Published in final edited form as:

Saf Sci. 2012 February 1; 50(2): 333–341. doi:10.1016/j.ssci.2011.09.013.

COMPARING THE IMPAIRMENT PROFILES OF OLDER DRIVERS AND NON-DRIVERS: TOWARD THE DEVELOPMENT OF A FITNESS-TO-DRIVE MODEL

Jonathan F. Antin^a, Thurmon E. Lockhart^b, Laura M. Stanley^c, and Feng Guo^d

^aVirginia Tech Transportation Institute, 3500 Transportation Research Plaza (0536), Blacksburg, VA USA 24061, jantin@vtti.vt.edu

^bVirginia Tech, Industrial and Systems Engineering Department, 557 Whittemore Hall, Blacksburg, VA USA 24061, lockhart@vt.edu, Tel: + 1 540-231-9088; fax: +1 540-231-3322

^cMontana State University, Mechanical & Industrial Engineering Department, 220 Roberts, P.O. Box 173800, Bozeman, MT USA 59717-3800, laura.stanley@ie.montana.edu

^dVirginia Tech, Statistics Department, 415B Hutcheson Hall, Blacksburg, VA USA 24061, fguo@vtti.vt.edu

Abstract

The purpose of this research effort was to compare older driver and non-driver functional impairment profiles across some 60 assessment metrics in an initial effort to contribute to the development of fitness-to-drive assessment models. Of the metrics evaluated, 21 showed statistically significant differences, almost all favoring the drivers. Also, it was shown that a logistic regression model comprised of five of the assessment scores could completely and accurately separate the two groups. The results of this study imply that older drivers are far less functionally impaired than non-drivers of similar ages, and that a parsimonious model can accurately assign individuals to either group. With such models, any driver classified or *diagnosed* as a non-driver would be a strong candidate for further investigation and intervention.

Keywords

Functional Impairment; Senior Mobility; Driver Assessment; Driving Assessment

1. BACKGROUND

According to the Organization for Economic Co-operation and Development (OECD) report on Aging and Transport (OECD, 2001), by 2030 one out of every four drivers will be aged 65 or over. Furthermore, by 2050 the population of those over 80 years of age is expected to triple in most OECD member countries. The traffic safety impacts associated with this demographic trend have been debated for decades (TRB, 2005), but taken as a whole, drivers over 65 years of age have historically experienced higher rates of crashes on a per mile driven basis than their younger counterparts (NHTSA, 2000).

Correspondence to: Thurmon E. Lockhart.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

However, Janke (1991) found that drivers of all ages who travel shorter distances have a greater crash risk per unit of distance compared to those who drive greater distances. More recently, Hakamies-Blomqvist and colleagues (Hakamies-Blomqvist, Raitanen & O'Neill, 2002; Langford, Methorst & Hakamies-Blomqvist, 2006) have also pointed to a “low-mileage bias” which indicates that it is the older drivers with the lowest mileage driven who are primarily responsible for the increased crash rate observed for this age group, in alignment with Janke’s conclusion. The efficacy of the “low mileage bias” construct for explaining senior driver crash rate has been debated (see Langford, Koppel, McCarthy, & Srinivasanc, 2008 and Staplin, Gish, & Joyce, 2008).

A mitigating factor which may provide the underlying explanation for the low mileage bias is the notion of functional impairment. That is, for many low mileage seniors, it may be their degree of functional impairment which underlies both their lower mileage and their increased crash rate. An early study by Hakamies-Blomqvist (1998) suggested that the higher crash rate of older drivers is related to degree of functional impairment rather than to chronological age alone. It is well known that aging can lead to declines in perceptual, cognitive and psychomotor functions, but age alone cannot predict the degree of impairment (see Eby, Trombley, Molnar & Shope, 1998 for a review). Accurately assessing these impairments and relating them to crash risk has long been a goal of traffic safety professionals. Many have directed their efforts toward developing tests that can be used to screen drivers in order to identify those who are “at-risk”.

The assessment of older drivers can be framed as a signal detection problem as illustrated in Figure 1, with the distribution on the left representing those who have already ceased driving, and the distribution on the right representing those who are still driving. The X-axis represents functional ability – the further to the right on this conceptual scale, the more fit and capable the older driver. The Safe Driving Criterion is the conceptual defining point - those falling to its right are considered safe enough to drive today, whereas those drivers to its left (represented by the shaded region) are in need of further assessment and are our main focus in the current research effort. The problem is not currently a set of metrics which can definitively indicate on which side of the line any particular individual belongs, so the objective of this study is to further the effort to define such a set of metrics.

The ideal fitness to drive assessment must be accepted by the public, easy to administer, feasibly and reliably duplicated, suitable for both men and women, objectively scored, and, most importantly, predictive of real-world driving safety outcomes. Determining fitness to drive is important for all age groups, but it becomes most controversial for the oldest drivers. Should the testing be the same for all age groups? Should the testing increase in frequency after an age milestone has been reached? Because of the aging trend in many countries around the world (Fildes, 2008; U.S. Census Bureau, 2008), the need for an empirically validated fitness to drive assessment is ever growing in importance.

It must be recognized that fitness to drive is just one aspect of the larger set of senior mobility concerns faced by society. Devising a fitness-to-drive process that merely restricts or removes driving privileges from large numbers of older drivers would be easy, but neither prudent nor productive. First, doing such would merely serve to limit seniors’ mobility or simply transfer some portion the driving crash risk to other modes of travel (Langford, 2008; Whalen, Langford, Oxley, Koppel, & Charlton, 2006). Beck, Dellinger, and O’Neil (2007) compiled crash data over the period from 1999–2003. They found that the annualized fatality rate per trip was highest for those 65 and older for walking, cycling, and bus riding. Pucher and Dijkstra (2003) reported that based on 2001 data, in the U.S., cycling and walking entailed 23 and 12 times the risk of driving, respectively. They also compare the behavior of older individuals in the U.S. with those of Germany and the Netherlands. They

found that these European countries were safer for older pedestrians and cyclists by at least a factor of two compared with the U.S., which may be due to a wide variety of cultural and infrastructural differences between the U.S. and these European countries. This is especially salient when considering that the brunt of the injuries and fatalities associated with the generally increased crash risk observed in older drivers are primarily borne not by society at large, but by the older drivers themselves, as well as their typically older passengers (Braver & Trempe, 2004).

