Published in final edited form as:
Hum Factors. 2012 August ; 54(4): 663–674.

Naturalistic Validation of an On-Road Driving Test of Older Drivers

Brian R. Ott,
Warren Alpert Medical School of Brown University, Providence, Rhode Island

George D. Papandonatos,
Brown University, Providence, Rhode Island

Jennifer D. Davis, and
Warren Alpert Medical School of Brown University, Providence, Rhode Island

Peggy P. Barco
Washington University School of Medicine, St. Louis, Missouri

Abstract

Objective—The objective was to compare a standardized road test to naturalistic driving by older people who may have cognitive impairment to define improvements that could potentially enhance the validity of road testing in this population.

Background—Road testing has been widely adapted as a tool to assess driving competence of older people who may be at risk for unsafe driving because of dementia; however, the validity of this approach has not been rigorously evaluated.

Method—For 2 weeks, 80 older drivers (38 healthy elders and 42 with cognitive impairment) who passed a standardized road test were video recorded in their own vehicles. Using a standardized rating scale, 4 hr of video was rated by a driving instructor. The authors examine weighting of individual road test items to form global impressions and to compare road test and naturalistic driving using factor analyses of these two assessments.

Results—The road test score was unidimensional, reflecting a major factor related to awareness of signage and traffic behavior. Naturalistic driving reflected two factors related to lane keeping as well as traffic behavior.

Conclusion—Maintenance of proper lane is an important dimension of driving safety that appears to be relatively underemphasized during the highly supervised procedures of the standardized road test.

Application—Road testing in this population could be improved by standardized designs that emphasize lane keeping and that include self-directed driving. Additional information should be sought from observers in the community as well as crash evidence when advising older drivers who may be cognitively impaired.

Keywords

driving; aging; dementia; Alzheimer’s disease; cognitive impairment

Address correspondence to Brian R. Ott, Alzheimer’s Disease & Memory Disorders Center, Rhode Island Hospital, APC-6, 593 Eddy Street, Providence, RI 02903; BOtt@lifespan.org.

INTRODUCTION

A major risk factor for hazardous driving in the elderly is cognitive impairment. Road testing has long been a time-honored national standard for licensing new drivers. Given this apparent face validity, road testing has been widely adapted as a tool to assess driving competence of older people who may be at risk for unsafe driving because of accompanying medical illnesses, such as dementia (Carr & Ott, 2010). Test characteristics for on-road driving tests in elders without dementia show good interrater and test–retest reliability, internal consistency, and correlation between global and total performance scores in research settings (DiStefano & Macdonald, 2003; Justiss, Mann, Stav, & Velozo, 2006; Kay, Bundy, Clemson, & Jolly, 2008). However, the validity of this approach in the special population of cognitively impaired older drivers has not been as rigorously evaluated.

A wide range of variables may affect the reliability and validity of the road test. These factors include, but are not limited to, uncontrollable environmental factors within or across courses (e.g., changes in traffic density, course familiarity), driver experience (e.g., occupations that require driving at a very high skill level), driver intrinsic factors (e.g., personality traits, lifelong driving habits, test-taking anxiety, fatigue), and assessor differences (e.g., training, biases, experience). Road test designs make comparisons across studies difficult, and differences include the number of assessors in the vehicle, scoring procedures, and the environment, duration, and course (Wheatley & Di, 2008).

A large number of research studies have used road tests as an outcome measure in neuropsychological studies attempting to define a screening tool for impaired driving. Carr, Jackson, Madden, and Cohen (1992) demonstrated stability of driving skills for older drivers using a low-intensity standardized road test on a college campus. Odenheimer et al. (1994) later examined 30 older drivers with a range of cognitive function including dementia on a combined closed course and open-traffic course, demonstrating a significant relationship between performances in the two environments. Interrater reliabilities were .84 for the closed course and .74 for the in-traffic component. Internal consistency for the closed course was .78 and for the in-traffic component was .89. This landmark study demonstrated the potential for road testing as a valid and reliable outcome in research studies of older drivers.

More recently, on-road tests in mixed normal and dementia samples from two large studies have shown interrater reliabilities ranging from a kappa of .83 to a kappa of .92 (Hunt et al., 1997b; Ott et al., 2008). Retest reliability, however, was lower (.53 for global ratings and .76 for total scores) probably because of traffic changes that occurred between the two road tests or variability in human performance (e.g., because of medication, fatigue, etc.; Hunt et al., 1997a).

