the driver's risk using a five point scale (very high risk to very low risk).

Duration: The test consists of 39 yes or no questions. It can be completed in approximately 10-15 minutes.

Administration: The test is administered one-on-one between the driver and the administrator. The administrator reads through the risk checklist, prompting the driver to answer the questions honestly. The administrator then records the answers (yes, no, unknown).

Level of expertise: The test was originally designed for administration by "Driver Safety Referees" in a DMV setting. Driver Safety Referees are responsible for driver licencing decisions in medical cases and are trained in risk assessment (Helander 1986).

4. METHODS

The following section describes in detail the process that was used in order to conduct the meta-analysis. The first step discussed is the review of the literature, which involved the collection of various articles related to the study topic. From these, particular articles were selected based on their relevance, i.e., whether they evaluated a cognitive tool or not. Once the applicable tools were highlighted within the selected articles, the data within these articles were analyzed in order to develop feasible data points. Once information regarding tools, their evaluation in relation to driving, and data points were documented, a database was developed to conduct the meta-analysis.

The search for evaluation studies of cognitive screening tools for older drivers began with three separate researchers independently identifying articles that were relevant to the project. The search was conducted using the following key words, search engines, journals, proceedings, and libraries.

Keywords: assessment; attention; cognition; cognitive; decision; decision making; disorder; disorders; driver assessment; driver screening; elderly drivers; executive; executive functioning; fitness; fitness to drive; functional abilities; geriatric assessment; memory; mental; mental ability; metacognition; neuropsychological; neuropsychology; older drivers; perception; performance; predictor; predictors; psychometrics; psychomotor; psychomotor performance; screen; screening; screening battery; self-report; task performance and analysis; test; testing.

Keyword combinations: Geriatric assessment + older driver OR screening battery; Driver screening + cognitive + older drivers; Psychometrics + driver screening OR functional abilities + older drivers; Functional abilities + older drivers; Task performance and analysis + older drivers; Self-report + driver assessment; fitness to drive OR cognitive + elderly drivers; cognitive OR mental ability + testing + elderly drivers; disorders + assessments + fitness to drive.

Journals: Accident Analysis and Prevention; Alzheimer's Disease and Associated Disorders; American Journal of Occupational Therapy; American Journal of Public Health; British Columbia Medical Journal; Canadian Family Physician; Clinics in Geriatric Medicine; Cochrane Database for Systematic Reviews; Human Factors; Injury Prevention; International Journal of Geriatric Psychiatry; Journal of Gerontology; Journal of Occupational Therapy; Journal of Safety Research; Journal of the American Geriatrics Society; Journal of the American Medical Association; Journal of the International Neuropsychological Society; Neurology; Traffic Injury Prevention; Transportation Research; Transportation Research Record.

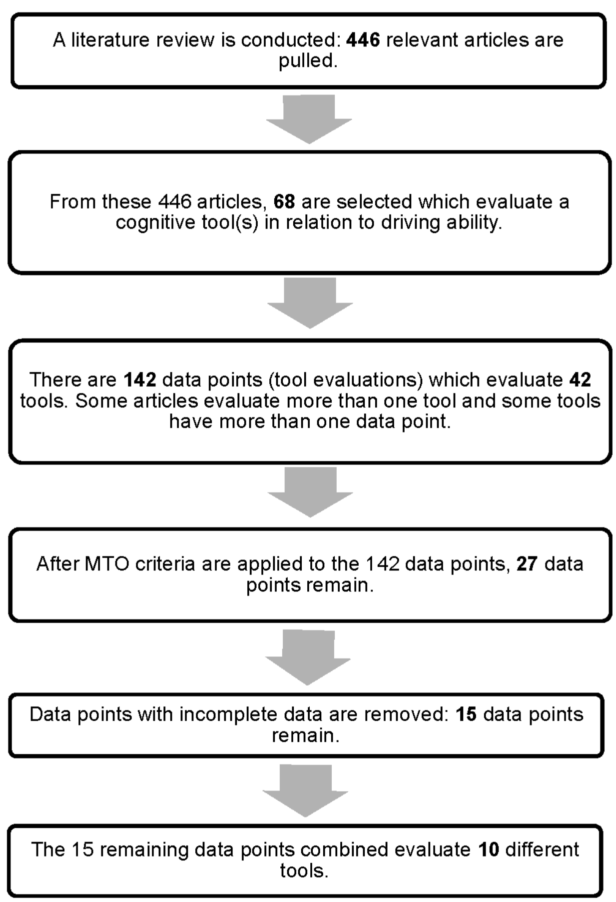
Proceedings: Association for the Advancement of Automotive Medicine; Fit to Drive; Traffic and Transport Behavior Psychology; Transportation Research Board; International Council on Alcohol, Drugs and Traffic Safety

Search engines and online catalogues: APA PsycNet; Factiva; Google; Hein Online; Karger; Medline; Proquest; PsycINFO; Pubmed; Safety Lit; Sage Journals; Scholar's Portal; Science Direct; Springer Link; Wolters Kluwer

Libraries: TIRF Library; Austroads (Australian Department of Transportation); Canadian Association of Road Safety Professionals (CARSP); SWOV Institute for Road Safety Research; IMOB at University of Hasselt; Insurance Institute for Highway Safety (IIHS); National Highway Traffic Safety Administration (NHTSA); University of Michigan Transportation Research Institute (UMTRI); University of North Carolina Highway Safety Research Center (HSRC); University of Toronto

Findings were cross-referenced with a few other sources including an independent search conducted by MTO and an independent search conducted in a TIRF project that had previously been conducted for the University of Hasselt. Some additional references from these independent sources were included (see Appendix A). Below is Figure 1 which is a flow chart describing each step of the methodology.

Figure 1: Methodology Flow Chart



Once the literature search was completed and the researchers combined their respective findings, there were 446 articles listed for further analysis (see Appendix B of this report for this complete list of references that were initially identified). The researchers then extracted all the articles from the literature list, either from the TIRF library or from the Internet. Articles were sorted into three categories. First, an article was deemed as "unusable" if it did not discuss any type of assessment tool or if it was not about the topic of elderly driver assessment or elderly drivers. Second, an article was deemed "of some use" if it directly related to the topic at hand but did not actually evaluate a cognitive screening tool – these articles were still reviewed for the purposes of writing the report, but they were not included in the meta-analysis. The last category was for articles that evaluated the predictive value of a cognitive screening tool or tools for elderly drivers. This category included 142 such evaluations of cognitive screening tools for elderly drivers with and/or without cognitive impairment; these were coded for the meta-analysis.

Several studies included the evaluation of more than one tool. Since the unit of analysis in the meta-analysis is an evaluation of a tool, rather than a study per se, several studies produced more than one data point for inclusion in the meta-analysis. It warrants mentioning that this can cause problems when estimating the precision of the summary effect size in the meta-analysis, due to the fact that results from the same study are correlated. We account for this in the analysis.

Two independent readers reviewed the 142 evaluation outcomes and coded them according to a set of variables of interest to our meta-analysis. Because MTO had specific requirements regarding the cognitive screening tool intended for use in GES, a two-tiered approach was used when coding the evaluations. In a first step, evaluations of tools were coded according to variables enabling us to select those tools that satisfy MTO's requirements, along with some basic descriptive variables about the study that evaluated the tool, such as authors of the study, title of the study, and publication year. Once these 142 evaluation outcomes were coded using these tier-1 variables, a further selection was made according to MTO's requirements and additional information was coded for this sub-selection using tier-2 variables.

MTO's requirements are related to logistical aspects of cognitive screening tools in a GES setting. For example, due to time constraints MTO is only interested in tools that can be administered in a short period of time to avoid taking too much time out of the GES for the screening. Also, it has to be possible to administer the tool during the GES, which is by definition a group activity with only one administrator per session. Thus, a cognitive tool that requires one-on-one administration is not feasible in this setting. More precisely, the following selection criteria were used to select a subset of evaluation outcomes out of the 142 data points that were identified in the literature search. Note that all of these conditions had to be satisfied for a tool to be included.

- **Duration of test** – a variable was created to select tools of different duration; initially we selected those tools that can be administered in less than 30 minutes. Note that this effectively led to the selection of tools that can be administered in less than 10 minutes due to the fact that all tools included in the meta-analysis that take less than 30 minutes actually also take less than 10 minutes;

- **Administration of test** – a variable was created to select between tools that are not limited to individual (one-on-one) administration versus tools that can only be administered individually (e.g., they require the participant to recall words to the administrator or the test is timed); we selected those studies of tools that do not require individual administration;
- **Feasibility of administration** – a variable was created to distinguish between tools according to their ease of administration and scoring; we used it to select those studies of tools that are either easy to administer (i.e., simple instruction) or easy to score (i.e., minimal calculation needed to arrive at the score) while tools that were difficult to administer and difficult to score were not included;
- **Computer** – a variable was created to distinguish between tools that only require pencil-and-paper versus tools that require an expensive piece of hardware like a computer or simulator; only tools that do not require an expensive piece of hardware, but that can be administered with pencil-and-paper were included; and,
- **Expertise** – a variable was created to distinguish between tools that are easily administered without specialized training versus tools that do require such training; only tools that do not require specialized training were included.

After coding the 142 evaluation outcomes using tier-1 variables, researcher one and researcher two compared their coding choices. Any discrepancies were discussed. If the two researchers coded the variables differently, they came to an agreement about the appropriate code. Note that "discrepancy" is defined as an instance in which two different codes were encountered by the researchers and a discussion or debate was required in order to decide upon the appropriate code. This does not include non-matching codes which could be resolved simply by referring back to the original study. Instead, discrepancies needed subjective discussion over the nature of a variable in the article. For tier-1, the researchers encountered a discrepancy once every 33 codes (about 3% of the time). In summary, three discrepancies were encountered:

- **Expertise**: the expertise needed to administer an assessment tool was rarely described in detail in the publications, meaning the researchers had to use their own personal judgment of the expertise needed. In order to come to agreement when codes did not match, it was decided to err on the conservative side, meaning the expertise needed for each tool was estimated as higher, rather than lower, given that the staff administering the tools for MTO will not necessarily have any psychological or medical training. For example, if a cognitive screening tool used perceptual analysis tests (in such tests participants have to explain their perceptions of an image far away and up close) or involved any type of physical/medical examination (e.g., eye exam), researchers categorized this as high expertise and, as a consequence, the tool would be excluded.
- **Feasibility:** similar to expertise, feasibility was often not described in the publications. Researchers, when in disagreement, decided to once again err on the side of caution and estimate the difficulty of instruction to participants who are taking the test as well as the difficulty in scoring the test

as higher, rather than lower, to ensure tools that were not feasible to administer would not be included in the meta-analysis.

> **Age range:** not every publication was specific in mentioning the age range of the participants tested, but instead would use terms such as "older drivers" or "elderly". Therefore, age was described using a variety of different formats, i.e., age range, minimum age or mean.

After using these tier-1 variables to select evaluations of tools that satisfy MTO's requirements, 27 evaluation outcomes remained (coming from 15 different studies). The same two independent researchers used the same coding process as previously discussed for tier-2 variables. When comparing notes between both researchers, there were very few discrepancies this time. The researchers encountered a discrepancy once every 65 codes (about 1.5% of the time). Tier-2 variables included more detailed information, most notably the outcome measures of the original studies.

The 27 evaluation outcomes of tools selected for inclusion are identified in the table below. Note that their identification number used throughout this report is listed in the second column.

Table 3: Overview of evaluation outcomes selected for inclusion in the meta-analysis and their identification number

Cognitive screening tool	ID
Rey-Osterrieth Complex Figure Test - Copy	2c
Rey-Osterrieth Complex Figure Test - Immediate	2d
Clock Drawing test	8b
Digit Symbol Substitution Test	32c
Visual Form Discrimination Test	32d
Maze Task - time	33
Maze Task - time	38c
Single Letter Cancellation Test - Error	38e
Visual Form Discrimination Test - total score	38f
Visual Form Discrimination Test - peripheral errors	38g
Probable Risk Checklist	43
Digit Symbol Substitution Test	62
Rey-Osterrieth Complex Figure Test - Copy	69a
Visual Form Discrimination Test	69d
Digit Symbol Substitution Test	69e
Maze Task - time	78a
Clock Drawing test	78b
Eight item informant interview	78e
Self-Report Questionnaire	94
Clock Drawing test	96
Clock Drawing test	103a
Traffic Sign Recognition Test	107
Charron Test - Error	108b
Single Letter Cancellation - Error	108g
Double Letter Cancellation Test - Error	108i
AARP Reaction Time	113c
Traffic Sign Recognition Test	113d

The 15 studies from which these evaluation outcomes were obtained are:

- Goode, K.T., Ball, K.K.,Sloane, M., Roenker, D.L., Roth, D.L., Myers, R.S., Owsley, C. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings*, 5(4): 425-440. (ID=2c, 2d)
- McCarthy, D.P. (2005). Outcomes Evaluation of the Assessment of Driving Related Skills (ADReS). Gainesville, FL.: University of Florida. (ID=8b)
- Szlyk, J.P., Myers, L., Zhang, Y., Wetzel, L., Shapiro, R. (2002). Development and assessment of a neuropsychological battery to air in predicting driving performance. *Journal of Rehabilitation Research and Development*, 39(4): 483-496. (ID=32c, 32d)
- Snellgrove, C.A. (2005). Cognitive Screening for the Safe Driving Competence of Older People with Mild Cognitive Impairment or Early Dementia. Civic Square, Australian Capital Territory: Australian Transport Safety Bureau. (ID=33)
- Whelihan, W.M., DiCarlo, M.A., Paul, R.H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology*, 20: 217-228. (ID=38c, 38e, 38f, 38g)
- Helander, C.J. (1986). Development and Evaluation of a Risk Assessment Strategy for Medically Impaired Drivers. Volume 8 of an Evaluation of the California Drunk Driving Countermeasure System. Sacramento, CA.: California Office of Traffic Safety. (ID=43)
- Lafont, S., Marin-Lamellet, C., Paire-Ficout, L., Thomas-Anterion, C., Laurent, B., Fabrigoule, C. (2010). The Wechsler Digit Symbol Substitution Test as the Best Indicator of the Risk of Impaired Driving in Alzheimer Disease and Normal Aging. *Dementia and Geriatric Cognitive Disorders*, 29(2): 154-163. (ID=62)
- Bliokas, W., Taylor, J.E., Leung, J., Deane, F.P. (2011). Neuropsychological assessment of fitness to drive following acquired cognitive impairment. *Brain Injury*, 25(5): 471-487. (ID=69a, 69d, 69e)
- Carr, D.B., Barco, P.P., Wallendorf, M.J., Snellgrove, C.A., and Ott, B.R. (2011). Predicting road test performance in drivers with dementia. *Journal of the American Geriatrics Society*, 59(11): 2112-2117. (ID=78a, 78b, 78e)
- Eby, D.W., and Molnar, L.J. (2001). Older drivers: A comparison of self-reported driving-related decisions with observed driving problems. In: Proceedings of the 45th Annual Conference of the Association for the Advancement of Automotive Medicine, pp. 413-415. Barrington, IL.: Association for the Advancement of Automotive Medicine. (ID=94)
- Freund, B., Gravenstein, S., Ferris, R., Burke, B.L., and Shaheen, E. (2005). Drawing clocks and driving cars: Use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine*, 20(3): 240-244. (ID=96)

- Oswanski, M.F., Sharma, O.P., Raj, S.S., Vassar, L.A., Woods, K.L., Sargent, W.M., and Pitock, R.J. (2007). Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine and Rehabilitation*, 86(3): 190-199. (ID=103a)
- MacGregor, J.M., Freeman, D.H., Jr., and Zhang, D. (2001). A traffic sign recognition test can discriminate between older drivers who have and have not had a motor vehicle crash. *Journal of the American Geriatrics Society*, 49(4): 466-469. (ID=107)
- Mazer, B.L., Korner-Bitensky, N.A., and Sofer, S. (1998). Predicting ability to drive after stroke. *Archives of Physical Medicine and Rehabilitation*, 79(7): 743-750. (ID=108b, 108g, 108i)
- Stutts, J.C., Stewart, J.R., and Martell, C. (1998). Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention*, 30(3): 337-346. (ID=113c, 113d)

Not all of these studies provided enough information regarding their outcome measures that was required to conduct the meta-analysis. More precisely, the following studies did not provide information such as standard errors, confidence intervals (CI), or standard deviations:

- McCarthy, D.P. (2005). Outcomes Evaluation of the Assessment of Driving Related Skills (ADReS). Gainesville, FL.: University of Florida. (ID=8b)
- Szlyk, J.P., Myers, L., Zhang, Y., Wetzel, L., Shapiro, R. (2002). Development and assessment of a neuropsychological battery to air in predicting driving performance. *Journal of Rehabilitation Research and Development*, 39(4): 483-496. (ID=32c, 32d)
- Whelihan, W.M., DiCarlo, M.A., Paul, R.H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology*, 20: 217-228. (ID=38c, 38e, 38f, 38g)
- Helander, C.J. (1986). Development and Evaluation of a Risk Assessment Strategy for Medically Impaired Drivers. Volume 8 of an Evaluation of the California Drunk Driving Countermeasure System. Sacramento, CA.: California Office of Traffic Safety. (ID=43)
- Eby, D.W., and Molnar, L.J. (2001). Older drivers: A comparison of self-reported driving-related decisions with observed driving problems. In: Proceedings of the 45th Annual Conference of the Association for the Advancement of Automotive Medicine, pp. 413-415. Barrington, IL.: Association for the Advancement of Automotive Medicine. (ID=94)
- Stutts, J.C., Stewart, J.R., and Martell, C. (1998). Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention*, 30(3): 337-346. (ID=113c, 113d)

TThe final set of nine remaining studies and their evaluation outcomes of tools that satisfied the inclusion criteria were included in the meta-analysis. These nine studies (listed below) collectively produced 15 evaluation outcomes (or data points; also listed below). Note that "published in a peer-reviewed journal" has not been used as an inclusion/exclusion criterion in this review. Instead, the nine remaining studies were assessed individually and no critical flaws were identified that would warrant their exclusion. As such, the already small dataset of nine studies did not become smaller still, which would make the meta-analysis more challenging, if not problematic – one study included in our final selection of studies (id 33) has not been published in a peer-reviewed journal.

> Goode, K.T., Ball, K.K.,Sloane, M., Roenker, D.L., Roth, D.L., Myers, R.S., Owsley, C. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings*, 5(4): 425-440. (ID=2c, 2d)

> Snellgrove, C.A. (2005). Cognitive Screening for the Safe Driving Competence of Older People with Mild Cognitive Impairment or Early Dementia. Civic Square, Australian Capital Territory: Australian Transport Safety Bureau. (ID=33)

> Lafont, S., Marin-Lamellet, C., Paire-Ficout, L., Thomas-Anterion, C., Laurent, B., Fabrigoule, C. (2010). The Wechsler Digit Symbol Substitution Test as the Best Indicator of the Risk of Impaired Driving in Alzheimer Disease and Normal Aging. *Dementia and Geriatric Cognitive Disorders*, 29(2): 154-163. (ID=62)

> Bliokas, W., Taylor, J.E., Leung, J., Deane, F.P. (2011). Neuropsychological assessment of fitness to drive following acquired cognitive impairment. *Brain Injury*, 25(5): 471-487. (ID=69a, 69d)

> Carr, D.B., Barco, P.P., Wallendorf, M.J., Snellgrove, C.A., and Ott, B.R. (2011). Predicting road test performance in drivers with dementia. *Journal of the American Geriatrics Society*, 59(11): 2112-2117. (ID=78a, 78e)

> Freund, B., Gravenstein, S., Ferris, R., Burke, B.L., and Shaheen, E. (2005). Drawing clocks and driving cars: Use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine*, 20(3): 240-244. (ID=96)

> Oswanski, M.F., Sharma, O.P., Raj, S.S., Vassar, L.A., Woods, K.L., Sargent, W.M., and Pitock, R.J. (2007). Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine and Rehabilitation*, 86(3): 190-199. (ID=103a)

> MacGregor, J.M., Freeman, D.H., Jr., and Zhang, D. (2001). A traffic sign recognition test can discriminate between older drivers who have and have not had a motor vehicle crash. *Journal of the American Geriatrics Society*, 49(4): 466-469. (ID=107)

> Mazer, B.L., Korner-Bitensky, N.A., and Sofer, S. (1998). Predicting ability to drive after stroke. *Archives of Physical Medicine and Rehabilitation*, 79(7): 743-750 (ID=108b, 108g, 108i)

The final dataset of 15 data points (see below) combined evaluated 10 different cognitive screening tools that satisfy each of MTO's previously specified inclusion criteria. Note that some of them were administered or scored in different ways (e.g., 2c and 2d are really the same tool but they are administered differently).

Table 4: Final dataset of 15 data points for the meta-analysis

ID	Cognitive screening tool
33	Maze Task - time
69d	Visual Form Discrimination Test
107	Traffic Sign Recognition Test
69a	Rey-Osterrieth Complex Figure Test - Copy
108i	Double Letter Cancellation Test - Error
2d	Rey-Osterrieth Complex Figure Test - Immediate
108b	Charron Test - Error
2c	Rey-Osterrieth Complex Figure Test - Copy
108g	Single Letter Cancellation - Error
78a	Maze Task - time
103a	Clock Drawing test
62	Digit Symbol Substitution Test
78b	Clock Drawing test
78e	Eight item informant interview
96	Clock Drawing test

5. RESULTS

The results section will identify the outcome of each step of the statistical analyses. To begin, the outcome measures of each included study are highlighted, which is then followed with the conversion process that was conducted for each outcome variable, converting all measures to the same scale. Once all outcomes are comparable, it is possible to analyze their effectiveness in predicting safe/unsafe driving behaviour. This is done using the forest plot (see Figure 2) and the meta-regression analysis. Issues relevant to a meta-analysis are discussed such as variance due to type of outcome measure and the methodology of each study. The tools are then ranked based on their predictive abilities of driving according to the meta-analysis.

Table 5 contains an overview of the outcome measures as reported in the original studies. As can be seen, the studies included in the meta-analysis used four different approaches to reporting outcomes: log-odds ratios (id=33); odds ratio (id=107, 62, 69a, 69d); means data (id=2c, 2d, 108b, 108g, 108i, 78a, 78b, 78e, 103a); and, correlational data (id 96).