Second, driving is of great importance to many older individuals. In many developed countries, driving has long been the key to mobility, mobility is the key to independence, and independence is one of the keys to living a fulfilled life (Marshall, 2008; Oxley & Whelan, 2008). In a longitudinal study, Ragland, Satariano, & MacLeod (2005) interviewed 1,953 drivers, aged 55 and older, on factors related to depression and driving status, among other topics. It was discovered that former drivers reported higher levels of depression than the active drivers. Three years later, the researchers once again interviewed the participants and found that, even after controlling for health status and cognitive function, those who ceased driving during the three-year interval reported higher levels of depressive symptoms than those who remained active drivers.

Some older individuals may have sufficient access to transportation provided by family or friends, taxicabs, or senior mobility programs. For others, the private car may be the only truly accessible transportation mode due to cost, lack of transit options in rural areas, poor health, or personal mobility/locomotion limitations (Oxley & Whelan, 2008). That is, even if public transport is readily available, the rider must still be able to get to and from the stops. That may involve several blocks of walking with personal items in tow, steps to negotiate, etc. All of these may impose significant mobility barriers to older individuals, while driving a private vehicle may impose few or none.

Therefore, the goal of determining fitness to drive is not first and foremost to remove drivers from the roads. Rather, it is to identify those drivers who may benefit from any of a variety of interventions intended to help keep them driving longer and more safely (Langford, 2008). While much work has been done in this area, several that have demonstrated some success are described briefly below.

Staplin, Lococo, Gish, and Decina (2003) developed and evaluated a model driver screening and evaluation program. Their gross impairment screening (GRIMPS) included physical as well as perceptual-cognitive measures. Four aspects of perceptual-cognitive ability emerged as being predictive of real world safety-related outcomes. These were directed visual search, information processing speed for divided attention tasks, the ability to visualize missing information in an image, and working memory. Two physical measurements also emerged as important. These were lower limb strength and head/neck rotation.

Hennessy and Janke (2005) described the development and validation of an integrated three-tier system for assessing driving wellness (freedom from driving-relevant functional limitations) and driving fitness (the degree to which a driver compensates for any such limitations while driving). Tier 1 is a quick screen comprised of visual tests of acuity and contrast sensitivity, a scan for obvious physical limitations, and a brief memory test. Tier 2 includes the standard California DMV knowledge test and a test of perceptual response time. Tier 3 is a road test. Each successive tier is invoked when the lower tier is failed. They found the counterintuitive result that of those had crashed at least once in the 3 years prior to the study, the percentage of those classified as *somewhat* functionally limited was more than twice as high as those who were *extremely* functionally limited. Hennessy and Janke (2005) concluded that this was due to the extra degree of compensation employed by the more

severely impaired group. However, it may simply have been a function of the more extremely limited group driving less than their more capable counterparts, thus experiencing less exposure.

Ball et al. (2006) also sought to determine if older driver fitness to drive could be determined in a DMV setting. These researchers administered the GRIMPS Battery to participants and compared the results prospectively with crashes. They found that the two best predictors of future crashes were the UFOV® Subtest 2 and the Motor-Free Visual Perception Visual Closure subtest, but the odds ratio for each was less than 1.25.

2. OBJECTIVE

The objective of this study was to compare the multi-dimensional functional abilities of a group of older drivers with that of a group of older individuals who had recently ceased driving in an effort to point to metrics which may provide a basis for a suitable fitness-to-drive assessment protocol. It was hypothesized that these two groups could be distinguished based on their respective functional abilities. Eby, Trombley, Molnar, and Shope (1998) noted that the most important high level abilities for driving include visual perception, cognition, psychomotor ability. The current study was designed to assess both the driving and non-driving participants on each of these dimensions as well as a variety of physical and psychological dimensions.

3. METHODS

3.1 Participants

Participants were 49 older individuals living in the vicinity of the New River Valley area of Virginia. Recruitment was accomplished using newspaper ads, posted flyers, word of mouth, and personal appearances at senior gathering locations. The minimum age cutoff was 65, but an attempt was made to recruit as many as possible who were 75 and older.

The driving group was recruited to participate in a naturalistic driving study, and so this group was comprised of those who had valid driver licenses and reported driving regularly (i.e., at least a few times a week). In several cases, husband and wife pairs were both recruited, as long as they each met the criteria for participation. The non-drivers reported having ceased driving approximately within the past two years. Drivers who had ceased driving prior to that were excluded because it was felt that these individuals may be too different from current drivers to form a meaningful or fair comparison. The non-drivers in this study reported ceasing driving due to fairly typical reasons; almost 45% indicated that they “just felt it was time.” The most prominent specific reasons included: physical problems, vision problems, and family member recommendation (> 30% each). Note participants could select as many reasons as were applicable.

An attempt was made to recruit an equal number of men and women in each group. Participant age information is summarized in Table 1. A t-test showed no significant difference between the driving and non-driving groups in terms of age ($p = 0.13$), leading to the conclusion that the two groups are comparable in terms of that dimension.

All participants were required to be able to walk at least a short distance during the study, though regular use of a cane, walker, or wheelchair did not preclude participation. Participants also had to be judged competent to provide informed consent, and a few of the non-driving candidates were politely dismissed when it became apparent that they could not meet this standard. For example, if the participant could not remember where he or she was, then that participant would have been determined to not be competent to provide consent.

Transportation was provided for all participants as required, and participants were compensated at a rate of \$20/hour.

It is important to note that for 20 of the driving participants, the current study served as the beginning of their participation in a year-long naturalistic driving study (NDS). In the NDS, the older drivers' vehicles have been instrumented with a continuous data acquisition system, including four video cameras and a variety of sensors, which continuously record their driving behaviors and safety outcomes for one year each. During the informed consent process, all driving participants were assured that their data would not be used against them legally or otherwise, and a Certificate of Confidentiality was secured from NIH to protect their data from subpoena if necessary.