The validity of road testing cognitively impaired older drivers as a measure of driving competence can be challenged by the lack of comparisons to real-world driving. Test-taking anxiety in some participants may affect road test results (Bhalla, Papandonatos, Stern, & Ott, 2007). Some older adults have had poor lifetime driving habits that could contribute to impaired driving outcomes yet could not be attributed to their new onset dementia. This construct has not been adequately defined, although a history of an at-fault crash may be a simple proxy (O'Connor, Kapust, Lin, Hollis, & Jones, 2010). To our knowledge, only one study has validated the standardized road test against motor vehicle crash frequency in a mixed sample that included drivers with dementia. In this study, the number of collisions and moving violations per 1,000 miles was significantly correlated with driving test performance during the 2 years preceding the examination. The participants included 13 participants with Alzheimer's disease (AD), 12 participants with vascular dementia, 15 diabetic patients, 24 normal elders, and 16 younger controls (Fitten et al., 1995).

An innovative aspect of current research is the addition of naturalistic study methods. Recently, a significant number of errors on a standardized road test in a university setting were accounted for by lack of driver familiarity with the traffic environment (Uc et al., 2009). We recently reported results of research comparing standardized road test performance to naturalistic driving using video recordings, which found only a modest correlation between the two methods. This may be attributed in part to differences in the driving evaluation environment (Davis et al., 2011). In this study of 80 participants with either no cognitive impairment or a diagnosis of questionable to mild dementia, global ratings of safety on the road test and naturalistic driving were comparable; however, cognitively impaired drivers were more likely to receive a pass without restrictions in naturalistic driving compared to the road test, where restrictions were more likely to be recommended. A higher frequency of errors detected in the naturalistic driving environment for both normal and cognitively impaired drivers did not affect global ratings. Error scores between settings were only modestly correlated (Spearman's $\rho = .39$, $p < .001$).

The current article uses factor analytic techniques to identify key attributes of both a standardized road test typically performed in unfamiliar surroundings by a driving assessment specialist and naturalistic driving, as captured by direct observation. In addition, we critically examine test characteristics of a new scale that we introduced for evaluating naturalistic driving errors and determine specific scale components that were more heavily weighted by instructors as important measures of driving competency. Based on these analyses, recommendations are offered for optimal road test design and for clinical evaluations of normal and cognitively impaired older drivers.

METHOD

Participants

Study participants were required to be 55 to 80 years of age, with a valid driver's license and no at-fault accidents within the past year. Healthy older adults had no history of dementia and a Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) score greater than 26. Cognitively impaired participants were recruited primarily from a hospital-based memory disorders program. Dementia severity was measured with the Clinical Dementia Rating scale (CDR; Morris, 1993); only participants with CDR scores of 0.5 (questionable dementia) or 1 (mild dementia) were eligible to participate. Diagnosis of AD was based on the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984). Exclusion criteria for both groups included reversible causes of dementia, physical or ophthalmologic disorders that might impair driving abilities, mental retardation, schizophrenia, bipolar disorder, and alcohol or substance abuse within the previous year. Anxiolytic and antipsychotic medications were permitted, but dosages were required to be stable for at least 6 weeks before study entry.

At the first study visit, written informed consent was obtained from all participants. The Rhode Island Hospital Institutional Review Board approved all procedures. Participants were screened for inclusion criteria at an office visit that included the neurological examination, vision screen, MMSE, and CDR. If the participants met inclusion criteria, they were scheduled for an on-road driving test. If they failed, ethical considerations dictated that they be dropped from the study and be advised to stop driving. There were five failures in the original group. Two retested and passed, and they are included in this sample. If participants passed the first test or a retest, cameras were installed in their vehicles to record real-world driving. Participants who subsequently received a fail rating on naturalistic driving were counseled by the driving instructor and research staff about driving cessation.