Table 5: Overview of outcome measures as reported in the original studies included in the meta-analysis

ID	Model	Outcome measure	Outcome	Interpretation
2c	Means	Means (s.d.; N) Crashers	28.07 (7.7; 124)	Higher score on Rey-O Copy is associated with fewer crashes as measured by state-reported crashes
		Means (s.d.; N) Non-crashers	30.57 (6.27; 115)	
2d	Means	Means (s.d; N.) Crashers	8.25 (7.22; 124)	Higher score on Rey-O Immediate is associated with fewer crashes as measured by state-reported crashes
		Means (s.d.; N) Fail	10.79 (7.94; 115)	
108b	Means	Means (s.d.; N) Pass	3.9 (4.3; 33)	Failing the driving assessment is associated with more errors on the Charron Test
		Means (s.d.; N) Fail	5.4 (5.9; 51)	
108g	Means	Means (s.d.; N) Pass	2.5 (5.1; 33)	Failing a driving assessment is associated with more errors on the single digit cancellation test
		Means (s.d.; N) Fail	6.6 (13.0; 51)	
108i	Means	Means (s.d.; N) Pass	5.1 (5.5; 33)	Failing a driving test is associated with more errors on the double letter cancellation test
		Means (s.d.; N) Fail	6.2 (7.9; 51)	
33	Logistic regression model	Log odds ratio (s.e.)	-0.02 (0.01)	More time required to complete Maze Task corresponds to smaller chance of passing driving assessment
96	Correlation	Correlation (95%-CI)	0.68 (0.57-0.77)	Driving errors during same-day driving assessment correlates with CDT scoring scale
62	Logistic regression model	Odds ratio (95%-CI)	6.5 (2.1-20.1)	Low performance on WDSST (-5.8 correct) corresponds to increased risk of unsafe driving according to composite driving indicator
69a	Logistic regression model	Odds ratio (95%-CI)	0.882 (0.782-0.994)	Higher score on Rey-O Complex Figure Test corresponds to smaller chance of failing road test

ID	Model	Outcome measure	Outcome	Interpretation
69d	Logistic regression model	Odds ratio (95%-CI)	0.964 (0.843-1.103)	Higher score on Visual Form Discrimination corresponds to smaller chance of failing a road test
78a	Means	Means (s.d.; N) Pass	35.2 (12.3; 35)	Failing a test is associated with requiring more than 60 seconds to complete the Maze Test - time
		Means (s.d.; N) Fail	62.5 (43.9; 65)	
78b	Means	Means (s.d.;N) Pass	6.2 (1.2; 35)	A higher score on the Clock Drawing Test corresponds to a higher chance of passing a road test
		Means (s.d.;N) Fail	4.2 (2.5; 65)	
78e	Means	Means (s.d.; N) Pass	4.3 (1.5; 35)	Failing a driving test is associated with a higher score on the Eight Item Information Interview
		Means (s.d.; N) Fail	5.8 (1.6; 65)	
103a	Means	Means (s.d.; N) Capable	3.5 (0.843; 131)	A higher score on the Clock Drawing Test corresponds to a higher chance of being a capable driver
		Means (s.d.; N) Incapable	2.77 (1.159; 101)	

Table 6 contains a description of the dataset used in the meta-analysis after conversion of the outcome measures onto the same scale, i.e., log-odds ratio. Also, from Table 5 it is apparent that the direction of the original outcome measure was not the same in every study, i.e., a higher score on some cognitive tools may be indicative of unsafe driving behaviour, while other cognitive tools may indicate such unsafe behaviour by assigning lower scores. Therefore, the direction of each outcome measure was also made comparable by reversing it when necessary.

More precisely, the following conversions were adopted:

> Studies reporting log-odds and standard errors: 95%-CIs were calculated (point estimate plus/minus 1.96 X s.e.);

> Studies reporting odds ratios and 95%-CIs: odds ratios were transformed into log-odds point estimates and log-odds of the 95%-CI. This was then used to calculate standard errors;

> Studies reporting means data: the Stata command -metan- was used to calculate Cohen's d point estimates and their 95%-CIs. These Cohen's d results were then converted into log-odds using the formula Log-odds=(Cohen's d) X (pi/sqrt(3)). The result was then used to calculate standard errors;

> Studies reporting correlations: The point estimate and its 95%-CI were transformed into Cohen's d using the formula Cohen's d=(2 X correlation)/sqrt(1-(correlation X correlation)). These Cohen's d results were then transformed into log-odds using the formula mentioned previously (see previous bullet). Standard errors were calculated from there. Note that converting the correlation and its 95%-CI into Cohen's d and then into log-odds did not yield a symmetric confidence interval around the resulting log-odds point estimate, so when calculating the standard error the average of the result with the lower bound on the one hand and the upper bound on the other was used.

After adopting these conversion rules, the effect size used during the analyses can be defined as the log-odds ratio of predicting unsafe driving behaviour; in the exponential form effect sizes greater than one in our study indicate that when cognitive screening tools predict unsafe driving behaviour, it is more likely that the driver will indeed exhibit unsafe driving behaviour. Conversely, effect sizes smaller than one would indicate that when cognitive screening tools predict unsafe driving behaviour, it would be more likely that the driver would exhibit safe driving behaviour.

Table 6 contains the following variables: id (an indication of which study the data come from; refer to the list of studies mentioned previously for a complete reference); age (age of subjects in study as reported in the original paper or report: a number followed by a plus-sign represents a minimum age; two numbers separated by a hyphen represent an age range; a number represents a mean age); N (the number of subjects in each study); cognitive screen (an indication of which cognitive screening tool is evaluated); log-odds (the log of the odds ratio after transformation of the original outcome measure; refer to Table 5 for a description of the original outcome measure per study); and, se (standard error of the log-odds). Note that information regarding education was not always available. Regarding gender, all studies used a mix of males and females.

Table 6: Age, sample size, log-odds ratios, and their standard errors used to evaluate the predictive value of ten cognitive screening tools

ID	Cognitive screen	Age	N	log-odds	se
103a	Clock Drawing Test	55+	232	1.333639	.2480941
107	TSRT	65-91	120	.1278	.0653316
108b	Charron test - error	61	84	.51016	.4071622
108g	SLC - error	61	84	.6989591	.4088679
108i	DLC - error	61	84	.2825575	.405809
2c	ROCF - Copy	55+	239	.643332	.236664
2d	ROCF - Immediate	55+	239	.6082014	.236468
33	Maze Task - time	65+	115	.02	.01
62	DSST	65-85	76	1.8718	.5762244
69a	ROCF - copy	61	104	.1256	.061199
69d	VFDT - total score	61	104	.0367	.0685714
78a	Maze Task -time	74.2	96	1.367534	.3926136
78b	Clock Drawing Test	74.2	98	1.694839	.399073
78e	Eight Item Interview	74.2	99	1.737323	.400004
96	Clock Drawing Test	60+	119	3.36433	.474811

Figure 2 contains a forest plot of a random-effects meta-analysis of the effectiveness of cognitive screening tools in predicting driving performance of elderly drivers. Note that a random-effects model is justified, not only because we are pooling information from different studies published in the literature using different methodologies, but also because we are pooling information from different cognitive screening tools. It is, therefore, reasonable to assume that the effect sizes would truly differ across studies and tools. Consequently, a model that acknowledges this (like the random-effects model), rather than one that assumes there is only one true effect size (like the fixed-effect model) is the preferred one.

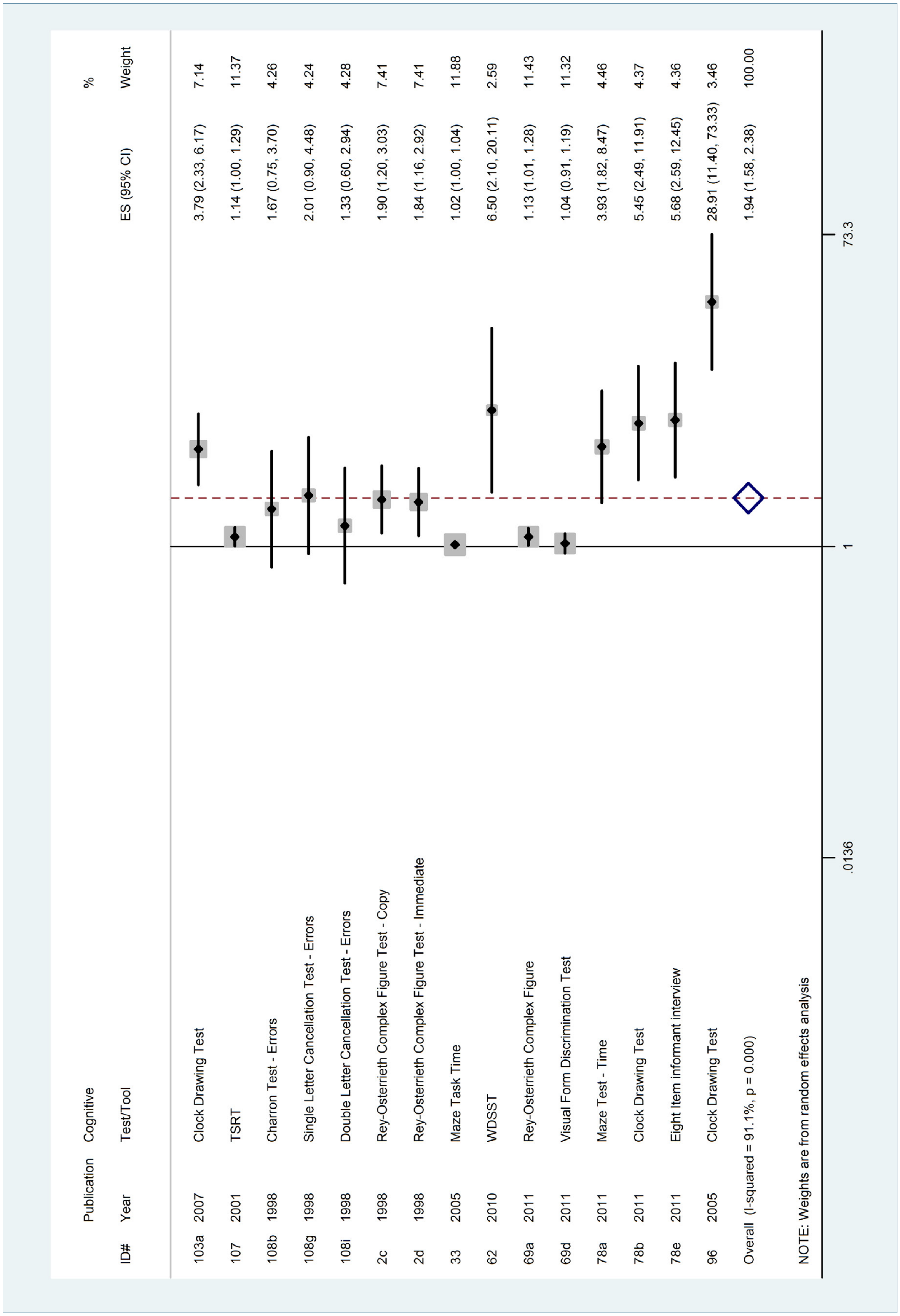
It can be seen that the pooled effect size, which represents the mean of the true effect sizes, is 1.940 in the exponential form, meaning that, on average, when cognitive screening tools predict a driver is unsafe, there is an increased chance of 94% ((1.940-1)X100) that this driver will exhibit unsafe driving behaviour, rather than safe driving behaviour (or, alternatively, if the cognitive screening tools predict that a driver is safe, on average, there is an increase of 94% in the likelihood that this driver will exhibit safe driving behaviour).

Note that “unsafe driving behaviour” in this instance refers to failing an on-road driving test, or a simulator driving test, or crashing, as the studies that are pooled in our meta-analysis used one of these three measures to assess the predictive value of a cognitive screening tool (conversely, “safe driving behaviour” refers to passing an on-road driving test, simulator driving test, or not crashing).

This pooled effect size is significant (Z=6.32, p=0.000). The 95% Confidence Interval (95%-CI) for this pooled effect size is 1.58-2.38. When transforming this result into Cohen’s d, this yields 0.37, which is between a small effect (d=0.2) and a medium effect (d=0.5), although it is somewhat closer to a medium effect.

The forest plot further shows that 91.1% of the variation in the summary effect is due to heterogeneity and that the test for heterogeneity is significant (p=0.000). This suggests – not surprisingly – that the between-study variance is not just the result of random variation but also of true differences in effect sizes.

Figure 2: Forest plot of the random-effects model



In an effort to help explain this heterogeneity, a meta-regression analysis has been conducted. It was hypothesized that five variables could potentially explain, at least partially, the heterogeneity. These variables were used as covariates in the meta-regression, i.e., they were included as independent variables in a random-effects model using Residual Maximum Likelihood (REML) estimation of effects. These variables are:

- **Typom** (type of outcome measure): 1 for studies that correlated performance on the cognitive screening tool to performance on a simulated driving test; 2 for studies that correlated performance on the cognitive screening tool to performance on an on-road driving test; and, 3 for studies that correlated this with state-reported crashes;
- **Blind** (blind assessment): 1 means the study includes a blind assessment whereby the driving evaluator does not know how the subject performed on the cognitive screening tool; and, 2 means there was no such blind assessment;
- **Outcome** (how outcome has been measured): 1 for studies that used a model directly calculating odds ratio; and, 2 for studies that did not use such a model but rather produced an outcome measure that had to be transformed into an odds ratio like a correlation or means data using one of the before-mentioned formulas;
- **Interval**: 1 for studies where the driving assessment is conducted prior to the cognitive assessment; 2 for studies where the driving assessment is conducted within 30 days following the cognitive assessment; and, 3 for studies where the driving assessment involves monitoring the driver's performance over a period of at least six months;
- **Control**: 1 for studies where no control group was used; and, 2 for studies where a safe or healthy group was compared to an unsafe or unhealthy group.

Table 7 provides an overview of the distribution of each covariates across studies. Periods in this table mean the data were missing (e.g., for study 96 the value of the variable interval is not known).

Table 7: Distribution of five covariates across studies included in the meta-analysis

ID	Typom	Blind	Outcome	Interval	Control
33	2	1	1	2	1
69d	2	2	1	2	1
107	3	2	1	3	2
69a	2	2	1	2	1
108i	2	1	2	2	1
2d	3	1	2	3	1
108b	2	1	2	2	1
2c	3	1	2	3	1
108g	2	1	2	2	1
78a	2	1	2	2	1
103a	2	.	.	2	1
62	2	1	1	2	2
78b	2	1	.	2	1
78e	2	1	2	2	1
96	1	2	2	.	1

One covariates in particular (typom) was found to explain a rather large proportion of the variance (approximately 47%; see below). Other models competing with this model did not explain the variance as well (note that the variable outcome did explain approximately 22% of the heterogeneity). Models with more than one covariates were also less satisfactory (diagnostics of the model with typom and outcome suggest the data do not fit the model as well). Given the low number of data points in our sample (15), including more than one covariate becomes problematic anyway – indeed, when including more than one independent variable, often at least one category of one of the variables was dropped due to co-linearity. Results of the model with typom are shown in Figure 3.

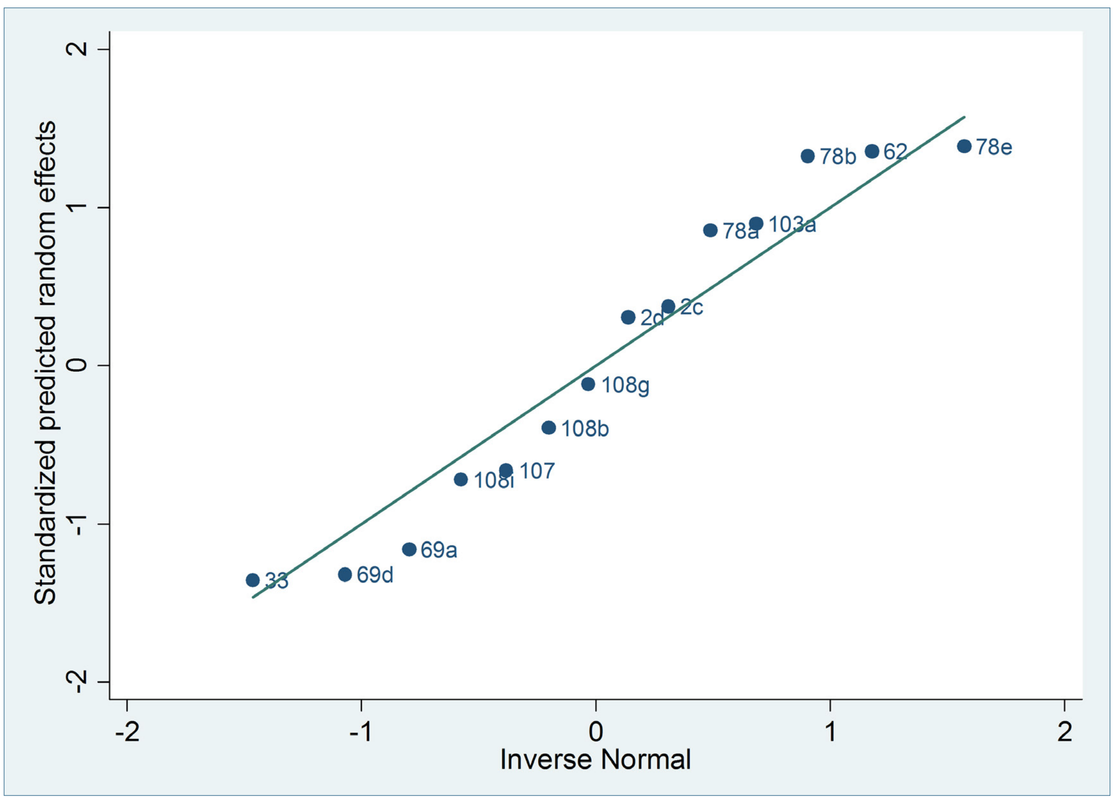
As can be seen, according to this model 88.08% of the variance is due to heterogeneity and the remaining 12% is due to within-study sampling variability. The proportion of between-study variance explained by this model is 47%. In other words, by introducing typom in the model the relative reduction in between-study variance is 47%, or, typom explains 47% of the between-study variance. The model is significant according to the joint test for all covariates (Model F=6.10, p=0.0149). Both dummy variables representing the variable outcome are also significant (p=0.007 and p=0.005, respectively). As such, the model suggests there is evidence for an association between the dummy variables and the outcome variable.

Figure 3: Results from a random-effects meta-regression with one covariate (typom; modeled as two dummies comparing their effect to the reference category= studies using a simulated driving test)

Meta-regression				Number of obs = 15		
REML estimate of between-study variance				tau2 = .3572		
% residual variation due to heterogeneity				I-squared_res = 88.08%		
Proportion of between-study variance explained				Adj R-squared = 47.04%		
Joint test for all covariates				Model F(2,12) = 6.10		
With Knapp-Hartung modification				Prob > F = 0.0149		
log-odds	**exp(b)**	**Std. Err.**	**t**	**P > | t |**	**[95% Conf. Interval]**	
_Itypom_2	.0756377	.059788	-3.27	0.007	.0135138	.4233509
_Itypom_3	.0539522	.0455932	-3.45	0.005	.0085579	.3401342

The Normal probability plot of the standardized shrunken residuals of this model, shown in Figure 4, suggests that the assumption of normal random-effects is adequate, and there are no notable outliers (i.e., no standardized shrunken residuals greater than two).

Figure 4: Normal probability plot of the random-effects model with one covariate (typom)



Results from this meta-regression analysis suggest that a large portion of the variation between studies, i.e., heterogeneity, is due to the fact that we have combined studies that used a different methodology to study the predictive value of cognitive screening tools on the driving performance of elderly drivers. In other words, there is evidence to suggest that there is no one effect size for the predictive value of the screening tools in our meta-analysis, but there are different effect sizes, and the effect size depends on whether on-road driving assessments, simulator driving assessments, or crash records are used in the evaluation.

Note that not all the between-study variance is explained by this covariate, which means there are other, unknown reasons why effect sizes differ. One other – logical – reason for heterogeneity would be the actual difference between the screening tools. Indeed, given the differences between these tools in terms of theory and implementation aspects, one would expect variation between their effect sizes. However, this hypothesis cannot be tested due to a small sample size (there are 15 data points in our sample and 10 different screening tools, which is not sufficient to use 'screening tool' as a covariate). Finally, other methodological differences between the 15 evaluation outcomes included in our meta-analysis can also help explain heterogeneity, but it was not possible to demonstrate this with our sample.

To better illustrate the effect of the covariate typom on the effect size, a subgroup analysis has been conducted representing the effect sizes for each category of typom. The results are shown in Figure 5 and are included for illustrative purposes only (a more accurate estimate of effect size by category of typom is provided below).

Figure 5: Subgroup random-effects analysis by categories of typom (1=studies using simulators; 2=studies using on-road driving tests; 3=studies using state-reported crashes)

Study	ES	[95% Conf. Interval]		% Weight
2				
Clock Drawing Test (2007)	3.795	2.334	6.171	7.14
Charron Test - Errors (1998)	1.666	0.750	3.700	4.26
Single Letter Cancellation Test - Errors (1998)	2.012	0.903	4.483	4.24
Double Letter Cancel Test (1998)	1.327	0.599	2.939	4.28
Maze Task Time (2005)	1.020	1.000	1.040	11.88
WDSST (2010)	6.500	2.101	20.110	2.59
Rey-Osterrieth Complex Figure Test (2011)	1.134	1.006	1.278	11.43
Visual Form Discriminination Test (2011)	1.037	0.907	1.187	11.32
Maze Test - Time (2011)	3.926	1.818	8.474	4.46
Clock Drawing Test (2011)	5.446	2.491	11.906	4.37
Eight Item Informant Interview (2011)	5.682	2.594	12.445	4.36
Sub-total D+L pooled ES	1.824	1.444	2.304	70.34
3				
TSRT (2001)	1.136	1.000	1.292	11.37
Rey-Osterrieth Complex Figure Test (1998)	1.903	1.197	3.026	7.41
Rey-Osterrieth Complex Figure Test (1998)	1.837	1.156	2.920	7.41
Sub-total D+L pooled ES	1.503	1.013	2.232	26.19
1				
Clock Drawing Test (2005)	28.914	11.401	73.329	3.46
Sub-total D+L pooled ES	28.914	11.401	73.329	3.46
Overall D+L pooled ES	1.940	1.580	2.383	100.00
Significance test(s) of ES=1				
2	z = 5.04		p = 0.000	
3	z = 2.02		p = 0.043	
1	z = 7.09		p = 0.000	
Overall	z = 6.32		p = 0.000	

Note that the effect size according to Figure 5 is smallest for tools that were evaluated using state-reported crashes (1.503), followed by tools evaluated in studies using on-road driving tests (1.824). The largest effect is for studies using simulated driving tests (28.914). Also, it is noteworthy that each of the three effect sizes is significant as can be seen both from the 95%-CI as well as the significance tests at the bottom of the Figure.

Finally, one might express concern about the very large effect of studies using simulated driving tests, especially because this result is based on one study only. However, this effect size has a very wide 95%-CI and this lack of precision is taken into account in the meta-analysis when weighing the results from each individual study. As can be seen, a smaller weight has been assigned to this study (3.46%; see also Figure 2).

The table below (Table 8) provides an overview of the effect sizes by category of typom based on the predicted values from the meta-regression and their predicted standard errors. The predicted value is presented in the exponential form (fitval) along with its 95%-CI (Lower and Upper). This provides a more accurate estimate of the group effects. The results have been sorted from smaller predicted values to larger predicted values.

Table 8: Predicted values in exponential form based on the meta-regression with typom as covariate

ID	Cognitive screen	Typom	Fitval	Lower	Upper
107	TSRT	3	1.559978	.7664511	3.175065
2d	ROCF - Immediate	3	1.559978	.7664511	3.175065
2c	ROCF - Copy	3	1.559978	.7664511	3.175065
78a	Maze Task - time	2	2.186997	1.462549	3.270288
69d	VFDT - total score	2	2.186997	1.462549	3.270288
33	Maze Task - time	2	2.186997	1.462549	3.270288
108i	DLC - error	2	2.186997	1.462549	3.270288
108g	SLC - error	2	2.186997	1.462549	3.270288
78e	Eight Item Interview	2	2.186997	1.462549	3.270288
69a	ROCF - copy	2	2.186997	1.462549	3.270288
78b	Clock Drawing Test	2	2.186997	1.462549	3.270288
62	DSST	2	2.186997	1.462549	3.270288
108b	Charron Test - error	2	2.186997	1.462549	3.270288
103a	Clock Drawing Test	2	2.186997	1.462549	3.270288
96	Clock Drawing Test	1	28.91411	6.476635	129.0833

As can be seen in Table 8, the lowest value is again for studies using crashes to evaluate the predictive value of cognitive screening tools (typom=3). This value is now 1.559 instead of 1.503 as reported in Figure 5. However, the 95%-CI is wider in Table 8 and includes values smaller than one, suggesting that findings from these studies are not significant according to the regression model. For studies using an on-road driving test (typom=2), the result is now 2.187 rather than 1.824. The confidence interval does not contain one in Table 8, nor did it contain one in Figure 5 (meaning the result is significant according to both approaches). Finally, the result for studies using simulators (typom=1) is the same as in Figure 5, but the confidence interval is much wider in Table 8.