3.2 Procedures

Participants' functional abilities were assessed during two sessions on two different days.

First Session—The first session was conducted at the Virginia Tech Transportation Institute and took approximately 1.5 to 2 hours to complete – two drivers took only an hour, and two non-drivers took 2.5 hours. Participants first signed the informed consent form. They then completed the following questionnaires and assessments:

1. Abbreviated Mental Test Score
2. Mental Sharpness Comparison – Participants were asked to “compare your mental sharpness now with how you were in your 40s or 50s. Would you say that you are... (circle a number below): “ on a scale from 1 to 7. Verbal anchors are shown below.
 - 1 - A lot less mentally sharp now
 - 4 - Somewhat less mentally sharp now
 - 7 - Just as mentally sharp now”
3. Snellen Acuity
4. Contrast Sensitivity Chart (L, R)
5. Color Perception Tests
6. Optec 2000 Tests (stereopsis, far acuity, vertical and lateral phorias)
7. Faces Pain Scale
8. Health Assessment Questionnaire
9. WHO (Five) Well-Being Index (1998)
10. Sleep Hygiene Questionnaire
11. From the DrivingHealth Inventory®
 1. Visualizing Missing Information
 2. Useful Field of View®
 3. Trail Making (Parts A and B)

Second Session—The second session included Locomotor Research Laboratory Activities and Visual Controls Laboratory Activities which were conducted within Virginia Tech's Industrial and Systems Engineering (ISE) Department. This session took one to three hours to complete – typically an hour for the drivers and two to three for the non-drivers

because they were generally slower to understand and perform the assessments. Because more walking was required to reach the two ISE laboratories, use of a wheelchair was offered to all participants during this session.

3.3 Locomotion Research Laboratory Activities

Strength was assessed as rotational torque generated around the hip and ankle joints, measured in Newton-meters (N-m) using a Biodex System 3 Dynamometer. A gross metric of upper body strength was similarly assessed. In addition, the time taken to generate the torque based on the presentation of a randomly timed visual stimulus was also recorded as reaction time. During every trial, the participant received each directional stimulus (i.e., extension/flexion or left/right) 3–4 times. Participants were allowed at least a 60-second rest period between trials. Maximum torque was recorded as the average of the peak torques for each type of exertion during a trial. Initial reaction times (IRT) are reported as the time from stimulus onset to the beginning of the exertion of torque, and are reported as the average of all IRTs for a given trial. Peak reaction time is the time from stimulus onset to the peak torque generated and was similarly averaged. Ankle and hip torque were both measured isometrically. To measure upper body torque and reaction time, the dynamometer was customized with an actual steering wheel and the participant was asked to rotate the wheel as fast and hard as possible in the direction indicated by a visual stimulus (i.e., left or right). Upper body torque and reaction time measurements were taken isokinetically at 150 degrees/second. LabVIEW® software was used to interface with the Biodex system and record reaction times. Raw data were smoothed using a 4th order, 40-sidepoint Savitzky-Golay filter and a 4th order, zero-lag Butterworth filter with a 6 Hz cutoff frequency.

To measure head-neck-torso flexibility, participants retained the same driving posture as was used for the upper body torque measurement described above, resembling the restraint system in a car. Participants were asked to grip the steering wheel over markings at approximately at 10 and 2 o'clock. To begin, a block with four directional arrows was positioned on a sliding arm to the far left of the participant. It was slowly moved to the right as far as the participant could reliably identify the direction of an arrow. The head-neck-torso flexibility was measured as the angle from the mid-sagittal plane to the target's location at the limit of the participant's ability to see and correctly identify the block's arrow direction.

3.4 Visual Controls Laboratory Activities

The participant was first dark adapted for 5 minutes before beginning the assessments in the Visual Controls Laboratory. To measure dynamic visual acuity (DVA) - A series of Landolt Rings was moved across the observer's field of view in a single horizontal direction at rotation velocities of 12, 24 and 36 degrees/second using a computer, projector, motor, and mirror apparatus. The participant was asked to identify the location of the gap in the ring (Up, Down, Left, or Right). The size of each successive ring was reduced during the trial, and values were observed until the participant could no longer reliably identify the direction of the gap. The largest target size corresponded to 20/200 acuity and the smallest size corresponding to 20/20 acuity.

Sensitivity to glare was measured using a mesopic contrast sensitivity chart placed 3 m from the participant. A halogen glare source seen through a neutral-density filter was also placed 3 m from the participant, 56 cm off the central line of sight. In addition, the participant was also asked to rate his or her discomfort due to the glare source using the DeBoer Discomfort Glare Rating Scale (de Boer & Schreuder, 1967).

4. RESULTS

The first phase of the analysis was to compare the drivers to the non-drivers on each variable individually. The main objective of this analysis is to investigate how the two groups differ across the various dimensions sampled with the assessments. A standard t-test was used and significant level was set to $\alpha = 0.05$. Altogether, 60 variables were tested (see Appendix A for a complete listing), and the ones that were statistically significant are shown in Table 2 through Table 5. To address the multiple comparison issue, possibly resulting in some false positive results, we also calculated the p value based on False Discovery Rate (FDR, Benjamini and Hochberg, 1995). The FDR is an alternative way of quantifying Type I errors in multiple comparison problems. It is especially useful for large number of comparison for exploratory purpose, which fits exactly with the context in which these comparisons are being conducted. Note seven variables identified as significant by t-test were not significant based on FDR p -values.

The results shown in Table 2 indicate significant differences between drivers and non-drivers in the physical dimension in terms of lower extremity and upper body strength.

All significant reaction time differences favored the drivers (Table 3). If these differences are translated into distance traveled at real-world speeds, then the potential impact of these seemingly subtle differences can be substantive. For example, the shortest of the significant reaction time differences between the two groups (0.404 s for the Upper Body Initial Reaction Time variable) translates to more than 20 feet of travel at 35 mph. These reaction time and derived differential distance traveled metrics lend support to the non-drivers' decision to cease driving.