The Standardized Road Test

The Rhode Island Road Test (RIRT) is an adaptation of the Washington University Road Test, a standardized driving measure with previously established reliability (Hunt et al., 1997b) that was adapted for use on comparable streets in Rhode Island. A total of 28 driving behaviors were rated on a trichotomous scale with higher scores reflecting poorer performance (0 = *unimpaired*, 1 = *mildly impaired*, 2 = *moderately to severely impaired*); the behaviors are described in detail in Table 1. By design, several of these behaviors were rated on multiple occasions at various points along the course, for a total of 480 possible behavioral ratings. Hence, total scores could range from 0 (*best*) to 960 (*worst*). Two dependent variables were generated from RIRT: an average score, reflecting the sum of observed behavioral ratings divided by their number, and a categorical rating of the participants' overall driving ability (*pass*, *marginal*, or *fail*).

The structure of the currently used RIRT examination is detailed in Table 2. Within 1 month of the baseline visit, participants were administered the RIRT by a professional driving instructor, who was blind to diagnosis, during daylight hours under good road conditions. The test covered 6.5 miles of urban terrain without highway driving and required 45 min to complete. A pretest of basic vehicle operation was performed in a parking lot prior to the on-road test. The driving instructor accompanied the participant in a specially fitted vehicle that had a brake on the passenger side for use in emergency situations. The driving instructor provided verbal instructions to complete the course but did not in any other way cue the driver.

Participants were rated by the driving instructor in the manner described by Hunt et al. (1997b). Previous research employing the RIRT and the same driving instructor showed adequate interrater reliability for 20 participants (rated by a second professional driving instructor in the back seat) with moderate to substantial agreement for the global rating ($\kappa = .83$ for linear weighted ratings and $\kappa = .92$ for quadratic weighted ratings). The Pearson correlation coefficient between the two raters for the average RIRT score was .87 (Brown et al., 2005).

Naturalistic Driving

The cars of participants who passed the RIRT were equipped with four small cameras that were placed on the dash and in the back of the vehicle. Cameras were placed in the participants' cars on the same day that they completed the RIRT. One camera faced the participant, another faced forward, and the back two cameras faced diagonally forward to capture the environment to the sides of the vehicle. Participants were instructed to drive in their typical environment and to follow their daily routine. Video from the cameras was automatically downloaded to a portable digital video recording device that was placed under the passenger seat. The cameras were low profile (3.5 in. wide by 2.5 in. deep by 3 in. tall) and installed by the study research assistant. Cameras were installed for a 2-week period to capture at least 4 hr of daytime driving trips from home. If the participant fell short of 4 hr, the cameras remained in the car for another week to capture a full 4 hr of driving. Of the participants, 64% consented to audio recording of the home driving segments. Only daytime driving was viewed, as the nighttime video quality was suboptimal for coding. The research assistant reviewed the driving segments for the 2-week period and provided the first 4 hr of consecutive daytime driving to the professional driving instructor for review.

The video ratings were done at least 1 month after the road test, and the instructor was not allowed to review his ratings from the road test to minimize bias in review of the naturalistic video. The same driving instructor rated the home driving video and the road test to maximize comparability of ratings in the different driving environments. The driving instructor rated the naturalistic driving according to the Composite Driving Assessment Scale (CDAS), a scale specifically developed for the purpose of this study whose items are described in Table 3. Because the number and types of driving situations in naturalistic driving assessment cannot be controlled in the same fashion as on a standardized road test, each item was rated on a global basis as *unimpaired*, *mildly impaired*, or *moderately to severely impaired*. Scale content was based on input from the driving instructor's standard assessment as well as from the RIRT to enhance generalizability of the study rating method. Behaviors were divided into discrete events (i.e., maneuvers) and global events (i.e., attention, attitude, reaction time). Each specific driving behavior was rated on a 3-point Likert-type scale (0–2). Total scores could range from 0 to 60, with higher scores reflecting poorer performance. An average score was calculated by summing the scores on each individual driving behavior and dividing by the number driving behaviors that could be assigned a score because they occurred during the 4-hr driving epoch. Behaviors that were not encountered even once during naturalistic driving (e.g., "response to emergency vehicles") were excluded from these calculations. It should be noted that each driving behavior received the same weight in calculating the CDAS score, whereas the RIRT score weighted each behavior by the frequency with which it was encountered during the road test. In addition, the driving instructor made a global rating for the CDAS using the same global rating scale (*pass*, *marginal*, or *fail*) used on the RIRT.