A comparable approach can be used to rank order tools, based on the prediction including random-effects, also known as the empirical Bayes estimates of the effects for each study. Table 9 provides an overview of these predictions. Studies have been sorted in ascending order and are presented in exponential form (empbayes), along with their 95%-CI (Lower, Upper). The rank ordering is fairly consistent with the results from the forest plot in Figure 2, although there are differences regarding the 95%-CI and some of the rankings (e.g., 69a and 107 are reversed, and so is 108b and 2d).

Table 9: Empirical Bayes estimates in exponential form based on the meta-regression with typom as covariate

ID	Cognitive screen	Typom	Empbayes	Lower	Upper
33	Maze Task - time	2	1.020419	.3162805	3.292189
69d	VFDT - total score	2	1.047482	.3265297	3.360244
107	TSRT	3	1.140585	.3541389	3.673514
69a	ROCF - copy	2	1.141583	.3554437	3.666437
108i	DLC - error	2	1.553199	.564237	4.275554
2d	ROCF - Immediate	3	1.79691	.571066	5.654137
108b	Charron Test - error	2	1.815742	.6601525	4.99418
2c	ROCF - Copy	3	1.852259	.588681	5.828051
108g	SLC - error	2	2.065971	.7519042	5.676569
78a	Maze Task - time	2	3.291008	1.186053	9.131742
103a	Clock Drawing Test	2	3.499566	1.162411	10.53583
62	DSST	2	3.846148	1.546854	9.563189
78b	Clock Drawing Test	2	4.110431	1.487148	11.3611
78e	Eight Item Interview	2	4.228985	1.530901	11.68221
96	Clock Drawing Test	1	28.91411	6.476635	129.0833

It is clear that the results from this meta-analysis alone are not sufficient to unambiguously select a cognitive screening tool as the best-performing one. For example, the lower bound of the Clock Drawing Test's 95%-CI (id=96) are below the upper bound of other cognitive screening tools, suggesting they may not be significantly different. However, this rank ordering can be augmented with other relevant information such as sensitivity, specificity, positive and negative predictive value, and area under the Receiver Operator Characteristic (ROC) curve. Table 10 contains such information whenever it was available in the studies included in the meta-analysis.

Table 10: Sensitivity, specificity, positive predictive value, negative predictive value and area under ROC curve

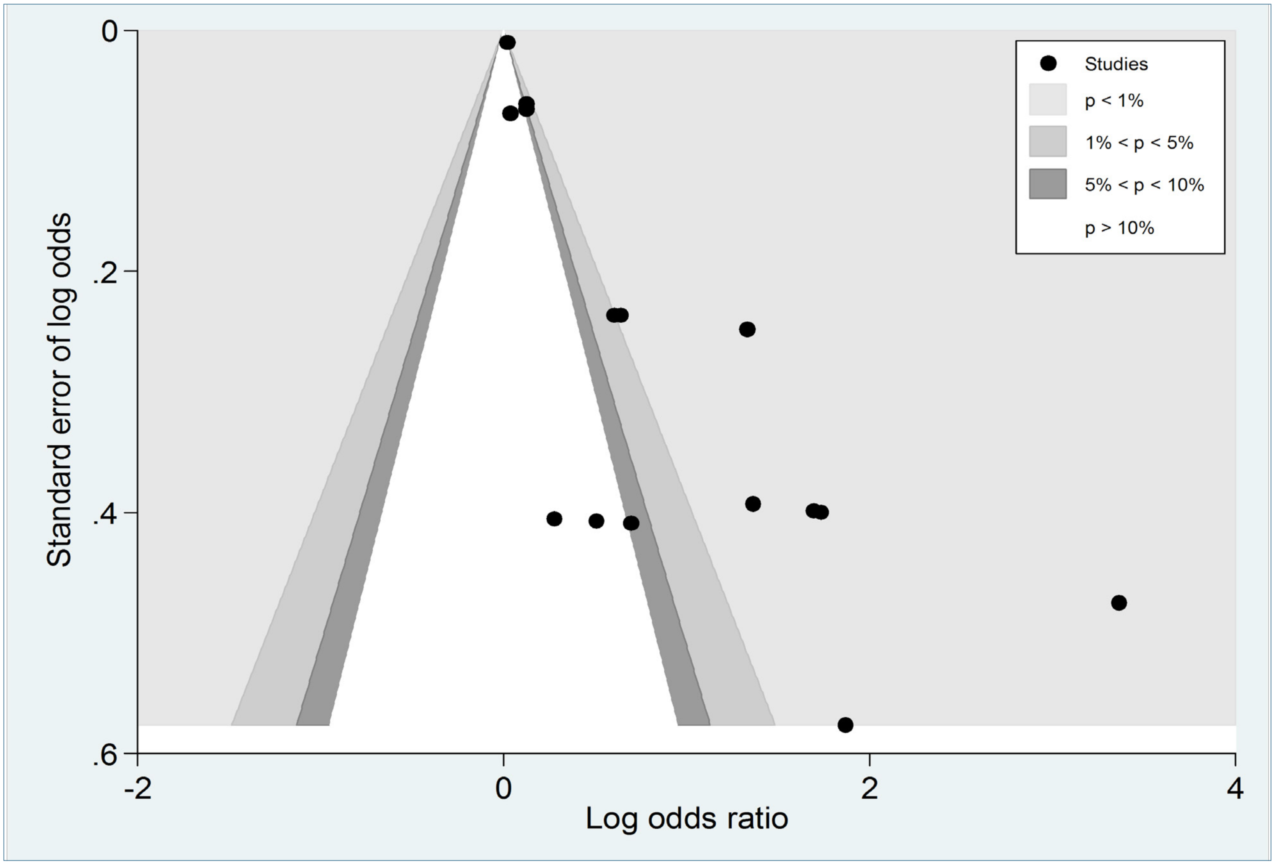
ID	Cognitive Screening Tool	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	ROC (AUC)
33	Maze Task - time	77.8	82.4	NA	NA	NA
62	DSST	91.7	81.2	NA	NA	NA
96	Clock Drawing Test	64.2	97.7	NA	NA	0.9
103a	Clock Drawing Test	70	65	NA	NA	NA
107	Traffic Sign Recognition Test	60	63	NA	NA	0.636
108b	Charron Test	NA	NA	72.4	45.5	NA
108g	SLC	NA	NA	78.9	44.6	NA
108i	DLC	NA	NA	64.9	42.6	NA

When combining information from Tables 9 and 10, one may consider choosing DSST or the Clock Drawing Test the best suited for GES. These two tests are in the top four performing tests according to Table 9. If public safety is considered more important, implying fewer false negative predictions of unsafe driving behaviour are preferable (rather than fewer false positives), DSST would be the tool of choice, given its high sensitivity (91.7%). If, however, concerns regarding the individual senior driver are considered a priority, implying fewer false positive predictions of unsafe driving behaviour are preferable, a better choice would be the Clock Drawing Test, given its high specificity (97.7%), at least according to the results from study 96.

Concerns regarding bias, notably publication bias are addressed based on the following figure. Figure 6 contains a contour-enhanced funnel plot. This is a classic funnel plot to test for bias, but it is enhanced by adding contours of statistical significance to aid in the interpretation of the plot. "If studies appear to be missing in areas of low statistical significance, then it is possible that the asymmetry is due to publication bias. If studies appear to be missing in areas of high statistical significance, then publication bias is a less likely cause of the funnel asymmetry" (Palmer et al. 2009: p. 124).

It is clear from Figure 6 that studies are missing from the left-hand side of the plot, suggesting a strong asymmetry. Note that the centre of the plot is very close to a log-odds ratio of zero (or odds ratio of one). This is because the effect with the greatest weight is very close to zero (study id=33; odds ratio=1.020; weight=11.88%). Studies on the left-hand side of the plot would suggest cognitive screening tools have no predictive value or, worse, that they would significantly predict the opposite, i.e., prediction of safe driving performance according to the screening tool would correlate with unsafe driving behaviour.

Figure 6: Contour-enhanced funnel plot



Upon closer inspection it can be seen that the area where studies are missing in Figure 6 includes both regions of high and low statistical significance on the left-hand side. On the right-hand side, there are studies in all regions of significance levels, although the majority is in the region of statistical significance. This means there is likely publication bias in our dataset but that publication bias alone cannot explain the asymmetry (publication bias refers to the fact that studies with non-significant findings have a smaller chance of publication than studies with significant findings).

A formal test for small-study effects using Egger's test is available in Figure 7. This test also confirms that there is bias (bias coefficient=3.07, t=5.87, p=0.000).

Figure 7: Egger's test for small-study effects

Egger's test for small-study effects: Regress standard normal deviate of intervention effect estimate against its standard error						
Number of studies = 15				Root MSE = 1.819		
Std_Eff	**Coef.**	**Std. Err.**	**t**	**P > | t |**	**[95% Conf. Interval]**	
slope	-.0157056	.0194851	-0.81	0.435	-0.578004	.0263893
bias	3.070871	.5230928	5.87	0.000	1.940797	4.200944
Test of HO: no small-study effects				P = 0.000		

Further interpretation of findings regarding bias appears to coincide with expectations regarding publication of studies in this field. First, it is not surprising that there would be publication bias, as only cognitive screening tools with significant predictive value would be perceived to be of interest to the field and, as a consequence, have a higher chance of publication. This likely explains why most of the data points on the right-hand side are in the significant regions.

Second, it is perhaps even less surprising that studies that would have found a result indicating that the cognitive tool predicts the opposite would never have been published, especially when such a result would be significant. In this regard, findings about the predictive value of cognitive screening tools are perhaps different from findings about many other treatments in that any significant finding about most treatments, good or bad, would be relevant to public health, and, as a consequence relevant to publish, whereas significant findings regarding tools that are opposite to expectations are not relevant as they are likely indicative of the tool's flawed design.

Another way to obtain a clearer perspective is to calculate the number of studies with a small, non-significant finding that would have to be conducted before the current finding would become insignificant. This is called a Fail Safe N. When applying Orwin's formula for Fail Safe N (N=15*((ln(1.940)*(sqrt(3)/pi))-0.2))/0.2), this number is 12.40, meaning that 12 studies with a small effect would have to be conducted before our finding (i.e., a significant, small to medium-sized pooled effect size of 1.940 on the odds ratio scale) would become insignificant.

In summary, it certainly appears that our database is biased according to these tests. This bias is the result not only from publication bias, but also other sources of bias, notably the inherently asymmetric distribution of performance of cognitive screening tools around the null-value. It is possible that after correcting for publication bias the summary effect would no longer be significant. This further confirms that the meta-analysis alone will not be sufficient to provide the evidence supporting the selection of one single tool. Such a choice will have to be further bolstered with other information about individual tests such as sensitivity, specificity, and area under the ROC curve (see Table 10).

As previously mentioned, another concern with this meta-analysis relates to the fact that several evaluation outcomes came from the same study (e.g., 69a and 69d come from study 69; 2c and 2d come from study 2). These findings are not independent of one another and may have affected results regarding precision of the summary effect (if the outcomes in these studies are positively correlated, precision may have been overestimated, meaning that our summary significant effect may not be significant after all). To further investigate the effect of correlation between data points from the same study, a multilevel model has been fitted using the MLwiN software package. In this model, level 1 has been defined as the level of outcome measures, which are nested in level 2, the studies in which these outcome measures are published (for example, study outcome measures 69a and 69b are nested in study 69).

The results from this model are shown below.

Figure 8: A multilevel model

$$\text{logodds}_{ij} \sim \text{N}(XB, \Omega)$$

$$\text{logodds}_{ij} = \beta_{0ij}\text{CONS} + -2.487(0.297)\text{typom\_2}_j + -2.989(0.176)\text{typom\_3}_j$$

$$\beta_{0ij} = 3.364(0.000) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.354(0.106) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 0.081(0.018) \end{bmatrix}$$

$$-2{*}loglikelihood(IGLS\ Deviance) = 14.308(15 \text{ of } 15 \text{ cases in use})$$

According to this model, the effect size if typom is equal to 1 (for studies using simulators) is 3.364, or, in exponential form 28.90, which is almost exactly the same as the result obtained in our meta-regression analysis in Table 8. According to this multilevel model, this result is also significant, as it is at least twice as large as its standard error (value following the coefficient in between brackets). This model further shows that the summary effect when typom is equal to 2 is 3.364-2.487 or, in exponential form, 2.41 compared to 2.19 in Table 8. Finally, the effect in exponential form for typom equal to 3 is 1.46, compared to 1.56 in Table 8.

In other words, the results from the multilevel model, i.e., the model that accounts for the fact that several of our data points come from the same studies, are very consistent with the results from our model that does not account for this.

6. DISCUSSION

6.1 Overview of the results

The random-effects meta-analysis provided evidence suggesting that cognitive screening tools that meet MTO's logistical requirements for GES can be used to predict driving performance. A small to medium-sized, significant pooled effect of 1.940 was found, meaning that on average, when cognitive screening tools predict a driver is unsafe, there is a 94% greater chance that this driver will exhibit unsafe driving behaviour, rather than safe driving behaviour (or, alternatively, if the cognitive screening tools predict that a driver is safe, on average, it is 94% more likely that this driver will exhibit safe driving behaviour, rather than unsafe driving behaviour). Note that unsafe driving behaviour refers to unsafe performance during a road test, a simulator driving test, or crashing as evidenced by state-reported crashes (it follows that safe driving behaviour refers to safe performance during a road test, a simulator driving test, or not crashing).

Strong evidence for heterogeneity was found in our sample of studies, confirming that our choice for a random-effects model rather than a fixed-effects model is appropriate. Given that different studies from the literature have been summarized in this meta-analysis, that these studies used different methodologies, and that the cognitive screening tools themselves are also different, it is not surprising that there is heterogeneity. As such, our pooled effect size represents an average of the different true effect sizes, rather than one true effect size.

When using the information regarding how unsafe driving behaviour was measured across studies (i.e., using on-road driving tests, simulator driving tests or state-reported crashes) as a covariate in a random-effects meta-regression analysis, this covariate explained much of the observed between-study variance. Results from this meta-regression can be used to rank order cognitive screening tools using empirical Bayes estimates in an effort to identify better-performing tools. This exercise made clear that the results from this meta-analysis alone cannot be used for this purpose. It is recommended it is used in combination with other information such as sensitivity, specificity, and area under the ROC curve. However, such information is not available for all the tools included in this meta-analysis. Partial information was available in eight studies. Using this partial information, and depending on one's preferences regarding concerns for the public versus the individual, DSST or Clock Drawing can be considered as preferred cognitive screening tools.

Regardless of results based on the rank ordering of tools, one also has to consider the fact that other tools may be preferred depending on practical considerations. While all tools included in this study were selected based on a rigorous assessment using MTO's criteria for the use of such tools in a GES setting, it is clear that some of the included tools will fit better in a GES setting than others. As such, it has to be emphasized that all ten tools included in this meta-analysis passed the test using MTO criteria, and therefore, all ten of them could be considered for inclusion within the existing GES framework. The trade-off when using tools with a lower rank-order is that, while they may fit better from a practical point of view, they may be less effective in predicting unsafe or safe driving behaviour according to the results from this meta-analysis.

6.2 Limitations

Different sources of bias exist in our sample of studies. It is not surprising evidence for publication bias was found, and perhaps even less surprising that evidence for other sources of bias were found, given the asymmetric nature of performance of cognitive screening tools. An investigation of how this bias may have affected our findings suggests that our results are fairly robust. For example, our dataset did not exclusively consist of significant findings but did include some published results that were not significant, suggesting publication bias may be less pronounced. Also, the Fail Safe N, i.e., the number of studies with non-significant findings needed to overturn our significant finding is 12. While this number is not overly large, it means that a number of studies with non-significant findings almost equal to the number of studies conducted to date is required to overturn our significant finding.

Finally, one particular concern in our sample of evaluation outcomes is that several of these outcomes come from the same studies. Because such studies use the same sample of subjects to test the different cognitive screening tools, results about these tools are correlated, which can bias estimates of precision of the summary effect. A multilevel regression model was run, comparable to the one-level meta-regression and the results were consistent, suggesting that we can have confidence in the results from the meta-analysis.

7. CONCLUSION

In sum, the meta-analysis shows there is fairly robust evidence to suggest that cognitive screening tools have value in predicting driving behaviour. The effect is significant and low to medium-sized, and it differs, not only across different ways of measuring (un)safe driving behaviour, but likely also across cognitive screening tools themselves. It was not possible, however, to study in more detail any effect of the screening tools due to our small sample size. The results of this meta-analysis can be used in combination with other sources of knowledge to select a cognitive screening tool.

Ultimately, this selection will depend on choices that are more of a practical nature (e.g., Can the tool be efficiently and practically be incorporated into the GES framework?) and political nature (e.g., What is the trade-off between concerns for the public versus concerns for the individual?). Based on this investigation, and the information at hand to date, we recommend that any of the ten tools identified in this meta-analysis can be considered for use in a GES setting, acknowledging that not all of them are equally effective in predicting unsafe or safe driving behaviour.

REFERENCES

Akinwuntan, A.E., De Weerdt, W., Feys, H., Baten, G., Arno, P., and Kiekens, C. (2005). The Validity of the Road Test After Stroke. *Archives of Physical Medicine and Rehabilitation* 86(3): 421-426.

Alzheimer's Association. (2012). AD8 Dementia Screening Interview. Alz.org. Retrieved from: http://www.alz.org/documents_custom/ad8.pdf.

Anstey, K.J., Wood, J., Lord, S., and Walker, J.G. (2005). Cognitive, sensory, and physical factors enabling driving safety in older adults. *Clinical Psychological Review* 25: 45-65.

Ball, K.K., Roenker, D.L., Wadley, V.G., Cissell, G., Matthews, M., Ball, D., Vance, D., Frankel, M., and McConnell, K. (2005). Longitudinal assessment of older drivers in a DMV setting. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 57. Iowa City, IA.: University of Iowa Public Policy Center.

Ball, K.K., Roenker, D.L., Wadley, V.G., Edwards, J.D., Roth, D.L., McGwin, G. Jr., Raleigh, R., Joyce, J.J., Cissell, G.M., and Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? *Journal of the American Geriatrics Society* 54(1): 77-84.

Bettcher, B.M., Libon, D.J., Kaplan, E., Swenson, R., and Penney, D.L. (2013). Digit Symbol Substitution Test. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds. In Chief). *Encyclopedia of Clinical Neurology Part 4*. New York: Springer; 2011. p. 849-853.

Bliokas, W., Taylor, J.E., Leung, J., and Deane, F.P. (2011). Neuropsychological assessment of fitness to drive following acquired cognitive impairment. *Brain Injury* 25(5): 471-487.

Bryer, T. (2000). Characteristics of motor vehicle crashes related to aging. In K.W. Schaie and M. Pietrucha (Eds). Mobility and Transportation in the Elderly, 157-206. New York: NY: Springer Publishing Company Inc.

Byszewski, A.M., Molnar, F.J., and Aminzadeh, F. (2010). The Impact of Disclosure of Unfitness to Drive in Persons with Newly Diagnosed Dementia: Patient and Caregiver Perspectives. *Clinical Gerontologist* 33:152-163.

Canadian Stroke Network. ND. In Depth Review of SLCT. StrokeEngine. Retrieved from: http://strokengine.ca/assess/module_slct_indepth-en.html.

Canadian Stroke Network. ND. In Depth Review of DLCT. StrokeEngine. Retrieved from: http://strokengine.ca/assess/module_dlct_indepth-en.html.

Canham, R.O., Smith, S.L., and Tyrell, A.M. (2005). Automated Scoring of a Neuropsychological Test: The Rey Osterrieth Complex Figure. *Proceedings of the 26th EUROMICRO Conference* 2:2406.

Caplan, B., and Scultheis, M.T. (1998). An Interpretive Table for the Visual Form Discrimination Test. *Perceptual and Motor Skills* 87: 1203-1207.

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

Carr, D.B., and Ott, B.R. (2010). The Older Adult Driver with Cognitive Impairment: "It's a Very Frustrating Life". *Journal of the American Medical Association* 303(16): 1632-1641.

Classen, S., Wen, P.S., Velozo, C.A., Bédard, M., Winter, S.M., Brumback, B., and Lanford, D.N. (2012). Rater reliability and rater effects of the Safe Driving Behavior Measure. *American Journal of Occupational Therapy* 66(1): 69-77.

Cooper, P.J., Tallman, K. Tuokko, H., and Beattie, B.L. 1993. Vehicle Crash Involvement and Cognitive Deficit in Older Drivers. *Journal of Safety Research* 24: 9-17.

De Raedt, R., and Ponjaert-Kristoffersen, I. 2000. The relationship between cognitive/ neuropsychological factors and car driving performance in older drivers. *Journal of the American Geriatrics Society* 48: 1664-1668.

Dobbs, B.M., and Schopflocher, D. (2010). The introduction of a new screening tool for the identification of cognitively impaired medically at-risk drivers: the SIMARD A Modification of the DemTect. *Journal of Primary Care and Community Health* 1(2):119-127.

Dugbartey, A., Townes, B., and Mahurin, R. (2000). Equivalence of the Colour Trails Test and Trail Making Test in Nonnative English-Speakers. *Archives of Clinical Neuropsychology* 15(5): 425-431.

Dumitrescu, C.C. 2005. Exploratory factor analysis of the Gestalt Completion Test. *ProQuest Dissertation and Thesis* 30.

Eby, D.W., and Molnar, L.J. (2001). Older drivers: A comparison of self-reported driving-related decisions with observed driving problems. In: *Proceedings of the 45th Annual Conference of the Association for the Advancement of Automotive Medicine*, pp. 413-415. Barrington, IL.: Association for the Advancement of Automotive Medicine.

Evans, L. (2001). Age and fatality risk from similar severity impacts. *Journal of Traffic Medicine* 29: 10-19.

Ferreira, I.S., Simoes, M.R., and Maroco, J. (2012). The Addenbrooke's Cognitive Examination Revised as a potential screening test for elderly drivers. *Accident Analysis and Prevention*, 49:278-286.

Fonda S.J., Wallace R.B., Herzog A.R. (2001). Changes in driving patterns and worsening depressive symptoms among older adults. *J Gerontol*.; 56(6): S343-351.

Freeman E.E., Gange S.J., Munoz B., Wet S.K. (2006). Driving status and risk of entry into long-term care in older adults. *Am J Public Health*.; 96:1254-1259.

Freund, B., Gravenstein, S., Ferris, R., Burke, B.L., and Shaheen, E. (2005). Drawing clocks and driving cars: Use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine* 20(3): 240-244.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R.C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., Cummings, J.L., de Leon, M., Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M.C., Whitehouse, P., and Winbald, B. (2006). Mild Cognitive Impairment. Lancet 367(9518): 1262-1270.

Ginzburg, T., and Minichiello, R. (2011). License removal, driving assessments and the process of re-acquisition. Toronto, ON.: DriveAgain.

Golden, C. (2004). Stroop Colour and Word Test. Psychological Assessments and Services. Retrieved on January 21st, 2013. Retrieved from: http://www.mhs.com/product.aspx?gr=cli&pr od=stroop-cw&id=overview.

Goode, K.T., Ball, K.K., Sloane, M., Roenker, D.L., Roth, D.L., Myers, R.S., and Owsley, C. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings* 5(4): 425-440.

Helander, C.J. (1986). Development and Evaluation of a Risk Assessment Strategy for Medically Impaired Drivers. Volume 8 of an Evaluation of the California Drunk Driving Countermeasure System. Sacramento, CA.: California Office of Traffic Safety.

Howe, E. (2007). Initial Screening of Patients for Alzheimer's Disease and Minimal Cognitive Impairment. Update on Alzheimer's. July 2007. *Psychiatry* 24-27.

Human Resources and Skills Development Canada. (2012). Canadians in Context – Aging Population. Indications of Well Being in Canada. Retrieved from: http://www4.hrsdc.gc.ca/.3n dic.1+.4r@-eng.jsp?iid=33.

Hunt, L.A., Murphy, C.F., Carr, D., Duchek, J.M., and Buckles, V. (1997). Reliability of the Washington University Road Test. *Archives of Neurology* 54: 707-712.