The key findings illustrated in Table 4 are the differences seen in static visual acuity as well as all three rotational velocities of dynamic visual acuity, all favoring the drivers. Table 4 also reveals some apparently anomalous findings. First, glare contrast sensitivity is the lone statistically significant variable where the non-drivers demonstrated superior performance to that of the drivers. Second, notice that there are statistically significant contrast sensitivity differences between the two groups at three spatial frequencies for the left eye, but no similar differences were found for the right eye. We currently have no explanations for these interesting but unexpected results.

As seen in Table 5, the non-drivers had more than twice the number of health problems of the drivers (without respect to severity). Since both groups had driven, on average, well over 50 years, it is not obvious that years of driving represents a practically significant difference, especially since we know that the non-drivers had recently given up driving.

The Trail Making Part B scores are revealing. Giovagnoli et al. (1996) tested normal, healthy individuals in different age and education categories to develop norms for the Trail Making Part B test. Age groups in their study included individuals from below 65 up to 89. Education levels were separated into two groups: those with education up to 12 years, and those with greater levels of education. Each participant was assigned a percentile value relative to his or her age and education category as calculated by Giovagnoli et al. (1996). The one participant outside of this age range, a 93 year old, was grouped with the 85–89 group for comparison purposes. The average Trail Making Part B score for the drivers in our study was 45th percentile compared with an average 20th percentile for the non-drivers.

4.1 Model Development

The next and primary phase of the analysis was to develop a parsimonious logistic regression model to predict participant membership in driver or non-driver groups based on

performance on the assessments. The results of the comparisons discussed above were one of the key variable selection criteria. In addition, a correlation matrix was generated across all variables to avoid placing any two factors in the model that were highly correlated. The idea is to use the information inherent in the drive/no-longer drive decisions of our participants to develop logistic regression models to assign other individuals to either the driving or non-driving group. In particular, if an older driver were to undergo routine assessment and a validated model were to assign that individual, based on these objective assessments, to the non-driver group, then that individual would be a good candidate for further, more in-depth evaluation.

Variable selection—The comparisons, correlation matrix, and driver safety perspectives were all used to select the variables for inclusion in the model. Table 6 shows the correlations among the initially selected variables.

- Upper Body Maximum Torque (left)
- Upper Body Maximum Torque (right)
- Upper Body Initial Reaction Time
- Dynamic Visual Acuity (24 degrees/sec)
- Snellen Acuity
- Trail Making Test (Part B)

Logistic Regression—A cross validation approach was used to assess the validity of the model. The method assigns individuals to either the driving or non-driving groups based on the cross-validated individual predicted probability, which is a function of the participant assessment scores (covariates). Cross validation was accomplished by excluding each individual's data from the model as the model was developed and used to predict that individual's group membership.

Upper body maximum torque left and right were highly correlated, as might be expected, so these two variables were averaged to derive *Average Upper Body Maximum Torque*. With the remaining five variables, all 31 eligible participants were correctly categorized by the model as being either drivers or non-drivers. Participants were eligible if they had produced valid data for each of the five variables, as a complete data analysis approach was used. Not all participants were able to produce usable data on all assessments due to each individual's particular abilities or limitations. While the remaining data from those who were excluded would theoretically be interesting, their missing values would have to have been imputed which was felt to be inappropriate for the current modeling effort.

This implies that those five variables provide sufficient information to assign all eligible participants in our sample to driver or non-driver groups with 100% accuracy. Ironically, because of this total separation, there are no unique solutions for the model parameters.

Another even more parsimonious model was sought that would still predict group membership with high accuracy. Thus, a model was fitted using the following three variables: dynamic visual acuity (24 degrees/sec), average upper body maximum torque, and Trail Making (Part B). Those three variables represent three distinct dimensions related to driving: perception, physical ability, and visual-cognitive ability. Only three of 31 participants were misclassified, which implies majority of the variance in the models can be explained by these three variables alone, even though none of the three individual model parameters achieved statistical significance, likely due to the relatively small sample size. Model parameters are shown in Table 7.

5. DISCUSSION AND FUTURE WORK

The results of this study show a clear difference between the driving and non-driving groups in terms of their relative degree of functional impairment. Of the more than 60 metrics evaluated, 21 showed statistically significant differences across several dimensions as summarized in Tables 2–5. All but one of those (glare contrast sensitivity) favored the drivers, as was expected.

The ankle plantarflexion and dorsiflexion torque values in drivers are comparable to values seen in various studies which have examined ankle strength in older adults using similar protocols (Chandler, Duncan, Kochersberger, & Studenski, 1998; Kim, Lockhart, & Nam, 2010). However, the reported ankle torque values in this study for both drivers and non-drivers are lower than those reported in some of the existing literature for older adults (Simoneau, Martin, & Van Hoecke, 2005; Rubenstein et al., 2000). These discrepancies may be due to the different postures used to collect the isometric ankle strength data. For the purpose of current study, the participants were seated in a position similar to driving which may have limited their capacity to generate maximum strength at the ankle.

An important goal for those concerned with elder mobility is to find an easy-to-administer set of assessments that will meaningfully indicate whether an individual is fit to drive. This paper represents an initial attempt to develop a model that will fit those important criteria by using driving-related assessments of individuals who had already preselected themselves into driving and non-driving groups. This effort was successful in producing a five-variable model that accurately classified all eligible participants and a three-variable model that misclassified only three of 31 individuals. Both of these models include variables that sample physical, perceptual, and cognitive abilities which are all believed to be key elements to safe driving.

Research has been conducted which supports these findings. The aging individual tends to exhibit musculoskeletal degradation. This degradation has been shown to include both a decreased ability to generate muscular force as well as longer muscular latency times (Tang & Woollacott, 1998; Chambers and Cham, 2007; Lockhart & Kim, 2006). Previous studies have indicated that age-based vision loss can explain some proportion of increased crash rate for this group (Ball, Owsley, Sloane, Roenker, & Bruni, 1993; Shinar and Schieber, 1991). Among the pool of the tested visual characteristics, static visual acuity is most often tested during licensing, yet it may not be the most salient visual capacity in terms of driving and avoiding crashes. A measure of dynamic visual acuity may be more beneficial when dealing with activities with high dynamic visual acuity demands, for example, reading a highway-sign while driving (Banks, Moore, Liu, & Wu, 2004). Stutts, Stewart, and Martell (1998) investigated the use of brief tests to examine whether impaired cognitive function is associated with increased crash risk for older drivers. Data for the study were collected from 3,238 drivers aged 65 and older applying for renewal of their North Carolina driver's license. The specific cognitive assessments examined include the Trail Making Test parts A and B and others. They found that cognitive test performance in general remained significantly associated with crash risk even after controlling for driver age, race and measures of driving exposure.