Statistical Analyses

Most published driving studies use a total error score to reflect performance, and few studies distinguish between different types of errors. To address whether there are multiple dimensions of driving behavior that are evaluated in the on-road test and naturalistic driving, and to identify driving maneuvers that are most relevant to performance in each setting, we conducted principal factor analyses of the RIRT and CDAS using S-Plus™ 8.2 (Insightful Corporation, 2008) with factor loadings subjected to varimax rotation.

Given the sample size ($N \leq 80$), tests of dimensionality based on asymptotic procedures were deemed unreliable. Instead, we compared single-factor to two-factor solutions using a graphical method known as the biplot (Gabriel, 1971; Gower, 2006). In a biplot, each item of the measurement scale under consideration is presented by an arrow. Longer arrows reflect more informative items, whereas overlapping arrows are indicative of items that may be informative on their own but potentially contribute redundant information. Arrows pointing in the same direction are positively correlated, items that are negatively correlated point in the opposite direction, and uncorrelated items appear at right angles. When all arrows point in the same direction, this can be taken as evidence that a two-factor solution provides little additional information over a single-factor solution.

RESULTS

Sample Description

Demographic characteristics for all 80 participants as well as between-group comparisons of healthy and cognitively impaired participants are included in Table 4.

Interrater Agreement of Driving Instructors

To establish interrater reliability of the CDAS, home driving video for 25 participants was randomly selected and independently rated by two professional driving instructors. To promote consistency in ratings, prior to rating the substudy reliability sample, the instructors reviewed several other videos from different drivers and reached consensus on each after discussion of their initial CDAS ratings. The sample included 12 healthy elderly people and 13 patients (CDR = 0.5, n = 10; CDR = 1, n = 3). Given the small sample size, global ratings were dichotomized into *pass* versus *marginal* or *fail*. Interrater agreement was in the moderate range (κ = .45).

Importance of Individual Items in Deriving the Global Scores

To understand the basis for differences in global determinations found in previous analyses, the two instructors were asked to weight each of the items included in the CDAS on a 1 to 10 scale, according to whether they perceived the item as *extremely important* (10) versus *not important at all* (1) in helping them formulate their global rating. Item importance comparisons between instructors using these weights showed poor agreement ($\kappa = -.04$), and correlation between instructors for item weights was low and not significant (ρ = .26, p = .15), indicating that the two instructors emphasized different aspects of driving in arriving at their global impressions. There was perfect agreement on the weights for “follows directions” and “attitude” and discrepancy of 5 points or greater for “starting,” “backing,” and “parking,” suggesting that one of the two instructors primarily considered performance in traffic situations. Of note, the two items with perfect agreement are included in the RIRT, whereas the three discrepant items are included only in the road pretest.

Results of the Factor Analysis

A graphical summary of a two-factor solution of the RIRT is depicted in Figure 1, in which individual components of the 28-item driving ability scale are numbered in the same order as in Table 1. Results revealed a homogeneous cluster of 21 items with strong intraclass correlation (ICC = .40) and high internal consistency reliability (Cronbach’s α = .93). This primary dimension of driving ability (RIRT Factor 1) explained 31% of the variance in the overall scale and was dominated by *driving awareness* items (spatial, environmental, traffic, and signage). Items related to stopping and parking (5, 27, 28) were uninformative. Four other items (3, 4, 13, 21) clustered in a secondary dimension (RIRT Factor 2) that explained an additional 8% of the variability in the data and appeared related to *speed control*. These items showed an equally strong ICC (ICC = .45), resulting in acceptable reliability (Cronbach’s α = .80) for this subscale. In sum, driving performance on the RIRT was best explained by a single 21-item factor dominated by behaviors related to awareness of traffic situations.

A graphical summary of a two-factor solution of the CDAS in depicted in Figure 2, in which individual components of the 30-item driving ability scale are numbered in the same order as in Table 3. Results revealed a major factor (CDAS Factor 1) defined by 20 items that were numerous enough to produce a reliable subscale (Cronbach’s α = .89), despite their modest correlation (ICC = .30). This factor was dominated by behaviors related to *responding to traffic and maneuvering the vehicle* and explained 14% of the variance. Of almost equal importance was a minor factor (CDAS Factor 2) defined by four items (4, 14, 18, 25) with a stronger ICC (ICC = .39). However, this factor showed only moderate subscale reliability (Cronbach’s α = .73) because of the small number of items. It explained 12% of the variance and appeared to capture *proper lane keeping*. Finally, four items (11, 12, 19, 26) were uninformative because of low variability. Taken together, these results suggest that performance in the naturalistic setting was best explained by a two-factor solution that describes two distinct dimensions of driving.