Innes, C.R., Jones, R.D., Dalrymple-Alford, J.C., Hayes, S., Hollobon, S., Severinsen, J., Smith, G., Nicholls, A., and Anderson, T.J. (2007). Sensory-motor and cognitive tests predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences*, 260(1-2): 188-198.

Johnson, L. (2007). The ability of a visual-motor (Frostig) worksheet to predict the on-road performance of older drivers as evaluated by video and GPS. Paper presented at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Justiss, M.D., Mann, W.C., Stav, W., and 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.G., Bundy, A.C., and Clemson, L.M. (2008). Predicting fitness to drive using the visual recognition slide test (USyd). *American Journal of Occupational Therapy* 62(2): 187-197.

Kay, L.G., Bundy, A.C., and Clemson, L.M. (2009). Predicting fitness to drive in people with cognitive impairments by Using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation* 90(9): 1514-1522.

Kelly, R., Warke, T., and Steele, I. (1999). Medical restrictions to driving: the awareness of patients and doctors. *Post Graduate Medical Journal* 75: 537-539.

Lafont, S., Marin-Lamellet, C., Paire-Ficout, L., Thomas-Anterion, C., Laurent, B., and Fabrigoule, C. (2010). The Wechsler Digit Symbol Substitution Test as the Best Indicator of the Risk of Impaired Driving in Alzheimer Disease and Normal Aging. *Dementia and Geriatric Cognitive Disorders* 29(2): 154-163.

Konstantopoulos, K., Issidorides, M., and Spengos, K. (2013). A Normative Study of the Colour Trails Test in the Greek Population. *Applied Neuropsychology: Adult* 20: 47-52.

Lampman, D.V. (2002). Fun, Fun, Fun, 'Til Sonny (of the Government) Takes the T-Bird Away: Elder Americans and the Privilege to Be Independent. Symposium on Living Independently: Impact of Science and Technology on the Elderly. Comment. *Albany Law Journal of Science and Technology* 22: 863-917.

Langford, J., Fitzharris, M., Koppel, S., and Newstead, S. (2004). Effectiveness of Mandatory License Testing for Older Drivers in Reducing Crash Risk Among Urban Older Australian Drivers. *Traffic Injury Prevention* 5: 326-335.

Langford, J., Hakamies-Blomqvist, L., and Methordst, R. (2006). Older drivers do not have high crash risk –A replication of low mileage bias. *Accident Analysis and Prevention* 38: 574-578.

Lerner, N.D. (1993). Brake perception-reaction times of older and younger drivers. In: *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, pp. 206-210.

Lerner, N.D., Huey, R., McGee, H., and Sullivan, A. (1993). Older Driver Perception-Reaction Time for Intersection Sight Distance and Object Detection. Final Report. Washington D.C., Federal Highway Administration.

Li, G., Braver, E. R., and Chen, L. (2003). Fragility versus excessive crash involvement as determinants of high death rates pr. vehicle-mile of travel among older drivers. *Accident Analysis and Prevention* 35: 227–235.

MacGregor, J.M., Freeman, D.H., Jr., and Zhang, D. (2001). A traffic sign recognition test can discriminate between older drivers who have and have not had a motor vehicle crash. *Journal of the American Geriatrics Society* 49(4): 466-469.

Marottoli R.A., de Leon C.F.M., Glass T.A. (2000). Consequences of driving cessation: decreased out-of-home activity levels. *J Gerontol B Psychol Sci Soc Sci.*; 55:S334-340.

Martin, A.J., Marottoli, R., and O'Neill, D. (2010). Driving assessment for maintaining mobility and safety in drivers with dementia (Review). The Cochrane Collaboration. *The Cochrane Library* 10: 1-22.

Martinez, R. (1995). Older Drivers and Physicians. *Journal of the American Medical Association* 274(13): 1060.

Mathias, J.L., and Lucas, L.K. (2009). Cognitive predictors of unsafe driving in older drivers: a meta-analysis. *International Psychogeriatrics* 21(4), 637-653.

Mayhew, D.R., Simpson, H.M., an d Ferguson, S.A. (2006). Collisions Involving Senior Drivers: High-Risk Conditions and Locations. *Traffic Injury Prevention* 7: 117-124.

Mazer, B. L., Korner-Bitensky, N.A., and Sofer, S. (1998). Predicting Ability to Drive After Stroke. *Archives of Physical Medicine and Rehabilitation* 79(7): 743-750.

McCane, S.J. (2006). Test Review: Colarusso, R.P., & Hammill, D.D. (2003). "Motor-Free Visual Perception Test" (3rd ed.). Novata, CA: Academic Therapy Publications. *Journal of Psychoeducational Assessment* 24(3): 265-272.

McCarthy, D.P. (2005). Outcomes Evaluation of the Assessment of Driving Related Skills (ADReS). Gainesville, FL.: University of Florida.

McGwin, G., Owsley, C., and Ball, K. (1997). Visual and cognitive impairments and injurious motor vehicle crashes in the elderly. In: *Proceedings of the 41st Annual Conference of the Association for the Advancement of Automotive Medicine*, pp. 427-428. Des Plaines, IL.: Association for the Advancement of Automotive Medicine.

Meng, A. (2010). Age based population screening for fitness to drive does not produce any safety benefits. European Federation of Psychologists' Associations (EFPA). Brussels, Belgium.

Mevius, J. (2013). Elderly Driving Assessments. Assessment Comparison Matrix. Retrieved from: http://elderlydrivingassessments.com/assessment-comparison-chart.pdf.

Mezuk B, Rebok GW. (2008). Social integration and social support among older adults following driving cessation. *J Gerontol Series B Psychol Sci Social Sci*.; 63:S298-303.

Ontario Road Safety Annual Report (ORSAR). 2009. Driver Age Groups – Number Licensed, Collision Involvement and Percent Involved in Collisions. The People. Retrieved from: http://www.mt o.gov.on.ca/english/safety/orsar/orsar09/people.shtml.

Rapoport, M.J. Herrman, N., Molnar, F., Paula, A., Juurlink, D.N., Zagorski, B., Seitz, D., Morris, J.C., and Redelmeier, D.A. (2008). Psychotropic medications and motor vehicle collisions in patients with dementia. *Journal of the American Geriatric Society* 56(10): 1968-1970.

Ray, W.A. (1992). Medications and the safety of the older driver: Is there a basis for concern? *Human Factors* 34: 33-47.

Redelmeier, D.A., Yarnell, C.J., Thiruchelvam, D., and Tibshirani, R.J. (2012). Physicians' Warning for Unfit Drivers and the Risk of Trauma from Road Crashes. *The New England Journal of Medicine* 367(13): 1228-36.

Reger, M.A., Welsh, R.K., Watson, G.S., Cholerton, B., Baker, L.D., and Craft, S. (2004). The relationship between neuropsychological functioning and driving ability in dementia: A meta-analysis. *Neuropsychology* 18(1): 85-93.

Robertson, R. and Vanlaar, W. (2008). Elderly drivers: Future Challenges? *Accident Analysis and Prevention* 40: 1982-1986.

Romoser, M., and Fisher, D.L. (2009). Effects of Cognitive and Physical Decline on Older Drivers' Side-to-Side Scanning for Hazards while Executing Turns. In: Driving Assessment 2009: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Big Sky, Montana, June 22-25, 2009. Paper 09. Iowa City, IA: University of Iowa Public Policy Center.

Schanke, A., and Sundet, K. (2000). Comprehensive driving assessment: neuropsychological testing and on-road evaluation of brain injured patients. *Scandinavian Journal of Psychology* 41: 113-121.

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

Shapiro, J.P. (1999). Off the Road. US News and World Report. Lexis, October 25th 1999. Market Library.

Shua-Haim, J., Koppuzha, G., Shua-Haim, Vered, and Gross, J. (1996). A simple scoring system for clock-drawing in patients with Alzheimer's disease. *Journal of the American Geriatric Society* 37:725-729.

Shulman, K.I., Gold, D.P., Cohen, C.A., and Zucchero, C.A. (1993). Clock drawing and dementia in the community: a longitudinal study. *International Journal of Geriatric Psychiatry* 8:487-496.

Siren, A., and Meng, A. (2012). Cognitive screening of older drivers does not produce safety benefits. *Accident Analysis and Prevention* 45: 634-638.

Snellgrove, C.A. (2010) Cognitive screening for the safe driving competence of older people with mild cognitive impairment or early dementia Civic Square, Australian Capital Territory: Australian Transport Safety Bureau.

Strauss, E.H., Sherman, M.S.E., and Spreen, O. (2006). A Compendium of Neuropsychological Tests. Oxford University Press: New York.

Stutts, J.C., Stewart, R., and Martell, C.M. (1996). Can Screening for Cognitive Impairment at License Renewal Identify High Risk Older Drivers? 4*0th Annual Proceedings: Association for the Advancement of Automotive Medicine* , pp. 335-350. Des Plaines, IL.: Association for the Advancement of Automotive Medicine.

Stutts, J.C.; Stewart, J.R.; and Martell, C. 1998. Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention* 30(3): 337-346.

Szlyk, J.P., Myers, L., Zhang, Y., Wetzel, L., and Shapiro, R. (2002). Development and assessment of a neuropsychological battery to air in predicting driving performance. *Journal of Rehabilitation Research and Development* 39(4): 483-496.

Transport Canada. (2007). Canadian Motor Vehicle Traffic Statistics: 2006. TP3322. Transport Canada, Ottawa, ON.

Trenerry, M.R., Crosson, B., and DeBoe, J. (2012). Visual Search and Attention Test (VSAT). PAR. Retrieved from: http://www4.parinc.com/Products/Product.aspx?ProductID=VSAT.

UFOV User's Guide. (2002). UFOV User's Guide: Version 6.0.9. Visual Awareness. Birmingham, AL. Retrieved from: crag.uab.edu/VAI/PDF Pubs/UFOV_Manual_V6.0.6.pdf.

Van Elslande, P., and Fleury, D. (2000). Elderly Drivers: What Errors Do they Commit on the Road? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 44: 3-259.

Voelker, R. (1999). Crash Risk Among Older Drivers Studied. *Journal of the American Medical Association* 282: 1610.

Wang, C.C., and Carr, D.B. (2004). Older Driver Safety: A Report from the Older Drivers Project. *Journal of the American Geriatric Society* 52(1): 143-149.

Whelihan, W.M., DiCarlo, M.A., and Paul, R.H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology* 20: 217-228.

Wilson, K.A. (2007). What to do about Elderly Drivers: Part 1 of 2. Factiva: Special Report. *Autoweek* April 2nd, 2007.

Wilson, K.A. (2007). What to do about Elderly Drivers: Part 2 of 2. Factiva: Special Report. *Autoweek* April 2nd, 2007.

Wood, J.M., Horswill, M.S., Lacherez, P.F., and Anstey, K.J. (2013). Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. A*ccident Analysis and Prevention* 50: 1161-1168.

APPENDIX A

Additional references from our independent resources:

Adams, A.J. (2010). Driving and cognitive decline: a need for collaboration. *Canadian Family Physician* 56(11): 1185-6.

Aksan, N., Anderson, S.W., Dawson, J.D., Johson, A.M., and Rizzo, M. (2012). Cognitive functioning predicts driver safety on road tests 1 and 2 years later. *Journal of the American Geriatrics Society* 60(1): 99-105.

Barrash, J., Stillman, A., Anderson, S.W., Dawson, J.D., and Rizzo, M. (2010). Prediction of driving ability with neuropsychological tests: demographic adjustments diminish accuracy. *Journal of the International Neuropsychological Society* 16(4): 679-86.

Canadian Council of Motor Transport Administrators. (2004). CCMTA Medical Standards for Drivers. Ottawa, ON.: Canadian Council of Motor Transport Administrators. July.

Classen, S., Winter, S.M., Velozo, C.A., Bédard, M., Lanford, D.N., Brumback, B., and Lutz, B.J. (2010). Item development and validity testing for a self- and proxy report: The Safe Driving Behavior Measure. *American Journal of Occupational Therapy* 64(2): 296-305.

De Koning, I., van Kooten, F., Koudstaal, P.J., and Dippel, D.W.J. (2005). Diagnostic value of the Rotterdam-CAMCOG in post-stroke dementia. *Journal of neurology, neurosurgery, and psychiatry* 76(2): 263-265.

Gianutsos, R. (1994). Driving advisement with the Elemental Driving Simulator (EDS)- When less suffices. *Behavior Research Methods Instruments and Computers* 26(2): 183-186.

Hunt, L. (2011). Driving rehabilitation: Frequently asked questions and answers. *Occupational Therapy Now* 12(5):19.

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., and Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *American Journal of Occupational Therapy* 60(4): 428-434.

Lococo, K.H., and Staplin, L. (2005). Strategies for Medical Advisory Boards and Licencing Review. DOT HS 809 874. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. July.

Mosimann, U.P., Bachli-Bietry, J., Boll, J., Bopp-Kistler, I., Donati, F., Kressig, R.W., Martensson, B., Monsch, A.U., Muri, R., Nef, T., Rothenberger, A., Seeger, R., von Gunten, A., and Wirz, U. (2012). Consensus recommendations for the assessment of fitness to drive in cognitively impaired patients. *Praxis* 101(7): 451-64.

National Highway Traffic Safety Administration. (2009). Driving Transitions Education: Tools, Scripts, and Practice Exercises. DOT HS 811 152. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. November.

National Transportation Safety Board. (2004). Medical Oversight of Noncommerical Drivers. Highway Special Investigation Report. NTSB/SIR-04/01. Washington, DC.: National Transportation Safety Board. November.

Neitch, S.M., Madero, G., and Maynard, S. (2011). Driving assessment results in patients with a diagnosis of dementia. *West Virginia Medical Journal* 107(3): 54-8.

O'Neill, D., and Dobbs, B.M. (2004). Age-related disease, mobility, and driving. In: Transportation in an Aging Society: A Decade of Experience, November 7-9, 1999, Bethesda, Maryland, pp. 56-66. Washington, DC.: Transportation Research Board.

Patomella, A.H., Tham, K., Johansson, K., and 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.

Patomella, A.H., and Kottorp, A. (2005). An evaluation of driving ability in a simulator: A good predictor of driving ability after stroke? In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 15. Iowa City, IA.: University of Iowa Public Policy Center.

Petrucelli, E., and Malinowski, M. (1992). Status of Medical Review in Driver Licencing: Policies, Programs and Standards. DOT HS 807 892. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. March.

Pierce, S. (2003). The occupational therapist's roadmap to safety for seniors. Gerontology Special Interest Sectional Quarterly 26(3): 1-3. Porter, M.M.; and Tuokko, H.A. (2011). An evaluation of the Roadwise Review: A mixed methods approach. *Traffic Injury Prevention* 12(5): 451-458.

Rizzo, M. (2011). Impaired driving from medical conditions: a 70-year-old man trying to decide if he should continue driving. *Journal of the American Medical Association* 305(10): 1018-26.

Siren, A., and Meng, A. (2012). Cognitive screening of older drivers does not produce safety benefits. *Accident Analysis and Prevention* 45: 634-8.

Staplin, L., and Hunt, L. (2004). Driver programs. In: Transportation in an Aging Society: A Decade of Experience, November 7-9, 1999, Bethesda, Maryland, pp. 69-94. Washington, DC.: Transportation Research Board.

Unsworth, C.A., Pallant, J.F., Russell, K., Odell, M. and Coulson, M. (2011). Interrater reliability of the Road Law and Road Craft Test as part of the OT-DORA Battery for off-road driver assessement. B*ritish Journal of Occupational Therapy* 74: 394-98.

Unsworth, C.A., Pallant, J.F., Russell, K.J., Germano, C., and Odell, M. (2010). Validation of a test of road law and road craft knowledge with older of functionally impaired drivers. *American Journal of Occupational Therapy* 64(2): 306-315.

Vann, A.S. (2012). Please tell people with Alzheimer's disease not to drive. *Journal of Geriatrics Society* 60(3): 597-8.

Watanabe, T., Konagaya, Y., Yanagi, T. Miyao, M., and Shibayama, H. (2012). Study of daily driving characteristics of individuals with dementia using video-recording driving recorders. *Journal of the American Geriatrics Society* 60(7): 1381-3.

APPENDIX B

Complete reference list of all initially identified sources:

AAA Foundation for Traffic Safety. (1994). Drivers 55 Plus: Check Your Own Performance: A Self-Rated Form of Questions, Facts and Suggestions for Safe Driving. Washington, DC.: AAA Foundation for Traffic Safety.

Adams, A.J. (2010). Driving and cognitive decline: a need for collaboration. *Canadian Family Physician* 56(11): 1185-6.

Adler, G., Rottunda, S., and Dysken, M. (2005). The older driver with dementia: An updated literature review. *Journal of Safety Research* 36(4): 399-407.

Adler G., and Silverstein, N.M. (2008). At-risk drivers with Alzheimer's disease: Recognition, response, and referral. *Traffic Injury Prevention* 9(4): 299-303.

Adrian, J., Postal, V., Moessinger, M., Rascle, N., and Charles, A. (2011). Personality traits and executive functions related to on-road driving performance among older drivers. *Accident Analysis and Prevention* 43(5): 1652-1659.

Akinwuntan, E.A., De Weerdt, W., Feys, H., Verheyden, G., De Vooght, F., Devos, H., Baten, G., and Kiekens, C. (2007). Training of driving-related attentional performance after stroke using a driving simulator. In: Driving Assessment 2007: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Stevenson, Washington, July 9-12, 2007. Paper 19. Iowa City, IA.: University of Iowa Public Policy Center.

Akinwuntan, E.A., Devos, H., Feys, H., Verheyden, G., De Weerdt, W., Baten, G., and Kiekens, C.(2007). Short and predictive assessment battery of fitness-to-drive after stroke. In: Driving Assessment 2007: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Stevenson, Washington, July 9-12, 2007. Paper 45. Iowa City, IA.: University of Iowa Public Policy Center.

Akinwuntan, E.A., De Weerdt, W., Feys, H., Pauwels, J., Baten, G., Arno, P., and Kiekens, C. (2005). Effect of simulator training on driving after stroke: A randomized controlled trial. *Neurology* 65: 843-850.

Akinwuntan, E.A., Feys, H., De Weerdt, W., Baten, G., Arno, P., and Kiekens, C. (2006). Prediction of Driving after Stroke: A Prospective Study. *The American Society of Neurorehabilitation* 20(3): 417-423.

Aksan, N., Anderson, S.W., Dawson, J.D., Johson, A.M., and Rizzo, M. (2012). Cognitive functioning predicts driver safety on road tests 1 and 2 years later. *Journal of the American Geriatrics Society* 60(1): 99-105.

Alvarez, F.J., del Rio, M.C., Fierro, I., Vicondoa, A., and Ozcoidi, M. (2004). Medical Condition and Fitness to Drive: Prospective Analysis of the Medical-Psychological Assessment of Fitness to Drive and Accident Risk. Valladolid, Spain: Universidad de Valladolid. June.

American Association of Motor Vehicle Administrators. Jurisdictional Procedures, Practices & Programs. http://www.aamva.org/knoweldgecenter/Driver/At-Risk/JurisProceduresPractAndProgs.htm. Accessed November 30, 2007.

American Association of Retired Persons. (1992). Older Driver Skill Assessment and Resource Guide: Creating Mobility Choices. Washington, DC.: American Association of Retired Persons.

Anderson, S.W., Rizzo, M.,Shi, Q., Uc, E.Y., and Dawson, J.D. (2005). Cognitive abilities related to driving performance in a simulator and crashing on the road. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 41. Iowa City, IA.: University of Iowa Public Policy Center.

Anstey, K.J., Windsor, T.D., Luszcz, M.A., and Andrews, G.R. (2006). Predicting driving cessation over 5 years in older adults: Psychological well-being and cognitive competence are stronger predictors than physical health. *Journal of the American Geriatrics Society* 54(1): 121-126.

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

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

Antin, J.F., Lockhart, T.E., Stanley, L.M., and Guo, F. (2012). Comparing the impairment profiles of older drivers and non-drivers: Toward the development of a fitness-to-drive model. *Safety Science* 50(2): 333-341.

Artero, S., and Ritchie, K. (2003). The detection of mild cognitive impairment in the general practice setting. *Aging and Mental Health* 7(4): 251-258.

Arthur, W. Jr., Strong, M.H., and Williamson, J. (1994). Validation of a visual attention test as a predictor of driving accident involvement. *Journal of Occupational and Organizational Psychology* 67(Part 2): 173-182.

Ashman, R.D., Bishu, R.R., Foster, B.G., and McCoy, P.T. (1994). Countermeasures to improve the driving performance of older drivers. *Educational Gerontology* 20(6): 567-577.

Austroads. (2012). Assessing Fitness to Drive for Commercial and Private Vehicle Drivers. 4th Edition. AP-G56/12. Sydney, New South Wales. Austroads.

Avolio, B.J., Kroeck, K.G., and Panek, P.E. (1985). Individual differences in information-processing ability as a predictor of motor vehicle accidents. *Human Factors* 27(5): 577-587.

Bailey, T.J. (1994). A Comparative Study of Guidelines for Determining Medical Fitness to Drive. 2/97. Walkerville, South Australia: South Australian Department of Transport, Office of Road Safety. November.

Baldwin, C.L. (2002). Designing in-vehicle technologies for older drivers: Application of sensory-cognitive interaction theory. *Theoretical Issues in Ergonomics Science* 3(4): 307-329.

Ball, K.K., Berch, D.B., Helmers, K.F., Jobe, J.B., Leveck, M.D., Marsiske, M., Morris, J.N., Rebok, G.W., Smith, D.M., Tennstedt, S.L., Unverzagt, F.W., and Willis, S.L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association* 288(18): 2271-2281.

Ball, K.K., Clay, O.J., Wadley, V.G., Roth, D.L., Edwards, J.D., and Roenker, D.L. (2005). Predicting driving performance in older adults with the useful field of view test: A meta-analysis. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 08. Iowa City, IA.: University of Iowa Public Policy Center.

Ball,K.K., Edwards, J.D., and Ross, L.A. (2007). The impact of speed of processing training on cognitive and everyday functions. Journals of Gerontology: Series B 62B (Special Issue I): 19-31.

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

Ball, K.K., Roenker, D.L., McGwin, G. Jr., Wadley, V.G., Edwards, J.D., and Raleigh, R. (2003). Can high risk older drivers be identified in a DMV setting with a brief battery of functional tests? In: Driving Assessment 2003: International Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, eds. D.V. McGehee, J.D. Lee and M. Rizzo, p. 39. Iowa City, IA.: Public Policy Center, University of Iowa.

Ball, K.K., Roenker, D.L., Wadley, V.G., Cissell, G., Matthews, M., Ball, D., Vance, D., Frankel, M., and McConnell, K. (2005). Longitudinal assessment of older drivers in a DMV setting. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 57. Iowa City, IA.: University of Iowa Public Policy Center.

Ball, K.K., Roenker, D.L., Wadley, V.G., Edwards, J.D., Roth, D.L., McGwin, G. Jr., Raleigh, R., Joyce, J.J., Cissell, G.M., and Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? *Journal of the American Geriatrics Society* 54(1): 77-84.

Ball, K.K., Wadley, V.G., Edwards, J.D., Ball, D., and Roenker, D.L. (2001). Meta-analysis of crash risk factors among older drivers: Application to a model program of driver screening. In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 63. Iowa City, Iowa: University of Iowa Public Policy Center.

Bao, S., and Boyle, L.N. (2009). Driver safety programs: The influence on the road performance of older drivers. In: Traffic Safety: Roundabouts, Trucks, Older Drivers, and Traffic Law Enforcement 2009. *Transportation Research Record* 2096: 76-80.

Barrash, J., Stillman, A., Anderson, S.W., Dawson, J.D., and Rizzo, M. (2010). Prediction of driving ability with neuropsychological tests: demographic adjustments diminish accuracy. *Journal of the International Neuropsychological Society* 16(4): 679-86.