If a model can be found that is a good predictor of driver and non-driver group membership, then those who the model seemingly misclassifies may become the most interesting cases. Specifically, those drivers who are classified or *diagnosed* as non-drivers may become candidates for a wide variety of possible interventions including driver or cognitive training, physical therapy, medication adjustments, aftermarket vehicle add-ons, or driving restriction.

Self-imposed driving restrictions represent a logical step between unfettered driving and complete driving cessation, whereby driving habits are altered to avoid situations of perceived stress, difficulty, or danger. Typically, such restrictions might include avoiding or minimizing night driving, highway driving, or complex intersections. A study by Hakamies-Blomqvist (1994) demonstrated that such compensations can be successful at reducing crash risk for senior drivers.

Dellinger, Sehgal, Sleet, & Barrett-Connor (2001) found the counterintuitive result that a group of those who had ceased driving suffered from significantly *fewer* medical conditions than a group who were still driving. This seemingly anomalous finding does not coincide with the findings in the current study, but does point to the fact that a driver's decision to restrict or cease driving altogether may be based on a complex set of factors related to medical conditions, functional abilities, degree of dementia, self confidence, and input from others (e.g., family and physicians). Additionally, to the extent that such compensation leads to increased use of other modes of transportation (e.g., walking, cycling, public transit), overall societal crash risk may be reduced, but the overall risk for that individual may actually increase.

The situation may be further exacerbated by a lack of insight into one's own abilities. In a study conducted by Freund, Gravenstein, Ferris, Burke, & Shaheen (2005), older individuals rated their own driving abilities compared to that of their peers. Each participant's driving ability was then assessed in a STISIM Drive™ driving simulation system. Sixty-five percent of the participants reported a self-rating of "better than average" with regard to their driving abilities as compared to their peers. However, those who rated themselves as better drivers were four times more likely to be rated as an unsafe driver based on their performance in the simulator. Thus, a large portion of the older driver population may be overestimating their own driving capabilities based on an inability to effectively judge safety-related declines in their own functional capabilities. In the current study, participants were asked to rate how mentally sharp they were now compared with how they were in their 40s or 50s. A metric was derived from the ratio of their response to this question to their score on the Abbreviated Mental Test Score. The higher the score, the greater was the disparity between the participant's self image and reality. While there was no reliable difference between the drivers and non-drivers in terms of this metric, it will be interesting to compare the drivers' scores on this metric to their actual driving performance demonstrated in the NDS. This type of uncertainty regarding self-awareness further supports the need for definitive assessment.

Dynamic visual acuity (especially at 24 degrees/second) seems to be among the most promising of all the measures taken in terms of separating drivers from non-drivers. Interestingly, the correlations between Snellen Acuity and the three measures of dynamic acuity at 12, 24, and 36 degrees/second were surprisingly low: $r(12, 24, 36) = 0.14824, 0.44879, 0.27923$, respectively. This strongly implies that metrics of static acuity are not suitable surrogates for a dynamic acuity measure. It remains to be determined if the main source of the difference between the two groups on this metric is based upon the ability to track the moving targets with corresponding head movements, better anticipation, visual encoding, or cognitive processing. In either case, a digital presentation of the dynamic acuity test would likely come much closer to the criteria of ease of administration and mass dissemination.

Ultimately, for any model to truly fulfill the goal of meaningfully determining fitness to drive, it must not only be able to differentiate drivers from non-drivers, it must also be able to reliably predict relevant driving behaviors, performance, and safety outcomes. Future research efforts will seek to use the older drivers' naturalistic driving data to generate and validate such models.

5.1 Limitations

This study was conducted with a relatively small number of participants. For the models developed to have lasting value, they must be tested, refined, and validated on a much larger, more broadly representative sample. For instance, the driving participants in the current study knew they were being recruited to participate in a naturalistic driving study, so these individuals were conceivably in better health than they typical driver or more interested in research and technology. In addition, all participants were asked to use a mouse to perform the computerized tests of visual-cognitive ability. However, a few of the non-drivers had little or no experience using a computer, and so could not effectively use the mouse. For these few individuals, an experimenter positioned it for them, as guided by the participants. This likely served to mute the differences between the drivers and non-drivers on these metrics. In the future, a touch screen will be employed for these types of tests to minimize such problems.

This study was designed to exclude individuals who were deemed incapable of providing informed consent. While none of the driving participants were excluded for this reason, the results of this study cannot be extrapolated to individuals with any substantive degree of cognitive impairment.

Our effort to measure the ability of participants to dynamically accommodate between near and far targets was unsuccessful. The participants simply could not generate stable accommodation metrics using the autorefractor, although the apparatus routinely generates stable results with younger participants. As a driver must monitor objects at different distances while driving (including items on the dashboard as well as outside the vehicle), the failure to accommodate properly, due to either insufficient accommodative power or excessive accommodation time, may impair focus and thus affect driving performance or emergency cue detection. If the older individuals in this study were unable to perform a laboratory-based accommodation task well enough to be measured, then it is reasonable to expect that they may have a similar or related decrement when performing the key accommodation activities required for safe driving. Further study of this ability (or impairment) is warranted as a possible source of increased risk with older drivers.

While the models generated in the current study were successful at separating the two groups, the ease-of-administration criterion has not yet been achieved. The Trail Making Tests can be quickly and easily administered, and accurately and objectively scored. However, in this study upper body strength and dynamic visual acuity were measured with customized apparatus that would at best be difficult and expensive to reliably duplicate and disseminate. In the future, upper body strength may be approximated by grip strength, a far easier metric to attain. However, its ability to predict group membership and driving outcomes would both need to be verified.