To determine the contribution of driving environment, a factor analysis was completed on the previously video recorded subsample of 47 on-road tests. The driving instructor completed the RIRT in-car measure during the on-road test and later reviewed the videotape to complete the CDAS for the on-road test. He also completed the CDAS for the naturalistic driving. Therefore, each participant in this analysis had an RIRT, CDAS of the road test, and

CDAS of the home driving. Demographics for this subset of 47 participants are presented in Table 5.

Factor analysis of road test behavior measured with the CDAS (RCDAS) supported a single dimension of driving behavior rather than the two distinct dimensions that were captured when the CDAS was used to measure naturalistic driving behavior (HCDAS), arguing for an important contribution of environment when assessing driving behavior.

Since both naturalistic driving and road test course difficulty varied across participants, ratings of driving course difficulty were also made by the instructor on a 1–10 Likert-type scale, with a rating of 1 reflecting a *simple* driving environment and a rating of 10 reflecting an *extremely challenging* driving environment. Ratings for naturalistic driving course difficulty pertain to the HCDAS score and were averaged across the entire 4 hr of naturalistic driving; they averaged 5.8 ± 1.2 (range = 2–8). Ratings for road test course difficulty pertain to both the RCDAS and RIRT scores, as they were obtained simultaneously for each participant; they averaged 6.9 ± 0.4 (range = 6–8). After adjusting for variation in driving course difficulty, the Spearman rank correlation between the RIRT and RCDAS ($\rho = .62$, $p < .001$) was much stronger than the correlation between the RIRT and HCDAS ($\rho = .24$, $p = .11$) as well as that between the RCDAS and the HCDAS ($\rho = .43$, $p = .003$). In sum, results from the factor analyses and correlation analyses indicate that similarity in the driving environment appeared more relevant in ensuring a strong association than similarity in the measurement method.

DISCUSSION

Factor analysis of the standardized road test revealed a primary dimension of driving ability that explained almost 4 times as much variability in the data as the secondary, minor dimension (31% vs. 8%). This major dimension was characterized by maneuvers related to driving awareness of signage and traffic behavior, skills likely dependent on higher order executive skills, such as planning and selecting behaviors based on the environmental cues of traffic patterns and signage.

In contrast, factor analysis of home driving produced two dimensions of approximately equal importance (14% vs. 12%), one capturing the driver's ability to respond to traffic and the other proper lane keeping. The former may place greater demand on higher order executive skills and the latter on visual attention. It is interesting that evaluation of on-road test performance using the home driving CDAS measure showed a factor structure similar to the RIRT, suggesting that it was the environment that dictated performance rather than the different measurement approaches. Taken together, results suggest that adequate road test performance requires a more restricted set of driving skills than those required for naturalistic home driving.

Previous research has examined the frequency and types of errors made by cognitively impaired older drivers. In one study of normal older drivers and drivers with Parkinson's disease, participants drove a standardized route in urban and rural settings in an instrumented vehicle. Lane violations was the most common error category in both groups (Uc et al., 2009). In another study of drivers with AD, using the same methods, keeping lane position was the most common error encountered in older controls as well as patients (Dawson, Anderson, Uc, Dastrup, & Rizzo, 2009). A third study involving road tests of normal older drivers, drivers with AD, and drivers with Parkinson's disease found a high frequency of lane changing errors affecting up to 75% of AD drivers, with frequencies in Parkinson drivers being intermediate between those with AD and normal controls (Grace et al., 2005).

These and other studies suggest that inaccurate lane positioning and failure to check blind spots while changing lanes are important factors in older drivers that may potentially lead to motor vehicle accidents. In a study of 267 healthy older adults, lane positioning, approach, and blind spot monitoring were the most common errors. It is notable that 25% of the total duration of each of these road tests was conducted under self-directed navigation, which may more closely approximate naturalistic driving. Driving instructor interventions during the self-directed phase of the exam were associated with significantly higher retrospective as well as prospective crashes (Wood et al., 2009).