Bédard, M., Marshall, S., Man-Son-Hing, M., and Weaver, B. (2012). Application of the "Screen for the Identification of cognitively impaired Medically At-Risk Drivers": Preliminary results from the Candrive driver cohort. In: *Proceedings of the Canadian Multidisciplinary Road Safety Conference XIX. Paper Session 2A: Older Drivers* I. Banff, AB.: Canadian Association of Road Safety Professionals.

Bédard, M., Molloy, D.W., and Lever, J.A. (1998). Factors associated with motor vehicle crashes in cognitively impaired older adults. *Alzheimer Disease and Associated Disorders* 12(3): 135-139.

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

Bédard, M., Porter, M.M., Marshall, S., Isherwood, I., Riendeau, J., Weaver, B., Tuokko, H., Molnar, F., and Miller-Polgar, J. (2008). The combination of two training approaches to improve older adults' driving safety. *Traffic Injury Prevention* 9(1): 70-76. [*]

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

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

Berndt, A., May, E., and Clark, M. (2007). Drivers with dementia: Environment, errors and performance outcomes. In: Driving Assessment 2007: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Stevenson, Washington, July 9-12, 2007. Paper 67. Iowa City, IA.: University of Iowa Public Policy Center.

Bherer, L., Kramer, A.F., Peterson, M.S., Colcombe, S., Erickson, K., and Becic, E. (2005). Training effects of dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging* 20(4): 695-709.

Blanchard, R.A., Myers, A.M., and Porter, M.M. (2010). Correspondence between self-reported and objective measures of driving exposure and patterns in older drivers. *Accident Analysis and Prevention* 42(2): 523-529.

Bliokas, W., Taylor, J.E., Leung, J., and Deane, F.P. (2011). Neuropsychological assessment of fitness to drive following acquired cognitive impairment. *Brain Injury* 25(5): 471-487.

Borowsky, A., Shinar, D., and Oron-Gilad, T. (2007). Age, skill and hazard perception in driving. In: Driving Assessment 2007: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Stevenson, Washington, July 9-12, 2007. Paper 89. Iowa City, IA.: University of Iowa Public Policy Center.

Borson, S., Scanlan, J.M., Watanabe, J., Tu, S.P., and Lessig, M. (2005). Simplifying detection of cognitive impairment: Comparison of the Mini-Cog and Mini-Mental State Examination in a multiethnic sample. *Journal of the American Geriatrics Society* 53(5): 871-874.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist* 5(2): 125-142.

British Columbia Medical Association. (1997). Guide for Physicians in Determining Fitness to Drive a Motor Vehicle. British Columbia: British Columbia Medical Association.

British Columbia Ministry of Transportation and Highways. (2000). Bibliography of Research of Medical, Physical and Cognitive Conditions Affecting Driver Fitness. Victoria, BC.: B.C. Ministry of Transportation and Highways, Office of the Superintendent of Motor Vehicles. September.

British Psychological Society. (2001). Fitness to Drive and Cognition: A Document of the Multi-Disciplinary Working Party on Acquired Neuropsychological Deficits and Fitness to Drive 1999. Leicester, UK.: British Psychological Society. January.

Brodaty, H., Pond, D., Kemp, N.M., Luscombe, G., Harding, L., Berman, K., and Huppert, F.A. (2002). The GPCOG: A new screening test for dementia designed for general practice. *Journal of the American Geriatrics Society* 50(3): 530-534.

Bromberg, S., Oron-Gilad, T., Ronen, A., Borowsky, A., and Parmet, Y. (2012). The perception of pedestrians from the perspective of elderly experienced and experienced drivers. *Accident Analysis and Prevention* 44(1): 48-55.

Brooke, P., and Bullock, R. (1999). Validation of a 6 item cognitive impairment test with a view to primary care usage. *International Journal of Geriatric Psychiatry* 14(11): 936-940.

Brouwer, W. H. (2008). Aging related diseases and fitness to drive - assessment and rehabilitation. In: Wolf-Rüdiger Nickel, Martin Koran: Fit to Drive, 3rd International Traffic Expert Congress, Prague 2008, Proceedings, 73-77.

Brouwer, W. H., and Dotzauer, M. (2011). Compensating age-related physical and mental, impairments of older drivers by technological measures. In: Divera Twisk and Wolf-Rüdiger Nickel: Fit to Drive, 5th International Traffic Expert Congress, The Hague 2011, Proceedings, 89-92.

Buchstaller, R. (2007). Mental performance capacity of older drivers - Aspects of test fairness and age fairness. In:Ralf Risser & Wolf-Rüdiger Nickel: Fit to Drive, 2nd International Traffic Expert Congress, Vienna 2007, Proceedings, 84.

Bunce, D., Young, M. S., Blane, A., and Khugputh, P. (2012). Age and inconsistency in driver performance. *Accident Analysis and Prevention* 49(2): 293-299.

Buske, L. (2000). Driving: How old is too old? Canadian Medical Association 163(10): 1334.

Caird, J.K., and Dewar, R. (1999). A review of older drivers' perception of and performance with advanced transportation telematics (ATT). In: Keeping the Elderly Mobile. Outdoor Mobility of the Elderly: Problems and Solutions. Papers and Discussions of the Euroconference in Rolduc, June 1998, eds. M. Tacken, F. Marcellini, H. Mollenkopf, and I. Ruoppila, pp. 263-273. Delft, Netherlands: Delft University of Technology, Urban and Regional Planning Group.

California Department of Motor Vehicles. (1990). Conference on Driver Competency Assessment: Summary of Concurrent Workshop Recommendations; Sessions A, B, and C. San Diego, California, October 24-26, 1990.

California Department of Motor Vehicles. (1991). Summary of Proceedings of the Conference on Driver Competency Assessment. San Diego, California, October 24-26, 1990. Sacramento, CA.: California Department of Motor Vehicles.

California Department of Motor Vehicles. (2011). Senior Guide for Safe Driving. Sacramento, CA.: California Department of Motor Vehicles.

Callahan, C.M., Unverzagt, F.W., Hui, S.L., Perkins, A.J., and Hendrie, H.C. (2002). Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Medical Care* 40(9): 771-781.

Canadian Cardiovascular Society. (2003). Canadian Cardiovascular Society Consensus Conference 2003: Assessment of the Cardiac Patient for Fitness to Drive and Fly. Ottawa, ON.: Canadian Cardiovascular Society.

Canadian Council of Motor Transport Administrators. (2004). CCMTA Medical Standards for Drivers. Ottawa, ON.: Canadian Council of Motor Transport Administrators. July.

Canadian Medical Association. (1981). To Drive? Or Not to Drive? The Medical Answer. Guide to Physicians in Determining Fitness to Drive a Motor Vehicle. Ottawa, ON.: Canadian Medical Association.

Canadian Medical Association. (2000). Determining Medical Fitness to Drive. A Guide for Physicians. 6th Edition. Ottawa, ON.: Canadian Medical Association.

Cantin, V., Lavallière, M., Simoneau, M., and Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis and Prevention* 41(4): 763-771.

Carpenter, C.R., Bassett, E.R., Fischer, G.M., Shirshekan, J., Galvin, J.E., and Morris, J.C. (2011). Four sensitive screening tools to detect cognitive dysfunction in geriatric emergency department patients: Brief Alzheimer's Screen, Short Blessed Test, Ottawa 3DY, and Caregiver-completed AD8. *Academic Emergency Medicine* 18(4): 374-384.

Carr, D.B., Barco, P.P., Wallendorf, M.J., Snellgrove, C.A., and Ott, B.R. (2010). R1: Predicting road test performance in drivers with dementia. Paper presented at the International Conference for Alzheimer's Disease, July 2010, Honolulu, Hawaii. St. Louis, MO.: Washington University, School of Medicine.

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

Carr, D.B., Duchek, J., Meuser, T.M., and Morris, J.C. (2006). Older adult drivers with cognitive impairment. *American Family Physician* 73(6): 1029-1034.

Carr, D.B., Duchek, J., and Morris, J.C. (2000). Characteristics of motor vehicle crashes of drivers with dementia of the Alzheimer type. *Journal of the American Geriatric Society* 48(1): 18-22.

Carr D.B, and Ott, B.R. (2010). The older adult driver with cognitive impairment: "It's a very frustrating life." *Journal of the American Medical Association* 303(16): 1632-1641.

Carr, D., Schmader, K., Bergman, C., Simon, T.C., Jackson, T.W., Haviland, S., and O'Brien, J. (1991). A multidisciplinary approach in the evaluation of demented drivers referred to geriatric assessment centers. *Journal of the American Geriatrics Society* 39(11): 1132-1136.

Carr, D.B., Schwartzberg, J.G., Manning, L., and Sempek, J. (2010). Physician's Guide to Assessing and Counseling Older Drivers. 2nd Edition. DOT HS 811 298. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. June.

Carter, T. (2002). Medical aspects of fitness to driver. In: Behavioural Research in Road Safety: Twelfth Seminar, pp. 127-132. London, England: Department for Transport. {IN 17.12}

Cassavaugh, N.D., Domeyer, J.E., and Backs, R.W. (2009). The effect of age on decision making during unprotected turns across oncoming traffic. In: Driving Assessment 2009: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Big Sky, Montana, June 22-25, 2009. Paper 14. Iowa City, IA: University of Iowa Public Policy Center.

Ceci, R., Högman, L., and Patten, C. (2001). Measures of driver behavior and cognitive workload in a driving simulator and in real traffic environment: Experiences from two experimental studies in Sweden. In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 25. Iowa City, Iowa: University of Iowa Public Policy Center.

Cercy, S.P., Simakhodskaya, Z., and Elliott, A. (2010). Diagnostic accuracy of a new instrument for detecting cognitive dysfunction in an emergent psychiatric population: The Brief Cognitive Screen. *Academic Emergency Medicine* 17(3): 307-315.

Charlton, J.L., Koppel, S., Langford, J., Myers, A., Tuokko, H., Marshall, S., and Man-Son-Hing, M. (2012). Findings from the Candrive/Ozcandrive Study: Older drivers, cognitive function and self-regulatory driving practices. In: P*roceedings of the Canadian Multidisciplinary Road Safety Conference XIX. Paper Session 5C: Older Drivers III.* Banff, AB.: Canadian Association of Road Safety Professionals.

Chinn, A. (2002). Confidentiality, privacy and the reporting by health professionals of a patient's unfitness to drive: Legal dilemmas. In: 2002 Road Safety Research, Policing and Education Conference, November 4-5, 2002, Adelaide, Australia, pp. 315-320. Adelaide, South Australia: Transport South Australia.

Classen, S., Horgas, A., Awadzi, K., Messinger-Rapport, B., Shechtman, O., and Joo, Y. (2008). Clinical predictors of older driver performance on a standardized road test. *Traffic Injury Prevention* 9(5): 456-462.

Classen, S., McCarthy, D.P., Shechtman, O., Awadzi, K.D., Lanford, D.N., Okun, M.S., Rodriguez, R.L., Romrell, J., Bridges, S., Kluger, B., and Fernandez, H.H. (2009). Useful field of view as a reliable screening measure of driving performance in people with Parkinson's disease: Results of a pilot study. *Traffic Injury Prevention* 10(6): 593-598.

Classen, S., Velozo, C.A., Winter, S.M., Wang, Y., and Lanford, D.N. (2012). Validity and Usability of a Safe Driving Behaviors Measure for Older Adults. Gainesville, FL.: University of Florida, Department of Occupational Therapy. October.

Classen, S., Wen, P.S., Velozo, C.A., Bédard, M., Winter, S.M., Brumback, B., and 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.

Classen, S., Wen, P.S., Velozo, C.A., Bédard, M., Winter, S.M., Brumback, B., and Lanford, D.N. (2012). Rater reliability and rater effects of the Safe Driving Behavior Measure. *American Journal of Occupational Therapy* 66(1): 69-77.

Classen, S., Winter, S.M., Velozo, C.A., Bédard, M., Lanford, D.N., Brumback, B., and Lutz, B.J. (2010). Item development and validity testing for a self- and proxy report: The Safe Driving Behavior Measure. American *Journal of Occupational Therapy* 64(2): 296-305.

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

Cobb, R.W., and Coughlin, J.F. (1998). Are elderly drivers a road hazard?: Problem definition and political impact. *Journal of Aging Studies* 12(4): 411-427.

Cook, S.E., Marsiske, M., and McCoy, K.J.M. (2009). The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnestic mild cognitive impairment. *Journal of Geriatric Psychiatry and Neurology* 22(2): 103-109.

Cooper, P.J., Tallman, K., Tuokko, H., and Beattie, B.L. (1993). Vehicle crash involvement and cognitive deficit in older drivers. *Journal of Safety Research* 24(1): 9-17.

Cotter, V.T., Clark, C.M., and Karlawish, J.H. (2003). Cognitive function assessment in individuals at risk for Alzheimer's disease. *Journal of the American Academy of Nurse Practitioners* 15(2): 79-86.

Crizzle, A. M., Classen, S.B.M., Lanford, D., and Winter, S. (2012). MMSE as a predictor of on-road driving performance in community dwelling older drivers. *Accident Analysis and Prevention* 49(2): 287-292.

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

Dawson, J.D., Anderson, S.W., Uc, E.Y., Dastrup, E., and Rizzo, M. (2009). Predictors of driving safety in early Alzheimer disease. *Neurology* 72(6): 521-527.

de Jager, C.A., Budge, M.M., and Clarke, R. (2003). Utility of TICS-M for the assessment of cognitive function in older adults. *International Journal of Geriatric Psychiatry* 18(4): 318-324.

de Koning, I., Dippel, D.W.J., van Kooten, F., and Koudstaal, P.J. (2000). A Short Screening Instrument for Poststroke Dementia : The R-CAMCOG. *Stroke* 31: 1502-1508.

de Koning, I., van Kooten, F., Koudstaal, P.J., and Dippel, D.W.J. (2005). Diagnostic value of the Rotterdam-CAMCOG in post-stroke dementia. *Journal of neurology, neurosurgery, and psychiatry* 76(2): 263-265.

De Raedt, R., and Ponjaert-Kristofferson, I. (2000). The relationship between cognitive/ neuropsychological factors and car driving performance in older adults. *Journal of the American Geriatrics Society* 48(12): 1664-1668.

De Raedt, R., and Ponjaert-Kristoffersen, I. (2001). Predicting at-fault car accidents of older drivers. *Accident Analysis and Prevention* 33(6): 809-819.

De Raedt, R., and Ponjaert-Kristofferen, I. (2001). Short Cognitive/Neuropsychological Test Battery for First-Tier Fitness-To-Drive Assessment of Older Adults. *The Clinical Neuropsychologist* 15(3): 329-336.

Devlin, A., McGillivray, J., Charlton, J., Lowndes, G., and Etienne, V. (2012). Investigating driving behaviour of older drivers with mild cognitive impairment using a portable driving simulator. *Accident Analysis and Prevention* 49(6): 300-307.

Devos, H., Vandenberghe, W., Nieuboer, A., Tant, M., Baten, G., and De Weerdt, W. (2007). Predictors of fitness to drive in people with Parkinson disease. *Neurology* 69(14): 1434-1441.

DiStefano, M., and 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.

Dobbs, A.R. (1997). Evaluating the driving competence of dementia patients. Alzheimer's Disease and Associated Disorders, Suppl. 1: 8-12.

Dobbs, A.R. (2001). Lessons learned through the development and implementation of an evaluation for medically at-risk drivers. In: Road Safety Research, Policing and Education Conference Proceedings, November 18-20, 2001. Melbourne: Monash University.

Dobbs, A.R., Heller, R.B., and Schopflocher, D. (1998). A comparative approach to identify unsafe older drivers. *Accident Analysis and Prevention* 30(3): 363-370.

Dobbs, B.M. (2005). Medical Conditions and Driving: A Review of the Scientific Literature (1960-2000). DOT HS 809 690. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. September.

Dobbs, B.M., and Dobbs, A. (2001). Improving the safety and the mobility of older drivers: A conceptual framework. In: Road Safety Research, Policing and Education Conference Proceedings, November 18-20, 2001. Melbourne: Monash University.

Dobbs, B.M., and Schopflocher, D. (2010). The introduction of a new screening tool for the identification of cognitive impaired medically at-risk drivers: The SIMARD A Modification of the DemTect. *Journal of Primary Care and Community Health* 1(2): 119-127.

Donnelly, R.E., Karlinsky, H.J., Young, M.L., Ridgely, J.N., and Lamble, R.W. (1992). Fitness to Drive in Elderly Individuals with Progressive Cognitive Impairment. Toronto, Ont.: Ministry of Transportation. September.

Drachman, D.A., Swearer, J.M., Kane, K., Osgood, D., O'Toole, C., and Moonis, M. (1995). The cognitive assessment screening test (CAST) for dementia. *Journal of Geriatric Psychiatry and Neurology* (4): 200-208.

Eby, D.W., and Molnar, L.J. (1998). Constructing Guidelines for Appropriate Traffic Safety Messages for Youth: A Literature Review of Cognitive Development. Draft. Ann Arbor, MI.: University of Michigan Transportation Research Institute. May.

Eby, D.W., and Molnar, L.J. (1999). Matching Traffic Safety Strategies to Youth Characteristics: A Literature Review of Cognitive Development. DOT HS 808 927. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. June.

Eby, D.W., and Molnar, L.J. (2001). Older drivers: A comparison of self-reported driving-related decisions with observed driving problems. In: *Proceedings of the 45th Annual Conference of the Association for the Advancement of Automotive Medicine*, pp. 413-415. Barrington, IL.: Association for the Advancement of Automotive Medicine.

Eby, D.W., Molnar, L.J., Nation, A.D., Shope, J.T., and Kostyniuk, L.P. (2006). Development and Testing of an Assessment Battery for Older Drivers. UMTRI-2006-04. Ann Arbor, MI.: University of Michigan Transportation Research Institute. May.

Eby, D.W., Molnar, L.J., Shope, J.T., and Dellinger, A.M. (2007). Development and pilot testing of an assessment battery for older drivers. *Journal of Safety Research* 38(5): 535-543.

Eby, D.W., Molnar, L.J., Shope, J.T., Vivoda, J.M., and Fordyce, T.A. (2003). Improving older driver knowledge and self-awareness through self-assessment: The Driving Decisions Workbook. *Journal of Safety Research* 34(4): 371-381.

Eby, D.W., Shope, J.T., Molnar, L.J., Vivoda, J.M., and Fordyce, T.A. (2000). Improvement of Older Driver Safety Through Self Evaluation: The Development of a Self-Evaluation Instrument. Report no. UMTRI-2000-04. Ann Arbor, MI.: University of Michigan Transportation Research Institute. March.

Eby, D.W., Trombley, D.A., Molnar, L.J., and Shope, J.T. (1998). The Assessment of Older Drivers' Capabilities: A Review of the Literature. Report no. UMTRI-98-24. Ann Arbor, MI.: University of Michigan Transportation Research Institute. August.

Eck, R.W., and Winn, G. (2002). Older drivers' perceptions of problems at unsignalized intersections on divided highways. In: Transportation Research Board 81st Annual Meeting CD-ROM. TRB 02-2365. Washington, DC.: Transportation Research Board.

Edwards, J.D., Bart, E., O'Connor, M.L., and Cissell, G. (2010). Ten years down the road: Predictors of driving cessation. *Gerontologist* 50(3): 393-399.

Edwards, J.D., Delahunt, P.B., and Mahncke, H.W. (2009). Cognitive speed of processing training delays driving cessation. *Journal of Gerontology: Medical Sciences* 64(12): 1262-1267.

Edwards, J.D., Leonard, K.M., Lunsman, M., Dodson, J., Bradley, S., Myers, C.A., and Hubble, B. (2008). Acceptability and validity of older driver screening with the DrivingHealth® Inventory. *Accident Analysis and Prevention* 40(3): 1157-1163.

Edwards, J.D., Myers, C., Ross, L.A., Roenker, D.L., Cissell, G.M., McLaughlin, A.M., and Ball, K.K. (2009). The longitudinal impact of cognitive speed of processing training on driving mobility. *The Gerontologist* 49(4): 485-494.

Edwards, J.D., Perkins, M., Ross, L.A., and Reynolds, S.L. (2009). Driving status and three-year mortality among community-dwelling older adults. *Journal of Gerontology A: Biological Science and Medical Science* 64(2): 300-305.

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

Elkin-Frankston, S., Lebowitz, B.K., Kapust, L.R., Hollis, A.M., and O'Connor, M.G. (2007). The use of the Colour Trails Test in the assessment of driver competence: Preliminary report of a culture-fair instrument. *Archives of Clinical Neuropsychology* 22(5): 631-635.

Ellis, R.J., Kaining J., Kawas, C., Koller, W.C., Lyons, K.E., Jeste, D.V., Hansen, L.A., and Thal, L.J. (1998). Diagnostic validity of the Dementia Questionnaire for Alzheimer Disease. *Archives of Neurology* 55(3): 360-365.

Etienne, V., Marin-Lamellet, C., and Paire-Ficout, L. (2010). Simulator-based study to investigate executive functioning of older drivers at left-turn intersections. Paper 10-1238. In: Transportation Research Board 89th Annual Meeting Compendium of Papers DVD. Washington, DC.: Transportation Research Board.

Ferreira, I.S., Marmeleira, J.F., Godinho, M.B., and Simões, M.R. (2007). Cognitive factors and gender related to self-reported difficulties in older drivers. In: Driving Assessment 2007: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Stevenson, Washington, July 9-12, 2007. Paper 83. Iowa City, IA.: University of Iowa Public Policy Center.

Ferreira, I.S., Simoes, M.R., and Maroco, J. (2012). The Addenbrooke's Cognitive Examination Revised as a potential screening test for elderly drivers. *Accident Analysis and Prevention* 49:278-286.

Fildes, B., Pronk, N., Langford, J., Hull, M., Anderson, R., and Frith, B. (1999). A national licence assessment program for older drivers in Australasia. In: Speaker's Papers: 1999 Insurance Commission of Western Australia Conference on Road Safety "Green Light for the Future", pp. 118-138. Perth, Western Australia: Insurance Commission of Western Australia.

Fildes, B.N., Charlton, J., Newstead, S., Oxley, P., Pronk, N., Langford, J., and Frith, B. (2002). Evaluation of screening tests for older drivers. In: 2002 Road Safety Research, Policing and Education Conference, November 4-5, 2002, Adelaide, Australia, p. 233. Adelaide, South Australia: Transport South Australia.

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

Fildes, B.N., Langford, J., Deery, H., Hull, M., Anderson, R., and Frith, B. (1998). Licencing assessment for older drivers. In: Proceedings of Road Safety Research, Policing and Education Conference, Wellington, New Zealand, November 16-17, 1998, pp. 379-389. Vol. 2. Wellington, New Zealand: Land Transport Safety Authority.

Fildes, B.N., Pronk, N., Charlton, J.L., Langford, J., and Frith, B. (2001). Development of a national licence assessment program for older drivers in Australasia. In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 6. Iowa City, Iowa: University of Iowa Public Policy Center.

Fildes, B.N., Pronk, N., Charlton, J.L., Langford, J., and Frith, B. (2001). Development of a national licence assessment program for older drivers in Australasia. In: Proceedings of the Conference Road Safety on Three Continents, International Conference in Moscow, Russia, VTI Konferens 18B, pp. 151-161. Linköping, Sweden: Swedish Road and Transport Research Institute. November.

Fildes, B.N., Pronk, N., Langford, J., Hull, M., Frith, B., and Anderson, R. (2000). Model license re-assessment procedure for older and disabled drivers. AP-R176. Sydney, New South Wales: Austroads.

Finn, J. (2004). Driving Evaluation and Retraining Programs: A Report of Good Practices, 2004. Bethesda, MD.: American Occupational Therapy Association; and National Highway Traffic Safety Administration.

Finn, P., and Bragg, B.W.E. (1986). Perception of the risk of an accident by young and older drivers. *Accident Analysis and Prevention* 18(4): 289-298.