6. CONCLUSIONS

The results of this study imply that older drivers are very different from their non-driving counterparts in terms of degree of functional impairment, with the drivers demonstrating consistently less impairment in almost all instances where a statistically significant difference was observed. Also, it was shown that a logistic regression model can be derived which completely and accurately separates these two groups. This model was comprised of scores on five assessments of functional ability which were related to upper body strength and reaction time, static and dynamic visual acuity, and a metric of visual-cognitive ability. An even more parsimonious three variable model including only dynamic visual acuity, upper body strength, and a metric of visual-cognitive ability was almost as good at distinguishing these groups. This is a promising contribution to the effort to find a way to

assess fitness to drive which can be scored accurately and objectively. With such a model, any driver *diagnosed* as a non-driver would be a strong candidate for further investigation and intervention. Future research efforts will be directed towards validating the current model with larger sample sizes and against naturalistically observed driving safety outcomes. Finding suitable surrogate assessments to improve ease of administration and dissemination is another important goal.

Acknowledgments

The following individuals contributed to the development of the research protocols and the data collection conducted at Virginia Tech's Industrial and Systems Engineering Department: Wen Shi, Prakriti Parijat, Courtney Haynes, and Emily Miller. The following individuals also helped to collect data and manage logistical issues at the Virginia Tech Transportation Institute: Kim Shelton, Kristen Butler, Christine Link-Owens, Kathleen Smith, and Brian Wotring. Last, but certainly not least, we would like to acknowledge the efforts of all of our older volunteers. This study required a non-trivial amount of time and effort, and their participation and patience are greatly appreciated.

ROLE OF THE FUNDING SOURCE

This study was sponsored by the National Surface Transportation Safety Center for Excellence, and helpful reviews of the goals and protocols were provided by its Stakeholder Committee.

References

- Ball KK, Owsley C, Sloane ME, Roenker DL, Bruni JR. Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology and Visual Science*. 1993; 34:3110–23. [PubMed: 8407219]
- Ball KK, Roenker DL, Wadley VG, Edwards JD, Roth DL, McGwin G, et al. Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? *Journal of the American Geriatrics Society*. 2006; 54(1):77 – 84. [PubMed: 16420201]
- Banks PM, Moore LA, Liu C, Wu B. Dynamic visual acuity: a review. *The South African Optometrist*. 2004; 63(2):58–64.
- Beck LF, Dellinger AM, O'Neil ME. Motor vehicle crash injury rates by mode of travel, United States: Using exposure-based methods to quantify differences. *American Journal of Epidemiology*. 2007; 166(2):212–18. [PubMed: 17449891]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*. 1995; 57(1): 289–300.
- Braver ER, Trempe RE. Are older drivers actually at higher risk of involvement in collisions resulting in deaths or non-fatal injuries among their passengers and other road users? *Injury Prevention*. 2004; 10(1):27–32. [PubMed: 14760023]
- Chambers A, Cham R. Slip-related muscle activation patterns in the stance leg during walking. *Gait & Posture*. 2007; 25(4):565–72. [PubMed: 16876417]
- Chandler JM, Duncan PW, Kochersberger G, Studenski S. Is lower extremity strength gain associated with improvement in physical performance and disability in frail, community-dwelling elders? *Archives of Physical Medicine and Rehabilitation*. 1998; 79(1):24–30. [PubMed: 9440412]
- de Boer JB, Schreuder DA. Glare as a criterion for quality in street lighting, *Transactions of the Illuminating Engineering Society*. 1967; 32(2):117–35.
- Dellinger AM, Sehgal M, Sleet DA, Barrett-Connor E. Driving cessation: what older former drivers tell us. *Journal of the American Geriatrics Society*. 2001; 49(4):431 – 35. [PubMed: 11347787]
- Eby, DW.; Trombley, DA.; Molnar, LJ.; Shope, JT. The assessment of older drivers' capabilities: a review of the literature. (No. UMTRI-98-24). University of Michigan Transportation Research Institute; Ann Arbor, MI: 1998.
- Fildes B. Future directions for older driver research. *Traffic Injury Prevention*. 2008; 9(4):387–93. [PubMed: 18696396]