The results from our factor analyses suggest that two distinct, albeit correlated, dimensions of driving behavior are relevant for home driving, whereas only one dimension seems necessary for the road test. Road test designs that more closely approximate the driving skills required in the home environment of older drivers (such as being able to control one's own vehicle and self-direction) and that stress proper lane keeping, in addition to properly responding to traffic and signage, may improve the ecological validity of the road test.

Although previous road test studies have reported relatively high interrater reliability (DiStefano & Macdonald, 2003; Hunt et al., 1997b; Justiss et al., 2006; Kay et al., 2008; Odenheimer et al., 1994; Ott et al., 2008), these studies incorporated second raters from the same research team and same type of geographic region. Our data suggest that generalizability of these observations to road test practices nationwide is probably limited. Review of video driving segments by two research team driving instructors from different geographies and with different professional backgrounds—one an occupational therapist in Missouri (P.B.), the other a private driving instructor in Rhode Island (T.S.)—revealed only modest agreement. The apparent subjective nature of global impressions of driving performance suggests the need for development of special standards of experience and training for the evaluation of older and cognitively impaired drivers. Such standards, however, may not limit variability because of environment and regional driving differences between locales. This latter issue could be addressed by limiting evaluations to only standardized closed courses; however, such rigidly defined courses are unlikely to give an accurate picture of real-world driving in the natural home setting. To this end, our current study provides some insight into the optimal design for standardized road tests.

It is important to note that road tests themselves are done in different environments and can produce different outcomes for the same types of tested drivers. For example, some road tests have more stopping and starting in heavily populated areas (e.g., around a medical center). These provide less opportunity for lane keeping since the driver is closely following other vehicles, and those vehicles help to mark position and provide more visual cues. Other road tests have longer stretches of road (with fewer stops and starts in less heavily congested environments) and consequently fewer cues for lane position. These issues need to be considered in plans for standardization of road testing across sites.

A limitation of our study is the use of the same rater for both road test and home driving assessments. Recall bias by the rater may have limited our ability to see differences in these settings. This design, however, did enable us to draw conclusions about the differences between the two driving environments unaffected by interrater reliability issues, which, as mentioned earlier, may be significant.

Another limitation in our study design was the restriction of naturalistic video recordings to people who were able to pass a road test, which restricted the range of naturalistic observations we could have made on the most potentially hazardous drivers. Therefore, our observations may not be applicable to the most dangerous drivers. This design was required by the Institutional Review Board over concerns about safety since many authorities and published practice guidelines (Carr & Ott, 2010) feel that road test failure by a cognitively impaired older driver provides face-valid evidence that a person is an unsafe driver. Nevertheless, drivers regarded as marginal on the road test were permitted to participate in video recordings and thus provided useful information about drivers with impairment in the home environment.

The methods employed in this study were not completely naturalistic insofar as the mere presence of the cameras in the car could have influenced behavior. The cameras we employed were quite small, however, and the nature of some of the behaviors we observed suggested to the research team that drivers were focused on their driving and were not aware of the recording devices while driving. A survey of the participants about their awareness of the recordings would have strengthened the certainty of this conclusion.

The validity of assessment of driving skills using a structured road test versus assessment of home driving using video recordings or observations by family members for safe driving remains a topic of debate. Future studies of large numbers of participants looking at safety outcomes such as motor vehicle accidents would help to answer this question. Research of this type in the cognitively impaired at-risk driver is practically limited to retrospective analyses because of ethical and safety issues since cognitively impaired drivers are advised to stop or restrict driving when concerns are identified by these methods. In the meantime, we suggest that obtaining additional information related to driving in the home environment may be a useful supplement to standardized road tests in the evaluation of the older driver who may be cognitively impaired.

Some occupational therapists and certified driving rehabilitation specialists allow for an assessment option where the client is evaluated in the area where he or she habitually drives. Our data lend support to this practice and suggest that in some cases where cognitively impaired older drivers perform safely in their home environment, restriction of driving privileges to drive only in such situations may be appropriate.

Acknowledgments

Road tests as well as the video Composite Driving Assessment Scale ratings for the road test and home driving were performed by Timothy Souza from ABC/ACE Driving School, Brooklyn, Connecticut. Helpful review and comments during preparation of the article were provided by David B. Carr. These results were presented in part at the International Conference on Alzheimer's Disease in Paris on July 18, 2011. This work was supported by a grant from the National Institute on Aging at the National Institutes of Health (R01 AG016335 to B.R.O.).