Fitten, J.L. (2003). Driver Screening for Older Adults. *Archives of internal medicine* 163(18): 2129-2131.

Flicker,L., Logiudice, D., Carlin, J.B., and Ames, D. (1997). The predictive value of dementia screening instruments in clinical populations. *International Journal of Geriatric Psychiatry* 12(2): 203-209.

Fonda, S.J., Wallace, R.B., and Herzog, A.R. (2001). Changes in driving patterns and worsening depressive symptoms among older adults. *Journal of Gerontology B: Social Sciences* 56(6): S343-351.

Fortin-McCue, L.A., Saleheen, H., McQuay, J., Jacobs, L., and Lapidus, G. (2013). Implementation of a community-based mature driver screening and referral program: A feasibility study. *Accident Analysis and Prevention* 50: 751-757.

Fox, G.K., Bowden, S.C., Bashford, G.M., and Smith, D.S. (1997). Alzheimer's disease and driving: Prediction and assessment of driving performance. *Journal of the American Geriatrics Society* 45: 949-953.

Freund, B., Colgrove, L.A., Burke, B.L., and McLeod, R. (2005). Self-rated driving performance among elderly drivers referred for driving evaluation. *Accident Analysis and Prevention* 37(4): 613-618.

Freund, B., Gravenstein, S., Ferris, R., Burke, B.L., and Shaheen, E. (2005). Drawing clocks and driving cars: Use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine* 20(3): 240-244.

Freund, B., Gravenstein, S., Ferris, R., and Shaheen, E. (2002). Evaluating driving performance of cognitively impaired and healthy older adults: A pilot study comparing on-road testing and driving simulation. *Journal of the American Geriatric Society* 50(7): 1309-1310.

George, S., and Gélinas, I. (2012). Screening drivers with cognitive impairments using DriveSafe and Drive Aware: Promising tests to predict safe and unsafe drivers. *Australian Occupational Therapy Journal* 59(1): 105-106.

Gianutsos, R. (1994). Driving advisement with the Elemental Driving Simulator (EDS)- When less suffices. Behavior Research Methods Instruments and Computers 26(2): 183-186.

Ginzburg, T., and Minichiello, R. (2011). License removal, driving assessments and the process of re-acquisition. Toronto, ON.: DriveAgain.

Goode, K.T., Ball, K.K., Sloane, M., Roenker, D.L., Roth, D.L., Myers, R.S., and Owsley, C. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings* 5(4): 425-440.

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

Hagge, R., and Stylos, L. (1989). Development and Evaluation of a Risk Assessment Strategy for Medically Impaired Drivers - A Detailed Analysis.

Hebert, D. (2007). Assessing fitness to drive in older adults with stroke. Paper presented at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Helander, C.J. (1986). Development and Evaluation of a Risk Assessment Strategy for Medically Impaired Drivers. Volume 8 of an Evaluation of the California Drunk Driving Countermeasure System. Sacramento, CA.: California Office of Traffic Safety.

Henderson, S., Gagnon, S., Bélanger, A., Tabone, R., and Collin, C. (2010). Near peripheral motion detection threshold correlates with self-reported failures of attention in younger and older drivers. *Accident Analysis and Prevention* 42(4): 1189-1194.

Hennessy, D.F., and Janke, M.K. (2005). Clearing a Road to Driving Fitness by Better Assessing Driving Wellness: Development of California's Prospective Three-Tier Driving-Centered Assessment System. Summary Report. CAL-DMV-RSS-05-215. Sacramento, CA: Department of Motor Vehicles, Research and Development Branch. September.

Hill, L., Rybar, J., Baird, S., Concha-Garcia, S., Coimbra, R., and Patrick, K. (2011). Road safe seniors: Screening for age-related driving disorders in inpatient and outpatient settings. *Journal of Safety Research* 42(3): 165-169.

Hobson, P., and Meara, J. (1999). The detection of dementia and cognitive impairment in a community population of elderly people with Parkinson's disease by use of the CAMCOG neuropsychological test. *Age and Ageing* 28(1): 39-43.

Hodkinson, H.M. (1972). Evaluation of a mental test score for assessment of mental impairment in the elderly. *Age and Ageing* 1(4): 233-238.

Hoff, C.M. (2007). Review of Selected 2007 California Legislation: Vehicle: Putting Seniors in the Drivers' Seats: Improving California's Mature Drivers Program with Chapter 129. McGeorge Law Review, 38: 321-1113.

Hoffmann-Born, H., and Brenner-Hartmann, J. (2006). Decision-making in driving fitness diagnostics-medical and psychological criteria. In: Fit to Drive 1st International Traffic Expert Congress, eds. W.R. Nickel and P. Sardi, pp. 65-67. Bonn, Germany: Kirschbaum Verlag.

Hogan, D.B. (2005). Which older patients are competent to drive? Approaches to office-based assessment. *Canadian Family Physician* 51(3): 362-368.

Hogan, D.B., and Bédard, M. (2011). Papers that might change practice - Review of "The Introduction of a new screening tool for the identification of cognitively impaired medically at-risk drivers". *Canadian Geriatrics Journal* 14(2): 51-54.

Hoggarth, P.A., Innes, C.R.H., Dalrymple-Alford, J.C., Severinsen, J.E., and Jones, R.D. (2010). Comparison of a linear and a non-linear model for using sensory-motor, cognitive, personality, and demographic data to predict driving ability in healthy older adults. *Accident Analysis and Prevention* 42(6): 1759-1768.

Holland, C.A., and Rabbitt, P.M.A. (1994). The problems of being an older driver: Comparing the perceptions of an expert group and older drivers. *Applied Ergonomics* 25(1): 17-27.

Hopkins, R.W., Kilik, L., Day, D.J.A., Rows, C., and Tseng, H. (2004). Driving and dementia in Ontario: A quantitative assessment of the problem. *Canadian Journal of Psychiatry* 49(7): 434-438.

Horswill, M.S., Sullivan, K., Lurie-Beck, J.K., and Smith, S. (2013). How realistic are older drivers' ratings of their driving ability? *Accident Analysis and Prevention* 50: 130-137.

Howe, E. (2007). Initial screening of patients for Alzheimer's disease and minimal cognitive impairment. Psychiatry (Edgmont) 2007 4(7): 24-27.

Hull, M. (1998). Assessing fitness to drive: National uniform guidelines for health professionals. In: Proceedings of Road Safety Research, Policing and Education Conference, Wellington, New Zealand, November 16-17, 1998, pp. 374-378. Vol. 2. Wellington, New Zealand: Land Transport Safety Authority.

Hull, M., Howard, S., Leung, Y., and Cairney, P. (1998). Technological support for older driver testing. In: Proceedings of Road Safety Research, Policing and Education Conference, Wellington, New Zealand, November, 16-17, 1998, pp. 364-373. Vol. 2: Wellington, New Zealand: Land Transport Safety Authority.

Hunt, L. (2010). Driving rehabilitation: Frequently asked questions and answers. *Occupational Therapy Now* 12(5): 19.

Hunt, L.A., Murphy, C.F., Carr, D., Duchek, J.M., and Buckles, V. (1997). Reliability of the Washington University Road Test. *Archives of Neurology* 54: 707-712.

Ingley, S., Chinnaswamy, S., Devakumar, M., Bell, D., and Tranter, R. (2009). A community based survey of cognitive functioning, highway-code performance and traffic accidents in a cohort of older drivers. *International Journal of Geriatric Psychiatry* 24(3): 274-253.

Innes, C.R., Jones, R.D., Anderson, T.J., Hollobon, S.G., and 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.

Innes, C.R., Jones, R.D., Dalrymple-Alford, J.C., Hayes, S., Hollobon, S., Severinsen, J., Smith, G., Nicholls, A., and Anderson, T.J. (2007). Sensory-motor and cognitive tests predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences* 260(1-2): 188-198.

Innes, C.R., Jones, R.D., Dalrymple-Alford, J.C., Severinsen, J., and Gray, J. (2009). Prediction of driving ability in people with dementia- and non- dementia-related brain disorders. In: Driving Assessment 2009: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Big Sky, Montana, June 22-25, 2009. Paper 47. Iowa City, IA: University of Iowa Public Policy Center.

Inouye, S.K., Robison, J.T., Froehlich, T.E., and Richardson, E.D. (1998). The time and change test: a simple screening test for dementia. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 53(4): M281-M286.

Isaacs, B., and Akhtar, A.J. (1972). The set test: a rapid test of mental function in old people. *Age and Ageing* 1(4):222-226.

Israely, R. (2009). Cognitive and performance differences among novice and experienced drivers. In: Wolf-Rüdiger Nickel, Gunnar Meinhard and Ivika Born: Fit to Drive, 4th International Traffic Expert Congress, Tallinn 2009, Proceedings, 87-94.

Iverson, D.J., Gronseth, G.S., Reger, M.A., Classen, S., Dubinsky, R.M., and Rizzo, M. (2010). Practice Parameter update: Evaluation and management of driving risk in dementia. *Neurology* 74: 1316-1324.

Jagacinski, R.J., Greenberg, N., Liao, M.J., and Wang, J. (1992). Manual performance of older and younger adults with supplementary auditory cues. In: Proceedings of the Human Factors Society 36th Annual Meeting, Vol. I, pp. 161-165. Santa Monica, CA.: Human Factors Society.

Janke, M.K. (2001). Assessing older drivers. Two studies. *Journal of Safety Research* 32(1): 43-74.

Janke, M.K., and Eberhard, J.W. (1998). Assessing medically impaired older drivers in a licencing agency setting. *Accident Analysis and Prevention* 30(3): 347-361.

Johnson, L. (2007). The ability of a visual-motor (Frostig) worksheet to predict the on-road performance of older drivers as evaluated by video and GPS. Paper presented at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Jones, M.H. (1991). Perception and hazard detection tests. In: Summary of Proceedings of the Conference on Driver Competency Assessment, pp. 39-40. Sacramento, CA.: California Department of Motor Vehicles.

Jorm, A.F. (2004). The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review. *International Psychogeriatrics* 16(3): 275-293.

Jorm, A.F., Broe, G.A., Creasey, H., Sulway, M.R., Dent, O., Fairley, M.J., Kos, S.C., and Tennant, C. (1996). Further data on the validity of the informant questionnaire on cognitive decline in the elderly (IQCODE). *International Journal of Geriatric Psychiatry* 11(2): 131-139.

Justiss, M.D. (2005). Development of a behind-the-wheel driving performance assessment for older adults. Ph.D. Dissertation. Gainesville, FL.: University of Florida.

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

Kaleem, M.A., Munoz, B.E., Munro, C.A., Gower, E.W., and West, S. K. (2012). Visual characteristics of elderly night drivers in the Salisbury Eye Evaluation Driving Study. *Investigative Ophthalmology & Visual Science* 53(9): 5161-5167.

Kamenoff, I. (2008). Assessing elderly people to drive: Practical considerations. *Australian Family Physician* 37(9): 727-732.

Karrasch, M., Sinervä, E., Grönholm, P., Rinne, J., and Laine, M. (2005). CERAD test performances in amnestic mild cognitive impairment and Alzheimer's disease. *Acta Neurologica Scandinavica* 111(3): 172-179.

Kaufman, D.M., Weinberger, M., Strain, J.J., and Jacobs, J.W. (1979). Detection of cognitive deficits by a brief mental status examination: The Cognitive Capacity Screening Examination, a reappraisal and a review. *General Hospital Psychiatry* 1(3): 247-255.

Kawano, N., Iwamoto, K., Sasada, K. Kohmura, K., Yamamoto, M., Ebe, K., Iidaka, T., and Ozaki, N. (2011). The effects of mild cognitive impairment on driving performance in elderly drivers. *Neuroscience Research* 71: e389.

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

Kay, L.G., Bundy, A.C., Clemson, L., Cheal, B., and 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.

Kay, L.G., Bundy, A.C., and Clemson, L.M. (2009). Predicting fitness to drive in people with cognitive impairments by Using DriveSafe and DriveAware. *Archives of Physical Medicine and Rehabilitation* 90(9): 1514-1522.

Kay, L.G., Bundy, A.C., and Clemson, L.M. (2008). Predicting fitness to drive using the visual recognition slide test (USyd). *American Journal of Occupational Therapy* 62(2): 187-197.

Kay, L.G., Bundy, A.C., and Clemson, L.M. (2009). Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. *Disability and Rehabilitation* 31(13): 1074-1082.

Keall, M., and Frith, W. (2002). Association between number of driving test failures and crash liability of older drivers in New Zealand. In: 2002 Road Safety Research, Policing and Education Conference, November 4-5, 2002, Adelaide, Australia, p. 313. Adelaide, South Australia: Transport South Australia.

Keall, M., and Frith, W. (2004). Association between older driver characteristics, on-road driving test performance, and crash liability. *Traffic Injury Prevention* 5(2): 112-116.

Kipps, C.M., and Hodges, J.R. (2005). Cognitive assessment for clinicians. *Journal of Neurology, Neurosurgery and Psychiatry* 76(Suppl. I): i22-i30.

Klein, J.L. (1995). Elderly Drivers: The Need for Tailored License Renewal Procedures. *The Elder Law Journal* 3(10): 309-377.

Knopman, D.S., Knudson, D., Yoes, M.E., and Weiss, D.J. (2000). Development and standardization of a new telephonic cognitive screening test: the Minnesota Cognitive Acuity Screen (MCAS). *Neuropsychiatry, Neuropsychology and Behavioral Neurology* 13(4): 286-296.

Konstantopoulos, P., Chapman, P., and Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis and Prevention* 42(3): 827-834.

Korner-Bitensky, N. (2007). Older driver retraining: A systematic review of evidence of effectiveness. Paper presented at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Korner-Bitensky, N. (2007). Older individuals' perceptions regarding driving and a driver refresher program: Focus group findings. Paper presented at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Korner-Bitensky, N. (2008). A national blueprint for injury prevention in older drivers (Oral). In: Proceedings of the Canadian Multidisciplinary Road Safety Conference XVIII. Whistler, BC.: Canadian Association of Road Safety Professionals; and the Insurance Corporation of British Columbia.

Korner-Bitensky, N., and 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.

Korner-Bitensky, N., Bitensky, J., Sofer, S., Man-Son-Hing, M., and Gelinas, I. (2006). Driving evaluation practices of clinicians working in the United States and Canada. *American Journal of Occupational Therapy* 60(4): 428-434.

Korner-Bitensky, N., Menon, A., von Zweck, C., and Van Benthem, K. (2010). A national survey of older driver refresher programs: Practice readiness for a rapidly growing need. *Physical and Occupational Therapy in Geriatrics* 28(3): 205-214.

Korner-Bitensky, N., Menon, A., von Zweck, C., and van Benthem, K. (2010). Occupational Therapists' Capacity-Building Needs Related to Older Driver Screening, Assessment, and Intervention: A Canadawide Survey. *American Journal of Occupational Therapy* 64(2): 316-324.

Kowalski, K., and Tuokko, H. (2007). On-Road Driving Assessment of Older Adults: A Review of the Literature. Victoria, BC.: University of Victoria, Centre on Aging; and Justice Institute of British Columbia. December.

Krishnasamy, C., and Unsworth, C. (2012). Validity of the Occupational Therapy-Drive Home Maze Test for right and left handed test takers. *Accident Analysis and Prevention* 49(2): 274-277.

Krüger, H.P. (2006). The older driver - enabling lifelong mobility as a social task. In: Wolf-Rüdiger Nickel & Pierangelo Sardi: Fit to Drive, 1st International Traffic Expert Congress, Berlin 2006, Proceedings, 75-78.

Kuslansky, G., Buschke, H., Katz, M., Sliwinski, M., and Lipton, R.B. (2002). Screening for Alzheimer's Disease: The Memory Impairment Screen Versus the Conventional Three-Word Memory Test. *Journal of the American Geriatrics Society* 50(6): 1086-1091.

Lafont, S., Marin-Lamellet, C., Paire-Ficout, L., Thomas-Anterion, C., Laurent, B., and Fabrigoule, C. (2010). The Wechsler Digit Symbol Substitution Test as the Best Indicator of the Risk of Impaired Driving in Alzheimer Disease and Normal Aging. *Dementia and Geriatric Cognitive Disorders* 29(2): 154-163.

Lambert, L.D., Fleury, M. (1994). Age, cognitive style, and traffic signs. *Perceptual and Motor Skills* 78(2): 611-624.

Langford, J. (2004). Model Licence Re-Assessment Procedure for Older Drivers: Stage 2 Research. AP-R259/04. Sydney, NSW.: Austroads.

Langford, J. (2008). Usefulness of off-road screening tests to licencing authorities when assessing older driver fitness to drive. *Traffic Injury Prevention* 9(4): 328-335.

Langford, J., Bohensky, M., Koppel, S., and Newstead, S. (2008). Do age-based mandatory assessments reduce older drivers' risk to other road users? *Accident Analysis and Prevention* 40(6): 1913-1918.

Langford, J., Dow, J., and Turmel, E. (2011). Ageing and medical conditions: Implications for managing older driver safety. Proceedings of the CMRSC-XXI, Halifax, Nova Scotia, May 8-11, 2011. Halifax, NS.: Canadian Association of Road Safety Professionals.

Langford, J., Fitzharris, M., Koppel, S., and Newstead, S. (2004). Effectiveness of mandatory license testing for older drivers in reducing crash risk among urban older Australian drivers. *Traffic Injury Prevention* 5(4): 326-335.

Larish, J.F., Kramer A.F., DeAntona, J., and Strayer, D.L. (1993). Aging and dual-task performance. In: Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, pp. 162-166.

Larner, A.J. (2007). DemTect: 1-year experience of a neuropsychological screening test for dementia. *Age and Ageing* 36(3): 326-327.

Laux, L.F., and Brelsford, J. Jr. (1990). Age-Related Changes in Sensory, Cognitive, Psychomotor and Physical Functioning and Driving Performance in Drivers Aged 40 to 92. Washington, DC.: AAA Foundation for Traffic Safety. July.

Laycock, K.M. (2011). Driver Assessment: Uncertainties inherent in current methods. *British Columbia Medical Journal* 53(2): 74-78.

Lee, H.C. (2002). The validity of driving simulator to measure on-road driving performance of older drivers. Paper presented at the 24th Conference of Australian Institutes of Transport Research, Sydney, Australia.

Lee, H.C., Cameron, D., and Lee, A.H. (2003). Assessing the driving performance of older adult drivers: on-road versus simulated driving. *Accident Analysis and Prevention* 35(5): 797-803.

Lee, H.C., Lee, A.H., and Cameron, D. (2003). Validation of a driving simulator by measuring the visual attention skill of older adult drivers. *American Journal of Occupational Therapy* 57(3).

Lerner, N.D. (1993). Brake perception-reaction times of older and younger drivers. In: Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, pp. 206-210.

Lerner, N.D. (1994). Age and driver time requirements at intersections. In: Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, pp. 842-846. Santa Monica, CA.: Human Factors and Ergonomics Society.

Lerner, N.D. (1994). Research needs and driver risk perception. In: University of Michigan Symposium on Risk-Taking Behavior and Traffic Safety Proceedings, May 24-25, 1993, pp. 37-40. Ann Arbor, MI. University of Michigan Medical Center; and University of Michigan Transportation Research Institute.

Lesch, M.F. (2003). Comprehension and memory for warning symbols: Age-related differences and impact of training. *Journal of Safety Research* 34(5): 495-505.

Lesikar, S.E., Gallo, J.J., Rebok, G.W., and Keyl, P.M. (2002). Prospective study of brief neuropsychological measures to assess crash risk in older primary care patients. *Journal of the American Board of Family Practice* 15(1): 11-19.

Levin, L., Dukic, T., Henriksson, P., Mårdh, S., and Sagberg, F. (2009). Older Car Drivers in Norway and Sweden: Studies of Accident Involvement, Visual Search Behaviour, Attention and Hazard Perception. VTI Rapport 656A. Linköping, Sweden: Swedish National Road and Transport Research Institute (VTI).

Lincoln, N.B., Radford, K.A., Lee, E., and Reay, A.C. (2006). The assessment of fitness to drive in people with dementia. *International Journal of Geriatric Psychiatry* 21(11): 1044-1051.

Lincoln, N.B., Taylor, J.L., Vella, K., Bouman, W.P., and Radford, K.A. (2010). A prospective study of cognitive tests to predict performance on a standardised road test in people with dementia. International *Journal of Geriatric Psychiatry* 25(5): 489-496.

Lloyd, S., Cormack, C.N., Blais, K., Messeri, G., McCallum, M.A., Spicer, K., and Morgan, S. (2001). Driving and dementia: a review of the literature. *Canadian Journal of Occupational Therapy* 68(3): 149-156.

Lococo, K.H., and Staplin, L. (2005). Strategies for Medical Advisory Boards and Licencing Review. DOT HS 809 874. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. July.

Lonie, J.A., Tierney, K.M., and Ebmeier, K.P. (2009). Screening for mild cognitive impairment: a systematic review. *International Journal of Geriatric Psychiatry* 24(9): 902-915.

Lundberg, C., Caneman, G., Samuelsson, S., Hakamies-Blomqvist, L., and Almkvist, O. (2003). The assessment of fitness to drive after a stroke: The Nordic Stroke Driver Screening Assessment. Scandinavian *Journal of Psychology* 44: 23-30.

Lundberg, C., and Hakamies-Blomqvist, L. (2003). Driving tests with older patients: Effect of unfamiliar versus familiar vehicle. *Transportation Research Part F* 6F(3): 163-173.

Lundberg, C., Hakamies-Blomqvist, L., Almkvist, O., and Johansson, K. (1998). Impairments of some cognitive functions are common in crash-involved older drivers. *Accident Analysis and Prevention* 30(3): 371-377.

MacGregor, J.M., Freeman, D.H., Jr., and Zhang, D. (2001). A traffic sign recognition test can discriminate between older drivers who have and have not had a motor vehicle crash. *Journal of the American Geriatrics Society* 49(4): 466-469.

Mahncke, H.W., Connor, B.B., Appelman, J., Ahsanuddin, O.N., Hardy, J.L., Wood, R.A., Joyce, N.M., Boniske, T., Atkins, S.M., and Merzenich, M.M. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. *Proceedings of the National Academy of Sciences* 103(33): 12523-12528.

Maki, N., Ikeda, M., Hokoishi, K., Nebu, A., Komori, K., Hirono, N., and Tanabe, H. (2000). The validity of the MMSE and SMQ as screening tests for dementia in the elderly general population - A study of one rural community in Japan. *Dementia and Geriatric Cognitive Disorders* 11(4): 193-196.

Makishita, H., and Matsunaga, K. (2008). Differences of drivers' reaction times according to age and mental workload. *Accident Analysis and Prevention* 40(2): 567-575.

Manivannan, P., Czaja, S.J., Drury, C.G., and Ip, C.M. (1993). The impact of age on visual search performance. In: *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, pp. 172-176.

Man-Son-Hing M, Marshall SC, and Wilson KG. (2007). Systematic review of driving risk and the efficacy of compensatory strategies in persons with dementia. *Journal of the American Geriatric Society* 55(6):878-884.

Marottoli, R.A., and Richardson, E.D. (1998). Confidence in, and self-rating of, driving ability among older drivers. *Accident Analysis and Prevention* 30(3): 331-336.

Marottoli, R.A., Cooney, L.M.,Jr., Wagner, R., Doucette, J., and Tinetti,M.E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine* 121(11): 842-846.

Marshall, S. (2005). Assessing medical fitness to drive in the older driver. *Elder Care CMA Leadership Series*: 33-35.

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.

Marshall, S.C., Man-Son-Hing, M., Molnar, F., Hunt, L., and Finestone, H. (2004). Contribution of the on-road driving assessment evaluation elements for predicting pass or fail. In: Proceedings of the Canadian Multidisciplinary Road Safety Conference XIV, Paper 6a4. Ottawa, Ont.: Canadian Association of Road Safety Professionals.

Martin, A.J., Marottoli, R., and O'Neill, D. (2009). Driver assessment for maintaining mobility and safety in drivers with dementia. Cochrane Database for Systematic Reviews, 2009 (1): CD006222.