- Freund B, Gravenstein S, Ferris R, Burke BL, Shaheen E. Drawing clocks and driving cars use of brief tests of cognition to screen driving competency in older adults. *Journal of General Internal Medicine*. 2005; 20(3):240–4. [PubMed: 15836527]
- Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacina M, Capitani E. Trail making test: normative values from 287 normal adult controls. *Italian Journal of Neurological Science*. 1996; 17:305–9.
- Hakamies-Blomqvist L. Compensation in older drivers as reflected in their fatal accidents. *Accident Analysis & Prevention*. 1994; 26(1):107–12. [PubMed: 8110349]
- Hakamies-Blomqvist L. Older drivers' accident risk: conceptual and methodological issues. *Accident Analysis & Prevention*. 1998; 30(3):293–7. [PubMed: 9663287]
- Hakamies-Blomqvist L, Raitanen T, O'Neill D. Driver aging does not cause higher accident rates per km. *Transportation Research Part F*. 2002; 5:271–4.
- Hennessy, DF.; Janke, MK. California Department of Motor Vehicles Licensing Operations Division Summary Report No. RSS-05–215. 2005. Clearing a road to driving fitness by better assessing driving wellness development of California's prospective three-tier driving-centered assessment system.
- Janke MK. Accidents, mileage and the exaggeration of risk. *Accident Analysis and Prevention*. 1991; 23(2–3):183–8. [PubMed: 2029319]
- Kim S, Lockhart TE, Nam CS. Leg strength comparison among three different age groups; Are middle-age workers stronger than older workers? *International Journal of Industrial Ergonomics*. 2010; 40(3):315–20. [PubMed: 20436934]
- Langford J. Usefulness of off-road screening tests to licensing authorities when assessing older driver fitness to drive. *Traffic Injury Prevention*. 2008; 9(4):328–335. [PubMed: 18696389]
- Langford J, Koppel S, McCarthy D, Srinivasan S. In defence of the 'low-mileage bias'. *Accident Analysis and Prevention*. 2008; 40(6):1996–9. [PubMed: 19068306]
- Langford J, Methorst R, Hakamies-Blomqvist L. Older drivers do not have a high crash risk—A replication of the low mileage bias. *Accident Analysis and Prevention*. 2006; 38(3):574–8. [PubMed: 16426560]
- Lockhart TE, Kim S. Relationship between hamstring activation rate and heel contact velocity: Factors influencing age-related slip-induced falls. *Gait & Posture*. 2006; 24(1):23–34. [PubMed: 16112575]
- Marshall S. The role of reduced fitness to drive due to medical impairments in explaining crashes involving older drivers. *Traffic Injury Prevention*. 2008; 9(4):291–8. [PubMed: 18696384]
- National Highway Traffic Safety Administration (NHTSA). Report No. DOT HS 808 091. Washington, DC: U.S. Department of Transportation; 2000. Traffic Safety Facts 1999: Older Population.
- Organization for Economic Co-operation and Development (OECD). Report published by an OECD Scientific Expert Group. Paris: 2001. Ageing and transport: mobility needs and safety issues.
- Oxley J, Whelan M. It cannot be all about safety: the benefits of prolonged mobility. *Traffic Injury Prevention*. 2008; 9(4):367–78. [PubMed: 18696394]
- Pucher J, Dijkstra L. Promoting safe walking and cycling to improve public health: lessons from the Netherlands and Germany. *American Journal of Public Health*. 2003; 93(1):1509–16. [PubMed: 12948971]
- Ragland DR, Satariano WA, MacLeod KE. Driving cessation and increased depressive symptoms. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2005; 60(3):399–403.
- Rubenstein LZ, Josephson KR, Trueblood PR, Loy S, Harker JO, Pietruszka FM, et al. Effects of a group exercise program on strength, mobility, and falls among fall-prone elderly men. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2000; 55(6):M317–21.
- Shinar D, Schieber F. Visual requirements for safety and mobility of older drivers. *Human Factors*. 1991; 33(5):507–19. [PubMed: 1769671]

- Simoneau E, Martin A, Van Hoecke J. Muscular performances at the ankle joint in young and elderly men. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2005; 60(4):439–47.
- Staplin, L.; Lococo, K.; Gish, K.; Decina, L. Report No. DOT HS 809 58. 2003. Model Driver Screening and Evaluation Program Final Technical Report Volume II: Maryland Pilot Older Driver Study.
- Staplin L, Gish KW, Joyce J. 'Low mileage bias' and related policy implications—a cautionary note. *Accident Analysis and Prevention*. 2008; 40(3):1249–1252. [PubMed: 18460395]
- Stutts JC, Stewart JR, Martell C. Cognitive test performance and crash risk in an older driver population. *Accident Analysis and Prevention*. 1998; 30(3):337–46. [PubMed: 9663293]
- Tang PF, Woollacott MH. Inefficient postural responses to unexpected slips during walking in older adults. *Journal of Gerontology: Medical Sciences*. 1998; 53 (6):M471–80.
- Transportation Research Board (TRB). *Proceedings of Transportation in an Aging Society: A Decade of Experience*. Washington, DC: National Academy of Sciences; 2005.
- U.S. Census Bureau. *National Population Projections Released 2008 (Based on Census 2000)*. 2008. <http://www.census.gov/population/www/projections/summarytables.html>
- Whalen, M.; Langford, J.; Oxley, J.; Koppel, S.; Charlton, J. MUARC Report No. 255. 2006. *The Elderly and Mobility: a Review of the Literature*.

Biographies

Jonathan F. Antin earned his Ph.D. in Industrial Engineering and Operations Research (Human Factors Option) from Virginia Tech in 1987. He is currently a Human Factors Research Scientist and Light Vehicle Safety Group Leader in the Center for Automotive Safety Research at the Virginia Tech Transportation Institute.

Thurmon E. Lockhart earned his Ph.D. in Industrial and Systems Engineering from Texas Tech in 2000. He is currently an Associate Professor of Industrial and Systems Engineering at Virginia Tech.

Laura M. Stanley earned her Ph.D. in Industrial Engineering from Montana State University-Bozeman in 2006. She is currently an Assistant Professor of Mechanical & Industrial Engineering at Montana State University-Bozeman.

Feng Guo earned his Ph.D. in Statistics from the University of Connecticut in 2007. He is currently an Assistant Professor of Statistics department at Virginia Tech and at the Virginia Tech Transportation Institute.

Appendix A. Sixty Dependent Measures (Numbers in parentheses indicate total number of dependent measures included in that group or line item)

Physical Ability (13)

1. Ankle Torque Max/Plantar & Dorsiflexion (2)
2. Ankle Initial Reaction Time (RT, mean of Plantar & Dorsiflexion) (1)
3. Ankle Peak RT (mean of Plantar & Dorsiflexion) (1)
4. Hip Torque Max/Flex & Extend (2)
5. Hip Initial RT (mean of Flex & Extend) (1)
6. Hip Peak RT (mean of Flex & Extend) (1)
7. Upper Body Torque Max/Left & Right (2)

8. Upper Body Initial RT (mean of Left & Right) (1)
9. Upper Body Peak RT (mean of Left & Right) (1)
10. Head-Neck-Torso Flexibility (1)

Visual Ability (31)

1. Dynamic Visual Acuity @ 12, 24, & 36 Degrees per Second (DPS) (3)
2. Discomfort Glare Rating (1)
3. Glare Static Acuity (1)
4. Glare Contrast Sensitivity @ 4, 8, & 16 cycles per degree (CPD) (3)
5. Static Visual Acuity (Snellen) (1)
6. Contrast Sensitivity (Left Eye and Right Eye) @ 1.5, 3, 6, 12, & 18 CPD (10)
7. Color Vision (7 color plates) (7)
8. Total number of color vision plates correct (1)
9. Stereopsis (1)
10. Far Acuity (Optec) (1)
11. Far Vertical & Lateral Phoria (2)

General & Health-Related Information (10)

1. Height (inches) - self report (1)
2. Weight (lbs) - self report (1)
3. Total number of reported health problems (1)
4. Faces pain scale (1)
5. WHO (Five) Well-Being Index 1998 Version (1)
6. Total number of sleep problems (1)
7. Total number of sleep disorders (1)
8. Total hours of sleep estimated per day (1)
9. Education (1)
10. Total years driving (1)

Cognitive Ability (6)

1. Abbreviated Mental Test Score (AMTS) (1)
2. Self Estimate: How mental sharpness compared w/40s & 50s (1)
3. Meta-cognition: SE/AMTS (1)
4. Visualizing Missing Information (1)
5. Useful Field of View™ (1)
6. Trail Making B (1)

- We compare older driver and non-driver functional impairment profiles.
- Metrics showing statistically significant differences strongly favor the drivers.
- We develop assignment models based on a balanced set of key assessments.
- These models successfully assign individuals to driver or non-driver groups.