References

Bhalla RK, Papandonatos GD, Stern RA, Ott BR. Anxiety of Alzheimer's disease patients before and after a standardized on-road driving test. Alzheimer's and Dementia. 2007; 3(1):33–39.

Brown LB, Ott BR, Papandonatos GD, Sui Y, Ready RE, Morris JC. Prediction of on-road driving performance in patients with early Alzheimer's disease. Journal of the American Geriatric Society. 2005; 53(1):94–98.

Carr D, Jackson TW, Madden DJ, Cohen HJ. The effect of age on driving skills. Journal of the American Geriatric Society. 1992; 40(6):567–573.

Carr DB, Ott BR. The older adult driver with cognitive impairment: "It's a very frustrating life. Journal of the American Medical Association. 2010; 303(16):1632–1641. [PubMed: 20424254]

Davis, JD.; Papandonatos, GD.; Miller, LA.; Knott, S.; Festa, EK.; Heindel, WC.; Barco, P.; Ott, BR. Road test and naturalistic driving behavior in older adults with and without memory loss: Environment matters. Paper presented at the annual meeting of the International Neuropsychological Society; Boston, MA. 2011 Feb.

Dawson JD, Anderson SW, Uc EY, Dastrup E, Rizzo M. Predictors of driving safety in early Alzheimer disease. Neurology. 2009; 72(6):521–527. [PubMed: 19204261]

DiStefano M, Macdonald W. Assessment of older drivers: Relationships among on-road errors, medical conditions and test outcome. Journal of Safety Research. 2003; 34:415–429. [PubMed: 14636664]

Fitten LJ, Perryman KM, Wilkinson CJ, Little RJ, Burns MM, Pachana N, Mervis JR, Malmgren R, Siembieda DW, Ganzell S. Alzheimer and vascular dementias and driving: A prospective road and laboratory study. Journal of the American Medical Association. 1995; 273(17):1360–1365. [PubMed: 7715061]

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; (3):189–198. [PubMed: 1202204]

Gabriel KR. The biplot graphical display of matrices with application to principal component analysis. Biometrika. 1971; 58:453–467.

Gower, JDHJ. Biplots. London, UK: Chapman & Hall; 2006.

Grace J, Amick MM, D'Abreu A, Festa EK, Heindel WC, Ott BR. Neuropsychological deficits associated with driving performance in Parkinson's and Alzheimer's disease. Journal of the International Neuropsychological Society. 2005; 11(6):766–775. [PubMed: 16248912]

Hunt LA, Murphy CF, Carr D, Duchek JM, Buckles V, Morris JC. Environmental cueing may effect performance on a road test for drivers with dementia of the Alzheimer type. Alzheimer's Disease and Associated Disorders. 1997a; 11(Suppl 1):13–16.

Hunt LA, Murphy CF, Carr D, Duchek JM, Buckles V, Morris JC. Reliability of the Washington University Road Test: A performance-based assessment for drivers with dementia of the Alzheimer type. Archives of Neurology. 1997b; 54(6):707–712. [PubMed: 9193205]

Justiss MD, Mann WC, Stav WB, Velozo C. Development of a behind-the-wheel driving performance assessment for older adults. Topics in Geriatric Rehabilitation. 2006; 22(2):121–128.

Kay L, Bundy A, Clemson L, Jolly N. Validity and reliability of the on-road driving assessment with senior drivers. Accident Analysis and Prevention. 2008; 40(2):751–759. [PubMed: 18329430]

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34(7):939–944. [PubMed: 6610841]

Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology. 1993; 43(11):2412–2414. [PubMed: 8232972]

O'Connor MG, Kapust LR, Lin B, Hollis AM, Jones RN. The 4Cs (crash history, family concerns, clinical condition, and cognitive functions): A screening tool for the evaluation of the at-risk driver. Journal of the American Geriatric Society. 2010; 58(6):1104–1108.