Mate-Kole, C.C., Conway, J., Catayong, K., Bieu, R., Sackey, N.A., Wood, R., Fellows, R. 2009. Validation of the Revised Quick Cognitive Screening Test. *Archives of Physical Medicine and Rehabilitation*, 90(9): 1469-1477.

Matthews, M.L., and Moran, A.R. (1986). Age differences in male drivers' perception of accident risk: The role of perceived driving ability. *Accident Analysis and Prevention* 18(4): 299-313.

Mazer, B.L., Korner-Bitensky, N.A., and Sofer, S. (1998). Predicting ability to drive after stroke. *Archives of Physical Medicine and Rehabilitation* 79(7): 743-750.

McCarthy, D.P. (2005). Outcomes Evaluation of the Assessment of Driving Related Skills (ADReS). Gainesville, FL.: University of Florida.

McCarthy, D.P., Mann, W.C., and Lanford, D. (2009). Process and Outcomes Evaluation of Older Driver Screening Programs: The Assessment of Driving-Related Skills (ADReS) Older-Driver Screening Tool. DOT HS 811 113. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. May.

McCurry, S.M., Edland, S.D., Teri, L., Kukull, W.A., Bowen, J.D., McCormick, W.C., and Larson, E.B. (1999). The cognitive abilities screening instrument (CASI): data from a cohort of 2524 cognitively intact elderly. *International Journal of Geriatric Psychiatry* 14(10): 882-888.

McGee, P., and Tuokko, H. (2003). The Older and Wiser Driver: A Self-Assessment Program. Victoria, BC.: University of Victoria, Centre on Aging. January.

McGwin, G. Jr., Owsley, C., and Ball, K. (1997). Visual and cognitive impairments and injurious motor vehicle crashes in the elderly. In: Proceedings of the 41st Annual Conference of the Association for the Advancement of Automotive Medicine, pp. 427-428. Des Plaines, IL.: Association for the Advancement of Automotive Medicine.

McKenna, P., Jefferies, L., Dobson, A., and Frude, N. (2004). The use of a cognitive battery to predict who will fail an on-road driving test. *The British Journal of Clinical Psychology* 43(Pt.3): 325-336.

McKnight, A.J., and McKnight, A.S. (1999). Multivariate analysis of age-related driver ability and performance deficits. *Accident Analysis and Prevention* 31(5): 445-454.

Medical Commission on Accident Prevention. (1985). Medical Aspects of Fitness to Drive: A Guide for Medical Practitioners. Great Britain: Eaton Press Ltd.

Mendiondo, M.S., Ashford, J.W., Kryscio, R.J., and Schmitt, F.A. (2003). Designing a Brief Alzheimer Screen (BAS). *Journal of Alzheimer's Disease* 5(5): 391-398.

Meng, A. and Siren, A. (2012). Cognitive problems, self-rated changes in driving skills, driving-related discomfort and self-regulation of driving in old drivers. *Accident Analysis and Prevention* 49(2): 322-329.

Mercier, C.R., Mercier, J.M., O'Boyle, M.W., Strahan, R.F., Larpenteur-Gradwell, W., and Gill, H. (1997). Validation of Relationship of Cognitive Skills Losses to Driving Performance. Final Report. Ames, IA.: Midwest Transportation Center. December.

Mercier, C.R., Shelley, M.C. II, Adkins, G.H., and Mercier, J.M. (1999). Age and gender as predictors of injury severity in broadside and angle vehicular collisions. In: Highway and Traffic Safety, Enforcement, Older Person Mobility, and Women's Issues Research. Transportation Research Record 1693: 50-61.

Messinger-Rapport, B.J. (2003). Assessment and counseling of older drivers: A guide for primary care physicians. *Geriatrics* 58(12): 16-24.

Meulen, E.F.J., Schmand, B., van Campen, J.P., de Koning, S.J., Ponds, R.W., Scheltens, P., and Verhey, F.R. (2004). The seven minute screen: a neurocognitive screening test highly sensitive to various types of dementia. *Journal of Neurology, Neurosurgery and Psychiatry* 75: 700-705.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., and Hodges, J.R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry* 21: 1078-1085.

Mitchell, R.K., Castleden, C.M., and Fanthome, Y.C. (1995). Driving, Alzheimer's Disease and Ageing: A potential cognitive screening device for all elderly drivers. *International Journal of Geriatric Psychiatry* 10: 865-869.

Mitrushina, M., Abara, J., and Blumenfeld, A. (1994). The neurobehavioral cognitive status examination as a screening tool for organicity in psychiatric patients. *Hospital and Community Psychiatry* 45(3): 252-256.

Mohs, R., Petersen, R., Knopman, D., Ferris, S., Ernesto, C., Grundman, M., Sano, M., Bieliauskas, L., Geldmacher, D., Clark, C., and Thal, L. (1997). Development of Cognitive Instruments for Use in Clinical Trials of Antidementia Drugs: Additions to the Alzheimer's Disease Assessment Scale That Broaden its Scope. *Alzheimer Disease and Associated Disorders* 11(2): 13.

Molloy, D.W., Standish, T.I.M., and Lewis, D.L. (2005). Screening for Mild Cognitive Impairment: Comparing the SMMSE and the ABCs. *Canadian Journal of Psychiatry* 50(1): 52.

Molnar, F.J., Byszewski, A.M., Marshall, S.C., and Man-Son-Hing, M. (2005). In-office evaluation of medical fitness to drive: Practical approaches for assessing older people. *Canadian Family Physician* 51(3): 372-379.

Molnar, F.J., Marshall, S.C., Man-Son-Hing, M., Wilson, K.G., Byszewski, A.M., and Stiell, I. (2007). Acceptability and concurrent validity of measures to predict older driver involvement in motor vehicle crashes: an Emergency Department pilot case-control study. *Accident Analysis and Prevention* 39(5): 1056-1063.

Molnar, F.J., Patel, A., Marshall, S.C., Man-Son-Hing, M., and Wilson, K.G. (2006). Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: a systematic review. *Journal of the American Geriatric Society* 54(12): 1809-1824.

Molnar, L.J., Eby, D.W., Kartje, P.S., and St. Louis, R.M. (2010). Increasing self-awareness among older drivers: The role of self-screening. *Journal of Safety Research* 41(4): 367-373.

Monterde I., Bort, H., and Risser, R. (2007). Senior citizens - keep them driving the car? A presentation using results of the EU-project SIZE. In: Ralf Risser & Wolf-Rüdiger Nickel: Fit to Drive, 2nd International Traffic Expert Congress, Vienna 2007, Proceedings, 75-83.

Moorhouse, P., Hamilton, L., Fisher, T., and Rockwood, K. (2011). Barriers to assessing fitness to drive in dementia in Nova Scotia: Informing strategies for knowledge translation. *Canadian Geriatrics Journal* 14(3): 61-65.

Mosimann, U.P., Bachli-Bietry, J., Boll, J., Bopp-Kistler, I., Donati, F., Kressig, R.W., Martensson, B., Monsch, A.U., Muri, R., Nef, T., Rothenberger, A., Seeger, R., von Gunten, A., and Wirz, U. (2012). Consensus recommendations for the assessment of fitness to drive in cognitively impaired patients. *Praxis* 101(7): 451-64.

Mourant, R.R., Tsai, F.J., Al-Shihabi, T., and Jaeger, B.K. (2001). Measuring divided-attention capability of young and older drivers. In: Americans with Disabilities Act; Driver and Vehicle

Modelling; Situation Awareness; Licencing; Driver Behavior; Enforcement; Trucks; and Motorcycles. Transportation Research Record 1779: 40-45.

Mundt, J.C., Freed, D.M., and Greist, J.H. (2000). Lay Person-Based Screening for Early Detection of Alzheimer's Disease: Development and Validation of an Instrument. *The journals of gerontology Series B, Psychological sciences and social sciences* 55(3): 163-170.

Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., and Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53: 695-699.

National Highway Traffic Safety Administration. (2007). DriveWell: Promoting Older Driver Safety and Mobility in Your Community. DOT HS 809 838. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. January.

National Highway Traffic Safety Administration. (2009). Driving Transitions Education: Tools, Scripts, and Practice Exercises. DOT HS 811 152. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. November.

National Transportation Safety Board. (2004). Medical Oversight of Noncommerical Drivers. Highway Special Investigation Report. NTSB/SIR-04/01. Washington, DC.: National Transportation Safety Board. November.

Neitch, S.M., Madero, G., Maynard, S. 2011. Driving assessment results in patients with a diagnosis of dementia. *West Virginia Medical Journal* 107(3): 54-8.

Nelson, T.M., Evelyn, B., Taylor, R. 1993. Experimental intercomparisons of younger and older driver perceptions. *International Journal of Aging and Human Development* 36(3): 239-253.

Nemeto, H., Yanagishima, T., and Taguchi, M. (2001). Effect of physical changes in aging on driving performance. In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 24. Iowa City, Iowa: University of Iowa Public Policy Center.

New Zealand Transport Agency. (2009). Medical Aspects of Fitness to Drive: A Guide for Medical Practitioners. Palmerston North, New Zealand: New Zealand Transport Agency. July.

Ni, R., Kang, J.J., and Andersen, G.J. (2010). Age-related declines in car following performance under simulated fog conditions. *Accident Analysis and Prevention* 42(3): 818-826.

Nichols, A.L., Classen, S., McPeek, R., and Breiner, J. (2012). Does personality predict driving performance in middle and older age? An evidence-based literature review. *Traffic Injury Prevention* 13(2): 133-143.

Nouri, F.M., and Lincoln, N.B. (1992). Validation of a cognitive assessment: predicting driving performance after stroke. *Clinical Rehabilitation* 6: 275-281.

Odenheimer, G.L., Beaudet, M., Jette, A. M., Albert, M.S., Grand, L., and Minaker, K.L. (1994). Performance-Based Driving Evaluation of the Elderly Driver: Safety, Reliability, and Validity. *Journal of Gerontology* 49(4): 153-159.

O'Neill, D. (1992). Physicians, elderly drivers and dementia. *The Lancet* 339(8784): 41.

O'Neill, D., and Dobbs, B.M. (2004). Age-related disease, mobility, and driving. In: Transportation in an Aging Society: A Decade of Experience, November 7-9, 1999, Bethesda, Maryland, pp. 56-66. Washington, DC.: Transportation Research Board.

Oswanski, M.F., Sharma, O.P., Raj, S.S., Vassar, L.A., Woods, K.L., Sargent, W.M., and Pitock, R.J. (2007). Evaluation of two assessment tools in predicting driving ability of senior drivers. *American Journal of Physical Medicine and Rehabilitation* 86(3): 190-199.

Ott, B.R., Heindel, W.C., Whelihan, W.H., Caron, M.D., Piatt, P.L., and Dicarlo, M.A. (2003). Maze Test Performance and Reported Driving Ability in Early Dementia. *Journal of Geriatric Psychiatry and Neurology* 16(3): 151-155.

Ott, B., Papandonatos, G., Davis, J., and Barco, P. (2012). Naturalistic Validation of an On-Road Driving Test of Older Drivers. *The Journal of Human Factors and Ergonomics Society* 54(4): 663-674.

Owsley, C., and McGwin, G. Jr. (2004). Association between visual attention and mobility in older adults. *Journal of the American Geriatrics Society* 52(11): 1901-1906.

Owsley, C., Ball, K., McGwin, G.,Jr., Sloane, M.E., Roenker, D.L., White, M.F., and Overley, E.T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association* 279(14): 1083-1088.

Owsley, C., Ball, K., Sloane, M., Roenker, D., and Bruni, J.R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging* 6(3): 403-415.

Papp, K.V., Walsh, S.J., and Snyder, P.J. (2009). Immediate and delayed effects of cognitive interventions in healthy elderly: A review of current literature and future directions. *Alzheimer's and Dementia* 5(1): 50-60.

Parasuraman, R., and Nestor, P. (1993). Attention and driving: Assessment in elderly individuals with dementia. *Clinics in Geriatric Medicine* 9(2): 377-387.

Parker, D. (2001). Can we predict the on-road performance of older drivers? In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 36. Iowa City, Iowa: University of Iowa Public Policy Center.

Parker, D. (2002). Predicting the on-road performance of older drivers. In: Behavioural Research in Road Safety: Proceedings of the 11th Seminar, pp. 25-31. England: Department of the Environment, Transport and the Regions.

Parker, D., Macdonald, L., Sutcliffe, P., and Rabbitt, P. (2001). Confidence and the older driver. *Ageing and Society* 21(2): 169-182.

Parker, D., McDonald, L., Rabbitt, P., and Sutcliffe, P. (2000). Elderly drivers and their accidents: The Aging Driver Questionnaire. *Accident Analysis and Prevention* 32(6): 751-759.

Patomella, A.H., and Kottorp, A. (2005). An evaluation of driving ability in a simulator: A good predictor of driving ability after stroke? In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 15. Iowa City, IA.: University of Iowa Public Policy Center.

Patomella, A.H., Tham, K., Johansson, K., and 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.

Peel, N., and Westmoreland, J. (2000). Transport safety for older drivers: A study of their experiences, perceptions and management needs. In: Handbook and Proceedings of the Road Safety Research, Policing and Education Conference, pp. 223-228. Brisbane, Queensland: Queensland Transport.

Petrucelli, E., and Malinowski, M. (1992). Status of Medical Review in Driver Licencing: Policies, Programs and Standards. DOT HS 807 892. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. March.

Petzoldt, T., Weiß, T., Franke, T., Krems, J.F., and Bannert, M. (2013). Can driver education be improved by computer based training of cognitive skills? *Accident Analysis and Prevention* 50: 1185-1192.

Pierce, S. (2003). The occupational therapist's roadmap to safety for seniors. *Gerontology Special Interest Sectional Quarterly* 26(3): 1-3.

Pietras, T., Shi, Q., and Rizzo, M. (2005). Traffic entry behavior in older drivers with selective attention impairment. Paper 05-1286. In: Transportation Research Board 84th Annual Meeting Compendium of Papers CD-ROM. Washington, DC.: Transportation Research Board.

Pollatsek, A., Romoser, M.R.E., and Fisher, D.L. (2012). Identifying and remediating failures of selective attention in older drivers. *Current Directions in Psychological Science* 21(1): 3-7.

Ponsford, A.S., Viitanen, M., Lundberg, C., and Johansson, K. (2008). Assessment of driving after stroke- A pluridisciplinary task. *Accident Analysis and Prevention* 40(2): 452-460.

Popkin, C.L., Stewart, J.R., and Lacey, J.H. (1981). An examination of the driving performance of drivers with select medical restrictions. In: Proceedings of the 25th Annual Conference of the American Association for Automotive Medicine, pp. 201-214. Morton Grove, IL.: American Association for Automotive Medicine.

Popkin, C.L., Stewart, J.R., and Lacey, J.H. (1983). An examination of the driving performance of medically evaluated drivers. In: Proceedings of the 27th Annual Conference of the American Association for Automotive Medicine, pp. 263-274. Arlington Heights, IL.: American Association for Automotive Medicine.

Porter, M.M., and Tuokko, H.A. (2011). An evaluation of the Roadwise Review: A mixed methods approach. *Traffic Injury Prevention* 12(5): 451-458.

Pyo, G., Elble, R.J., Ala, T., and Markwell, S.J. (2006). The characteristics of patients with uncertain/mild cognitive impairment on the Alzheimer disease assessment scale-cognitive subscale. *Alzheimer Disease and Associated Disorders* 20(1): 16-22.

Racette, L., and Casson, E.J. (2005). The impact of visual field loss on driving performance: Evidence from on-road driving assessments. *Optometry and Vision Science* 82(8): 668-674.

Rackoff, N.J. (1975). An investigation of age-related changes in driver's visual search patterns and driving performance and the relation of tests to basic functional capacities. In: Proceedings of the Human Factors Society 19th Annual Meeting, eds. R.N. Hale, R.A. McKnight and J.R. Moss, pp. 285-288. Dallas, TX: Human Factors Society.

Rackoff, N.J., and Mourant, R.R. (1979). Driving performance of the elderly. *Accident Analysis and Prevention* 11(4): 247-254.

Raffle, P.A. (1977). Driving Fitness. Section of Occupational Medicine. *Proceedings of the Royal Society of Medicine* 70: 240-243.

Rami, L., Molinvevo, J.L., Sanchez-valie, R., Bosch, B., and Villar, A. (2007). Screening for amnestic mild cognitive impairment and early Alzheimer's disease with M @ T (Memory Alteration Test) in the primary care population. *International Journal of Geriatric Society* 22(4): 294-304.

Rami, L., Bosch, B., Sanchez-Valle, R., and Molinuevo, J.L. (2010). The memory alteration test discriminates between subjective memory complaints, mild cognitive impairment, and Alzheimer's disease. *Archives of Gerontology and Geriatrics* 50(2): 171-174.

Ranney, T.A., and Pulling, N.H. (1989). Relation of individual differences in information-processing ability to driving performance. In: Proceedings of the Human Factors Society 33rd Annual Meeting, pp. 965-969. Santa Monica, CA.: Human Factors Society.

Ranney, T.A., and Pulling, N.H. (1990). Performance differences on driving and laboratory tasks between drivers of different ages. In: Human Factors and Safety Research Related to Highway Design and Operation 1990. Transportation Research Record 1281, pp. 3-10. Washington, DC.: Transportation Research Board.

Ranney, T.A., and Simmons, L.A.S. (1992). The effects of age and target location uncertainty on decision making in a simulated driving task. In: Proceedings of the Human Factors Society 36th Annual Meeting, Vol. I, pp. 166-170. Santa Monica, CA.: Human Factors Society.

Ravaglia, G., Forti, P., Maioli, F., Servadei, L., Martelli, M., Brunetti, N., Bastagli, L., and Mariani, E. (2005). Screening for mild cognitive impairment in elderly ambulatory patients with cognitive complaints. Ageing *Clinical and Experimental Research* 17(5): 374-379.

Read, L. (2002). Developing protocols for assessing of fitness to drive in dementia and arthritis. In: Proceedings of the 16th International Conference on Alcohol, Drugs and Traffic Safety, Volume 2, eds. D.R. Mayhew and C. Dussault, p. 605. Quebec, QC.: Société de l'assurance automobile du Québec.

Reger, M.A., Welsh, R.K., Watson, G.S., Cholerton, B., Baker, L.D., and Craft, S. (2004). The relationship between neuropsychological functioning and driving ability in dementia: A meta-analysis. *Neuropsychology* 18(1): 85-93.

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. In: Human Performance, Information Systems, Simulation, and Visualization. Transportation Research Record 2138: 13-19.

Reisman, A. (2011). Surrendering the keys: a doctor tries to get an impaired elderly patient to stop driving. *Health Affairs (Project Hope)* 30(2): 356.

Rimmö, P.A. (2002). Aberrant driving behaviour: homogeneity of a four-factor structure in samples differing in age and gender. *Ergonomics* 45(8):569-582.

Risser, R., and Bell, D. (2011). Fitness to drive and mobility behaviour among senior road users. In: Divera Twisk and Wolf-Rüdiger Nickel: Fit to Drive, 5th International Traffic Expert Congress, The Hague 2011, *Proceedings*, 60-61.

Ritchie, K., and Fuhrer, R. (1992). A Comparative Study of the Performance of Screening Tests for Senile Dementia Using Receiver Operating Characteristics Analysis. *Journal of Clinical Epidemiology* 45(6): 627-637.

Ritchie, K., Robine, J.M., Letenneur, L. and Dartigues, J.F. (1994). Dementia-Free Expectancy in France. *American Journal of Public Health* 84(2): 232-236.

Rizzo, M., Reinach, S., McGehee, D., and Dawson, J. (1997). Simulated car crashes and crash predictors in drivers with Alzheimer disease. *Archives of Neurology* 54(5): 545-551.

Rizzo, M., Shi, Q., Dawson, J., Anderson, S.W., Kellison, I., and Pietras, T. (2005). Stops for Cops: Impaired response implementation in older drivers with cognitive decline. Paper 05-1095. In: Transportation Research Board 84th Annual Meeting Compendium of Papers CD-ROM. Washington, DC.: Transportation Research Board.

Rizzo, M., Shi, Q., Dawson, J.D., Anderson, S.W., Kellison, I., and Pietras, T. (2005). Stops for Cops: Impaired response implementation for older drivers with cognitive decline. In: Safety: Older Drivers; Traffic Law Enforcement; Management; School Transportation; Emergency Evacuation; Truck and Bus; and Motorcycles. Transportation Research Record 1922: 1-8.

Rizzo, M. (2011). Impaired driving from medical conditions: a 70-year-old man trying to decide if he should continue driving. *Journal of the American Medical Association* 305(10): 1018-26.

Rodriguez, L., and Guinea, F.S. (2004). Elderly and brain-damaged drivers in Spain. *Revista de neurologia* 38(8): 785.

Rogers, W.A., Gilbert, D.K., and Fisk, A.D. (1993). Ability-performance relationships in memory skill tasks for young and old adults. In: Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, pp. 167-171.

Romoser, M.R.E., and Fisher, D.L. (2009). Effects of cognitive and physical decline on older drivers' side-to-side scanning for hazards while executing turns. In: Driving Assessment 2009: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Big Sky, Montana, June 22-25, 2009. Paper 09. Iowa City, IA: University of Iowa Public Policy Center.

Romoser, M.R.E., Fisher, D.L., Mourant, R., Wachtel, J., and Sizov, K. (2005). The use of a driving simulator to assess senior driver performance: Increasing situational awareness through post-drive one-on-one advisement. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 67. Iowa City, IA.: University of Iowa Public Policy Center.

Romoser, M.R.E., Fisher, D.L., Mourant, R., Wachtel, J., Sizov, K., Kennedy, S.I., and Andra, M. (2005). Use of a driving simulator to assess older adults' critical driving skills. In: Transportation Research Board 84th Annual Meeting Compendium of Papers CD-ROM. Paper 05-0589. Washington, DC.: Transportation Research Board.

Rosenbloom, S. (2003). Older drivers: Should we test them off the road? *Access* 23(Fall): 8-13.

Rosenbloom, S. (2007). Differences in perceptions of driving skills: Older drivers and adult children of older drivers in the United Kingdom. In: Highway Safety, School Transportation, and Emergency Evacuation. Transportation Research Record 2009: 15-22.

Ross, L.A., Browning, C., Luszcz, M.A., Mitchell, P., and Anstey, K.J. (2011). Age-Based Testing for Driver's License Renewal: Potential Implications for Older Australians. *Journal of the American Geriatrics Society* 59(2): 281-285.

Rossi, D.G., and Flint, S.J. (1988). An Evaluation of Mature Driver Performance. New Mexico: Traffic Safety Bureau, New Mexico Highway and Transportation Department.

Rouleau, S., Antoun, M., Billette, V., Gautier, M., Hartman, J., Ménard, I., and Mazer, B. (2007). Driving and mental health disorders: Results of a survey on driving fitness and services. Poster presentation at the 7th National Workshop for Driver Rehabilitation Specialists, Ottawa, Ontario, May 4-5, 2007. TP 11489. Ottawa, ON.: Transport Canada; and the Ottawa Hospital.

Ruff, R.M., and Parker, S.B. (1993). Gender-specific and age-specific changes in motor speed and eye-hand coordination in adults- normative values for the finger tapping and grooved pegboard tests. *Perceptual and Motor Skills* 76(3) Part 2: 1219-1230.

Sager, M.A., Hermann, B.P., La Rue, A., and Woodward, J.L. (2006). Screening for Dementia in Community-based Memory Clinics. *Wisconsin Medical Journal* 105(7): 25.

Saka, E., Mihci, E., Topcuoglu, M.A., and Balkan, S. (2006). Enhanced Cued Recall has a high utility as a screening test in the diagnosis of Alzheimer's disease and mild cognitive impairment in Turkish People. *Archives of Clinical Neuropsychology* 21(7): 745-751.