Fig. 1.
Conceptual model of older driver assessment problem.

Table 1

Participant age information by driving group and gender.

	Drivers		Non-Drivers	
	Age Range (n)	Mean Age (s.d.)	Age Range (n)	Mean Age (s.d.)
<i>Male</i>	70–84 (14)	77 (4.6)	65–87 (11)	79 (7.7)
<i>Female</i>	71–85 (12)	79 (4.3)	72–93 (12)	82 (6.7)

Table 2

Significant physical variables ($\alpha = 0.05$; bold value represents better score between the two groups).

Variable	Driver Mean	Non-Driver Mean	t Value	Pr > t	FDR p value
Ankle Torque Max – Plantarflexion (N-m)	33.531	22.620	2.52	0.016	0.061 ^{NS}
Ankle Torque Max – Dorsiflexion (N-m)	30.650	17.084	3.66	0.001	0.005
Upper Body Torque Max – Left (N-m)	25.140	16.396	3.69	0.001	0.005
Upper Body Torque Max – Right (N-m)	25.101	16.207	4.17	0.001	0.002

“NS”, Indicates where FDR p value is not significant.

Table 3

Significant psychomotor variables ($\alpha = 0.05$; bold means represent better performance).

Variable	Driver Mean	Non- Driver Mean	t Value	Pr > t	FDR p value
Ankle Initial Reaction Time (seconds, s)	2.438	2.860	-2.90	0.006	0.028
Ankle Peak Reaction Time (s)	2.940	3.640	-4.23	0.001	0.001
Hip Initial Reaction Time (s)	2.363	2.994	-3.53	0.001	0.006
Hip Peak Reaction Time (s)	2.743	3.679	-4.93	<0.001	0.001
Upper Body Initial Reaction Time (s)	2.218	2.622	-2.48	0.018	0.060 ^{NS}
Upper Body Peak Reaction Time (s)	2.460	3.024	-2.93	0.006	0.028

“NS”, Indicates where FDR p value is not significant.

Table 4

Significant perceptual variables ($\alpha = 0.05$; bold means represent better performance).

Variable	Driver Mean	Non- Driver Mean	t Value	Pr > t	FDR p value
Dynamic Visual Acuity (12 degrees/s)	0.4043	0.331	2.48	0.018	0.061 ^{NS}
Dynamic Visual Acuity (24 degrees/s)	0.3304	0.122	6.23	<0.001	<0.001
Dynamic Visual Acuity (36 degrees/s)	0.1391	0.044	4.71	<0.001	0.001
Glare Contrast Sensitivity (16 cycles/degree)	3.9130	4.632	-2.86	0.007	0.029
Corrected Snellen Acuity (20/x)	25.1920	41.522	-4.23	0.001	0.001
Contrast Sensitivity Left (1.5 cycles/degree)	5.1154	4.348	2.50	0.016	0.061 ^{NS}
Contrast Sensitivity Left (3 cycles/degree)	5.5769	4.870	2.33	0.024	0.078 ^{NS}
Contrast Sensitivity Left (6 cycles/degree)	4.1923	3.565	2.22	0.031	0.095 ^{NS}

“NS” Indicates where FDR p value is not significant.

Table 5

Other significant variables ($\alpha = 0.05$; bold means represent better scores).

Variable	Driver Mean	Non- Driver Mean	t Value	Pr > t	FDR p value
Number of Health Problems	2.692	5.696	-3.86	0.001	0.003
Years Driving	59.692	53.750	2.12	0.039	0.114 ^{NS}
Trail Making Part B (s)	114.30	191.59	-4.35	0.001	0.001

“NS” Indicates where FDR p value is not significant.

Table 6

Correlation matrix of initial variables selected for inclusion in the model.

Spearman Correlation Coefficient Pr > r under H0: Rho=0 Number of Observations	Upper Body Max Torque (left)	Upper Body Max Torque (right)	Upper Body Initial RT	DVA (24 deg/s)	Snellen Acuity	Trail Making (Part B)
Upper Body Maximum Torque (left)	1	0.914 <.001 39	-0.377 0.018 39	0.553 0.001 39	-0.321 0.046 39	-0.425 0.017 31
Upper Body Maximum Torque (right)		1	-0.460 0.003 39	0.565 0.001 39	-0.363 0.023 39	-0.404 0.024 31
Upper Body Initial Reaction Time (RT)			1	-0.194 0.236 39	0.396 0.013 39	0.139 0.457 31
Dynamic Visual Acuity (DVA @ 24 degrees/sec)				1	-0.449 0.003 41	-0.225 0.208 53
Snellen Acuity					1	0.250 0.126 39
Trail Making (Part B)						1

Table 7

Three variable model parameters.^a

Parameter	DF	Estimate	Standard Error	Wald χ^2	Pr > χ^2
Intercept	1	10.326	8.164	1.600	0.206
Average Upper Body Maximum Torque	1	-0.178	0.186	0.912	0.340
Dynamic Visual Acuity (24 degrees/sec)	1	43.595	26.903	2.626	0.105
Trail Making (Part B)	1	-0.126	0.083	2.336	0.127

^aThe likelihood ratio test for the overall effect is: $\chi^2=33.3217$, df=33, $p < 0.0001$.