Salvucci, D.D., Boer, E.R., and Liu, A. (2001). Toward an integrated model of driver behavior in cognitive architecture. In: Americans with Disabilities Act; Driver and Vehicle Modelling; Situation Awareness; Licencing; Driver Behavior; Enforcement; Trucks; and Motorcycles. Transportation Research Record 1779: 9-16.

Sarasqueta, C., Bergareche, A., Arce, A., Lopez de Munain, A., Poza, J.J., De La Puente, E., Urtasun, M., Emparanza, J.E., and Martí Masso, J.F. (2001). The validity of Hodkinson's Abbreviated Mental Test for dementia screening in Guipuzcoa, Spain. *European Journal of Neurology* 8(5): 435-440.

Saxton, J., Morrow, L., Eschman, A., Archer, G., Luther, J., and Zuccolotto, A. (2009). Computer Assessment of Mild Cognitive Impairment. *Postgraduate Medicine* 121(2): 177-185.

Schlag, B. (1993). Elderly drivers in Germany- fitness and driving behavior. *Accident Analysis and Prevention* 25(1): 47-55.

Schubert, W., and Mattern, R. (2006). Criteria for evaluating future models for assessing physical and mental fitness for motor vehicle drivers. In: Fit to Drive 1st International Traffic Expert Congress, eds. W.R. Nickel and P. Sardi, pp. 106-110. Bonn, Germany: Kirschbaum Verlag.

Selander, H., Johansson, K., Lundberg, C., and 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.

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

Shanmugaratnam, S., Kass, S.J., and Arruda, J.E. (2010). Age differences in cognitive and psychomotor abilities and simulated driving. *Accident Analysis and Prevention* 42(3): 802-808.

Sharlin, E., Watson, B., Sutphen, S., Liu, L., Lederer, R., and Frazer, J. (2009). A tangible user interface for assessing cognitive mapping ability. I*nternational Journal of Human - Computer Studies* 67(3): 269-278.

Silva, M.T., Laks, J., and Engelhardt, E. (2009). Neuropsychological tests and driving in dementia: a review of recent literature. *Revista da Associçäo Médica Brasileira* 55(4): 484-488.

Silverman, J.M., Breitner, J.C., Mohs, R.C., and Davis, K.L. (1986). The Reliability of the Family History in Genetic Studies of Alzheimer's Disease and Related Dementia. *American Journal of Psychiatry* 143(10): 1279-1282.

Simoes, A. (2003). The cognitive training needs of older drivers. *Recherche Transports Securite* 79: 145-155.

Simoes, A. (1994). A training programme designed to improve the driving performance of elderly drivers. In: Road Safety in Europe and Strategic Highway Research Program (SHRP). Preprint. Traffic Safety. 27/9. Lille, France, September 26-28, 1994.

Sims, R.V., Owsley, C., Allman, R.M., Ball, K., and Smoot, T.M. (1998). Preliminary assessment of the medical and functional factors associated with vehicle crashes in older adults. *Journal of the American Geriatric Society* 46(5): 556-561.

Siren, A., and Meng., A. (2012). Cognitive screening of older drivers does not produce safety benefits. *Accident Analysis and Prevention* 45: 634-8.

Skaar, N., Rizzo, M., and Stierman, L. (2003). Traffic entry judgements by aging drivers. In: Proceedings of the 2nd International International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp. 92-97. Iowa City, IA.: University of Iowa Public Policy Center.

Skaar, N., Rizzo, M., Bateman, K., and Anderson, S.W. (2001). Effects of speed of visual processing training upon non-visual attention in "at risk" older drivers. In: Driving Assessment 2001: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Aspen, Colourado, August 14-17, 2001. Paper 64. Iowa City, Iowa: University of Iowa Public Policy Center.

Snellgrove, C.A. (2005). Cognitive Screening for the Safe Driving Competence of Older People with Mild Cognitive Impairment or Early Dementia. Civic Square, Australian Capital Territory: Australian Transport Safety Bureau.

Snyder, K.M., and Ganzini, L. (2009). Outcomes of Oregon's Law Mandating Physician Reporting of Impaired Drivers. *Journal of Geriatric Psychiatry and Neurology* 22(3): 161-165.

Soderstrom, C.A. (2008). Applying research to licencing agency screening of seniors for fitness to drive. *Gerontology and Geriatrics Education* 29(4): 310-325.

Soderstrom, C.A., and Joyce, J.J. (2008). Medical review of fitness to drive in older drivers: The Maryland experience. *Traffic Injury Prevention* 9(4): 342-349.

Solomon, P.R., and Pendlebury, W.W. (1998). Recognition of Alzheimer's Disease: the 7 Minute Screen. *Family Medicine* 30(4): 265-271.

Sotaniemi, M., Pullianen, V., Hokkanen, L., Pirttila, T., Hallikainen, I., Soininen, H., and Hanninen, T. (2012). CERAD-neuropsychological battery in screening mild Alzheimer's disease. *Acta Neurologica Scandinavica* 125(1): 16-23.

Staplin, L., Gish, K.W., Lococo, K.H., Joyce, J.J., and Sifrit, K.J. (2013). The Maze Test: A significant predictor of older driver crash risk. *Accident Analysis and Prevention* 50: 483-489.

Staplin, L., and Hunt, L. (2004). Driver programs. In: Transportation in an Aging Society: A Decade of Experience, November 7-9, 1999, Bethesda, Maryland, pp. 69-94. Washington, DC.: Transportation Research Board.

Staplin, L., and Lococo, K.H. (2003). Model Driver Screening and Evaluation Program: Guidelines for Motor Vehicle Administrators. DOT HS 809 581. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. May.

Staplin, L., and Lyles, R.W. (1991). Age differences in motion perception and specific traffic maneuver problems. In: Highway Safety: Older Drivers, Seat Belts, Alcohol, Motorcycles, and Pedestrians 1991. Transportation Research Record 1325, pp. 23-33. Washington, DC.: Transportation Research Board.

Staplin, L., Lococo, K.H., Gish, K.W., and Decina, L.E. (2003). Model Driver Screening and Evaluation Program. Final Technical Report. Volume 1: Project Summary and Model Program Recommendations. DOT HS 809 582. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. May.

Staplin, L., Lococo, K.H., Gish, K.W., and Decina, L.E. (2003). Model Driver Screening and Evaluation Program. Final Technical Report. Volume II: Maryland Pilot Older Driver Study. DOT HS 809 583. Washington, DC.: U.S. Department of Transportation, National Highway Traffic Safety Administration. May.

Staplin, L., Gish, K., Lococo, K. H., Joyce, J.J., Stutts, and Kathy J. (2012). The Maze Test: A significant predictor of older driver risk. Accident Analysis and Prevention, In Press, Corrected Proof - available June 2012.

Staplin, L., Lococo, K. H., Gish, K. W., and Joyce, J. (2012). Functional Assessments, Safety Outcomes, and Driving Exposure Measures for Older Drivers. National Highway Traffic Safety Administration. DOT HS 811 630.

Stav, W.B., Arbesman, M., and Lieberman, D. (2008). Background and methodology of the Older Driver Evidence-Based Systematic Literature Review. *American Journal of Occupational Therapy* 62(2): 130-135.

Stav, W.B., Justiss, M.D., McCarthy, D.P., Mann, W.C., and Lanford, D.N. (2008). Predictability of clinical assessments for driving performance. *Journal of Safety Research* 39(1): 1-7.

Stelmach, G.E., and Nahom, A. (1992). Cognitive-motor abilities of the elderly driver. *Human Factors* 34(1): 53-65.

Stephens, B.W., McCarthy, D.P., Marsiske, M., Shechtman, O., Classen, S., Justiss, M., and Mann, W.C. (2005). International Older Driver Consensus Conference on Assessment, Remediation and Counseling for Transportation Alternatives: Summary and recommendations. *Physical and Occupational Therapy in Geriatrics* 23(2/3): 103-121.

Stinchcombe, A., Gagnon, S., Zhang, J.J., Montembeault, P., and Bedard, M. (2011). Fluctuating attentional demand in a simulated driving assessment: The roles of age and driving complexity. *Traffic Injury Prevention* 12(6): 576-587.

Storey, J.E. (2005). The development, pilot and validation of a multicultural cognitive assessment scale. Thesis- University of New South Wales.

Stovel, L. (2012). DriveABLE tests are unfair to rural seniors, MLA says. *Revelstoke Current*: April 29.

Stutts, J.C., and Wilkins, J.W. (2003). On-road driving evaluations: A potential tool for helping older adults drive safely longer. *Journal of Safety Research* 34(4): 431-439.

Stutts, J.C., and Wilkins, J.W. (2012). Keeping Seniors Safe and Mobile: An Evaluation of a Local Drive Test Option. Washington, DC.: AAA Foundation for Traffic Safety. April.

Stutts, J.C., Stewart, J.R., and Martell, C. (1998). Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention* 30(3): 337-346.

Stutts, J.C., Stewart, J.R., and Martell, C.M. (1996). Can screening for cognitive impairment at license renewal identify high risk older drivers? In: Proceedings of the 40th Annual Conference of the Association for the Advancement of Automotive Medicine, pp. 335-350. Des Plaines, IL.: Association for the Advancement of Automotive Medicine.

Sullivan, K.A., Smith, S.S., Horswill, M.S., and Lurie-Beck, J.K. (2011). Older adults' safety perceptions of driving situations: Towards a new driving self-regulation scale. *Accident Analysis and Prevention* 43(3): 1003-1009.

Sümer, N., Ayvasik, H.B., and Er, N. (2005). Cognitive and psychomotor correlates of self-reported driving skills and behavior. In: Driving Assessment 2005: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Rockport, Maine, June 27-30, 2005. Paper 14. Iowa City, IA.: University of Iowa Public Policy Center.

Szlyk, J.P., Myers, L., Zhang, Y., Wetzel, L., and Shapiro, R. (2002). Development and assessment of a neuropsychological battery to air in predicting driving performance. *Journal of Rehabilitation Research and Development* 39(4): 483-496.

Szlyk, J.P., Seiple, W., and Viana, M. (1995). Relative effects of age and compromised vision on driving performance. *Human Factors* 37(2): 430-436.

Szlyk, J.P., Severing, K., and Fishman, G.A. (1991). Peripheral Visual Field Loss and Driving Performance. Washington, DC.: AAA Foundation for Traffic Safety. August.

Tasca, L. (1992). Review of the Literature on Senior Driving Performance. Ontario: Ministry of Transportation, Road User Safety Office. November.

Taylor, B.D., and Tripodes, S. (2001). The effects of driving cessation on the elderly with dementia and their caregivers. *Accident Analysis and Prevention* 33(4): 519-528.

Temple, L.L. (1990). Drivers licencing competency and assessment domains: Cognition and Perception. Presented at the Conference on Driver Competency Assessment. Las Vegas, NV.: University of Nevada at Las Vegas.

Trilling, J.S. (2001). Selections from current literature assessment of older drivers. *Family Practice* 18(3): 339-342.

Tsui, T., Atsumi, B., Kanamori, H., and Miyao, M. (2002). Required performance of display devices in consideration of visual characteristics of elderly people: About luminance and character size. *Journal of the Society of Automotive Engineers Review* 23(2): 265-270.

Tuokko, H.A., McGee, P., Gabriel, G., and Rhodes, R.E. (2007). Perception, attitudes and beliefs, and openness to change: Implications for older driver education. *Accident Analysis and Prevention* 39(4): 812-817.

Ulleberg, P., and Sagberg, F. (2003). Syn og kognitiv funksjon blant bilførere over 70 år: Betydning for kjøreferdighet (Vision and Cognitive Functioning Among Drivers Aged 70 Years and Older: Importance for Driving Performance). TØI report 668. Oslo, Norway: Institute of Transport Economics. July.

Unsworth, C.A., Baker, A., Taitz, Chan, S., Pallant, J.F., Russell, K.J., and Odell, M. (2012). Development of a standardized Occupational Therapy - Driver Off-Road Assessment Battery to assess older and/or functionally impaired drivers. *Australian Occupational Therapy Journal* 59(1): 2012.

Unsworth, C.A., Lovell, R.K., Terrington, N.S., and Thomas, S.A. (2005). Review of tests contributing to the occupational therapy of off-road driver assessment. *Australian Occupational Therapy Journal* 52(1): 57-74.

Unsworth, C.A., Pallant, J.F., Russell, K., Odell, M. and Coulson, M. (2011). Interrater reliability of the Road Law and Road Craft Test as part of the OT-DORA Battery for off-road driver assesement. *British Journal of Occupational Therapy* 74: 394-98.

Unsworth, C.A., Pallant, J.F., Russell, K.J., Germano, C., and Odell, M. (2010). Validation of a test of road law and road craft knowledge with older of functionally impaired drivers. *American Journal of Occupational Therapy* 64(2): 306-315.

Van Elsland, P., and Dominique, F. (2000). Elderly Drivers: What Errors do they Commit on the Road. *Human Factors and Ergonomics Society Annual Meeting Proceedings* 44(20): 259-259.

Vann, A.S. (2012). Please tell people with Alzheimer's disease not to drive. *Journal of Geriatrics Society* 60(3): 597-8.

Vance, D., Dawson, J., Wadley, V., Edwards, J., Roenker, D., Rizzo, M., and Ball, K. (2007). The Accelerate Study: The longitudinal effect of speed of processing training on cognitive performance of older adults. *Rehabilitation Psychology* 52(1): 89-96.

Veling, I.H. (1980). Cognitive testing in driver licencing. In: Human Factors in Transport Research. Volume II. User Factors: Comfort, the Environment and Behaviour, eds. D.J. Oborne and J.A. Levis, pp. 278-285. London, England: Academic Press.

Vrkljan, B.H., McGrath, C.E., and Letts, L.J. (2010). Assessment tools for evalutating fitness to drive: A critical appraisal of evidence. *Canadian Journal of Occupational Therapy* 78(2): 80-96.

Wagner, J.T., Muri, R.M, Nef, T., and Mosimann, U.P. (2011). Cognition and driving in older persons. *Swiss Medical Weekly* 140:13136.

Wallace, R.B. (1997). Cognitive change, medical illness, and crash risk among older drivers: An epidemiological consideration. *Alzheimer Disease and Associated Disorders* 11(Suppl. 1): 31-37.

Wang, D.Y.D., and Entsminger, S. (2009). Age and attentional capacity. In: Driving Assessment 2009: International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design, Big Sky, Montana, June 22-25, 2009. Paper 59. Iowa City, IA: University of Iowa Public Policy Center.

Watanabe, T., Konagaya, Y., Yanagi, T. Miyao, M., Shibayama, H. 2012. Study of daily driving characteristics of individuals with dementia using video-recording driving recorders. *Journal of the American Geriatrics Society* 60(7): 1381-3.

Weihong Guo, A., Brake, J.F., Edwards, S.J., Blythe, P.T., and Fairchild, R.G. (2010). The application of in-vehicle systems for elderly drivers. *European Transport Research Review* 2(3): 165-174.

Wheatley, C.J. (2001). Visual perceptual aspects of driving. *Physical Disabilities Special Interest Sectional Quarterly* 24: 1-3.

Wheatley, C.J., and 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.

Whilihan, W.M., DiCarloa, M.A., and Paula, R.H. (2005). The relationship of neuropsychological functioning to driving competence in older persons with early cognitive decline. *Archives of Clinical Neuropsychology* 20(2): 217-228.

Wilde, G.J.S. (1988). Risk taking in psychomotor and cognitive tasks as a function of probability of loss, skill and other person-related variables. Paper presented at the 2nd International Conference on Road Safety, Groningen, Netherlands, August 1987. In: Road User Behaviour: Theory and Research, eds. J.A. Rothengatter and R.A. de Bruin, pp. 120-126. Assen/Maastricht, Netherlands: Van Gorcum.

Williams, C., and Graham, J.D. (1995). Licencing standards for elderly drivers. *Consumer's Research Magazine* 78(12): 18.

Willis, S.L., Tennstedt, S.L., Marsiske, M., Ball, K.K., Elias, J., Koepke, K.M., Morris, J.N., Rebok, G.W., Unverzagt, F.W., Stoddard, A.M., and Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of the American Medical Association* 296(23): 2805-2814.

Wilson, K.A. (2007). What to do about Elderly Drivers. *AutoWeek* 57(14): 20.

Winter, S.M., Classen, S., Bédard, M., Lutz, B.J., Velozo, C.A., Lanford, D.N., and Brumback, B.A. (2011). Focus group findings for the self-report Safe Driving Behaviour Measure. *Canadian Journal of Occupational Therapy* 78(2): 72-79.

Withaar, F.K., Brouwer, W.H., and Van Zomeren, A.H. (2000). Fitness to drive in older drivers with cognitive impairment. *Journal of the International Neuropsychological Society*, 6: 480-490.

Wong, Ides Y., Smith, Simon S., Sullivan, and Karen A. (2012). The relationship between cognitive ability, insight and self-regulatory behaviors: Findings from the older driver population. *Accident Analysis and Prevention* 49(2): 316-321.

Wood, J.M., and Carberry, T.P. (2004). Older drivers and cataracts: Measures of driving performance before and after cataract surgery. In: Highway Safety: Older Persons; Traffic Law Enforcement, Management and Trucking. Transportation Research Record 1865: 7-13.

Wood, J.M., and Mallon, K. (2001). Comparison of driving performance of young and old drivers (with and without visual impairment) measured during in-traffic conditions. *Optometry and Vision Science* 78(5): 343-349.

Wood, J.M., Horswill, M.S., Lacherez, P.F., and Anstey, K.J. (2013). Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. *Accident Analysis and Prevention* 50: 1161-1168.

APPENDIX C

Tools Excluded from the Meta-Analysis

Tools Excluded from the Meta-Analysis

ID #	Tool	Duration	Administration	Feasibility	Computer/ Equipment	Expertise	Reason for Exclusion
2, 64, 80	Useful Field of View (UFOV)	15 min	One on one (computer)	Low due to computer	Computer required	Software training	Requires training and a computer
8, 109	Assessment of Driving-Related Skills (ADReS)	15 min	One on one (administrator)	Low due to medical examination	Snellen chart, tape, stopwatch	Driver rehabilitation specialist	Requires one-on-one administration from trained expert
18, 111	P-Drive	40 – 60 min	One on one (simulator)	Low due to computer	Computer and simulator	Occupational Therapist	Lengthy test which must be administered by an occupational therapist using a simulator
24	Nordic Stroke Driver Screening Assessment (NorDSA)	30 min	One on one (administrator)	Low, battery of tests	None	Training in cognition	Lengthy battery of tests, administered individually by a trained specialist
37	Visual Selective Attention Test (VSAT)	6 min	One on one (verbal)	Low due to computer	Computer is required	Minimal	A computer is required for the administration of the test
39	Washington University Road Test (WURT)	20 – 40 min	One on one (road test)	Low due to road test	None, except vehicle	Trained driving tester	Lengthy on-road test that involves the use of a vehicle and requires a trained driving tester
54	Drive ABLE	30 – 60 min	One on one (computer)	Low due to computer	Computer is required	Certified evaluators	Lengthy test requires a computer, and individual administration by a certified evaluator
55, 56	Drive Aware	Unknown	One on one (administrator)	Low due to difficulty in scoring	None	Clinician	The test must be administered individually by a clinician
56	Drive Safe	20 min	One on one (computer)	Low due to computer	Computer, projector and screen	Clinician	The test requires several technological devices, as well as individual administration by a clinician
57	Visual Recognition Slide Test – University of Sydney (VRST-USyd)	20 min	One on one (computer)	Low due to computer	Computer or projector	Clinician	The test requires a computer, and the administrator must be a clinician

74	Gross Impairment Screening Battery (GRIMPS)	11 min	One on one (physical and cognitive)	Low due to one on one testing	None	Motor vehicle administration staff	The administration must be on an individual basis due to physical/movement aspects of the test
81	Safe Driving Behavior Measure (SDBM)	Over 30 min	One on one (administrator)	Low due to length	None	Driving evaluator	Lengthy test requires individual administration by a trained driving evaluator
21, 68, 82, 117F	Mini Mental State Examination (MMSE)	10 min	One on one (administrator)	Low due to verbal recall	None	Occupational Therapist	Test requires individual administration by an occupational therapist
95	Driving Health Inventory (DHI)	Unknown	One on one (computer)	Low due to computer	Computer is required	Occupational Therapist	Test requires a computer and must be administered one-on-one by an occupational therapist
97	Addenbrooke's Cognitive Examination Revised (ACE-R)	15 – 20 min	One on one (administrator)	Low due to recall	None	Healthcare professional	Test requires individual administration and must be administered by a healthcare professional
114, 115	Sensory-Motor and Cognitive Tests (SMC)	45 min	One on one (computer)	Low due to simulator	Driving simulator	None	Lengthy battery of tests requires the use of a simulator
75, 88	Screen for the Identification of Cognitively Impaired Medically At-Risk Drivers (SIMARD)	7 min	One on one (administrator)	Low due to recall	None	Clinician	Test requires individual administration (involves substantial verbal recall)
2A, 32A, 38B, 41A, 69B, 74A, 78C, 92, 106A, 108E, 108F, 113A	Trail Making Test A	5 min	One on one (administrator)	Low: each individual must be timed separately; administrator must watch for errors	None	None	Test requires some individual supervision, i.e., tests must be timed accurately in seconds, and the administrator must watch for errors and bring them to the participant's attention immediately

2B, 8A, 32B, 38A, 41B, 69C, 74B, 78D, 92A, 100, 108C, 108D, 110, 113B	Trail Making Test B	5 min	One on one (administrator)	Low: each individual must be timed separately; administrator must watch for errors	None	None	Test requires some individual supervision, i.e., tests must be timed accurately in seconds, and the administrator must watch for errors and bring them to the participant's attention immediately
92B	Color Trails Test 1	5 min	One on one (administrator)	Low: must time individually and the administrator must watch for errors	None	None	Test requires some individual supervision, i.e., tests must be timed accurately in seconds, and the administrator must watch for errors and bring them to the participant's attention immediately
92C	Color Trails Test 2	5 min	One on one (administrator)	Low: must time individually and the administrator must watch for errors	None	None	Test requires some individual supervision, i.e., tests must be timed accurately in seconds, and the administrator must watch for errors and bring them to the participant's attention immediately
116	Frostig	Under 10 min	One on one (administrator)	Low: admin must time and count errors individually	None	None	Test requires individual administration
24 a-f	Dot Cancellation Test	15 min	One on one	Low due to complexity/ supervision	None	Psychologist	Test is part of a larger battery of tests (Nordic Stroke), all of which must be administered in order to obtain a score. Battery must be administered individually by a psychologist

117B	Hazard Change Perception	17 min	One on one (computer)	Low due to computer	Computer is required	None	A computer is required for this test
117A	Hazard Change Detection	Unknown	One on one (computer)	Low due to computer	Computer is required	None	A computer is required for this test
117C	Snowy	3 min	One on one (administrator)	Low due to recall	None	Clinician	Test must be administered by a clinician and requires individual administration due to verbal recall
117D	Gestalt	2 min	One on one (administrator)	Low due to recall	None	Psychologist	Test must be administered by a psychologist and requires individual administration due to verbal recall
113C	American Association of Retired Persons (AARP) Reaction Time	10 min	One on one (administrator)	Low: each individual must be timed separately	None	None	Test cannot be given to a group: each test must be timed individually
103	Motor Free Visual Perception Test (MFVP)	25 min	One on one (administrator)	Low: examiner presents images individually	None	Occupational therapist, psychologist, optometrist	Lengthy test can only be administered by a specialist and requires individual administration
62A	Stroop Test	5 minutes	One on one (administrator)	Low due to recall	None	Rehabilitation Specialist	Test can only be administered by a specialist and requires individual administration
51	Behind-the-Wheel Assessment	60 min	One on one (on-road exam)	Low due to on-road exam	None	Driving Rehabilitation Specialist	Lengthy on-road test can only be administered by a driving rehabilitation specialist, requires individual administration, and the use of a vehicle