Odenheimer GL, Beaudet M, Jette AM, Albert MS, Grande L, Minaker KL. Performance-based driving evaluation of the elderly driver: Safety, reliability, and validity. Journal of Gerontology. 1994; 49(4):M153–M159. [PubMed: 8014389]

Ott BR, Festa EK, Amick MM, Grace J, Davis JD, Heindel WC. Computerized maze navigation and on-road performance by drivers with dementia. Journal of Geriatric Psychiatry and Neurology. 2008; 21(1):18–25. [PubMed: 18287166]

Uc EY, Rizzo M, Johnson AM, Dastrup E, Anderson SW, Dawson JD. Road safety in drivers with Parkinson disease. Neurology. 2009; 73(24):2112–2119. [PubMed: 20018639]

Wheatley CJ, Di SM. Individualized assessment of driving fitness for older individuals with health, disability, and age-related concerns. Traffic Injury Prevention. 2008; 9(4):320–327. [PubMed: 18696388]

Wood JM, Anstey KJ, Lacherez PF, Kerr GK, Mallon K, Lord SR. The on-road difficulties of older drivers and their relationship with self-reported motor vehicle crashes. Journal of the American Geriatric Society. 2009; 57(11):2062–2069.

Biographies

Brian R. Ott is a professor of neurology at the Warren Alpert Medical School of Brown University in Providence. He received his MD from Thomas Jefferson Medical College, Philadelphia, in 1979.

George D. Papandonatos is associate professor of biostatistics at Brown University in Providence. He received his PhD in statistics from the University of Minnesota (Twin Cities), Minneapolis, in 1994.

Jennifer D. Davis is an assistant professor of psychiatry and human behavior at the Warren Alpert Medical School of Brown University in Providence. She received her PhD from the University of Utah in clinical psychology in 2002.

Peggy P. Barco is an instructor in the Program of Occupational Therapy and Department of Medicine at Washington University Medical School in St. Louis. She received her master's of science in occupational therapy from Washington University Medical School in 1987 and is currently pursuing her doctorate degree in occupational therapy.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

KEY POINTS

- Adequate road test performance requires a more restricted set of driving skills than do those required for naturalistic home driving.
- The major dimension examined by the standardized road test is characterized by maneuvers related to driving awareness of signage and traffic behavior, skills likely dependent on higher order executive function, such as planning and selecting behaviors based on the environmental cues of traffic patterns and signage.
- Driving in the natural home setting involves two dimensions of approximately equal importance, one involving the driver's ability to respond to traffic (similar to the road test) and the other involving proper lane keeping. The latter factor is probably more dependent on visual attention.

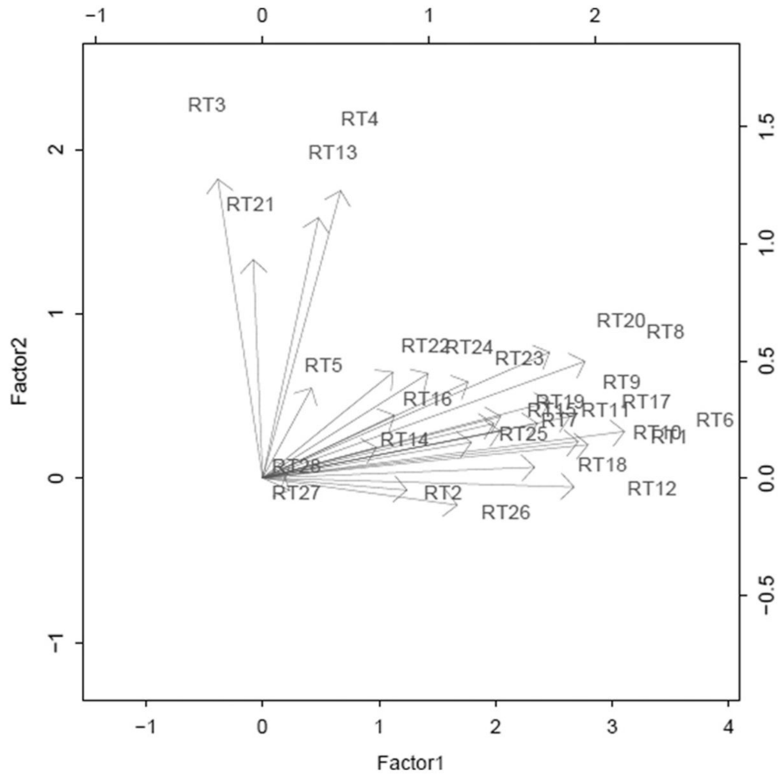

Figure 1.
Biplot of the two-factor solution for the Road Island Road Test (RIRT). RIRT items measuring road test performance are numbered in the same order as in Table 1.

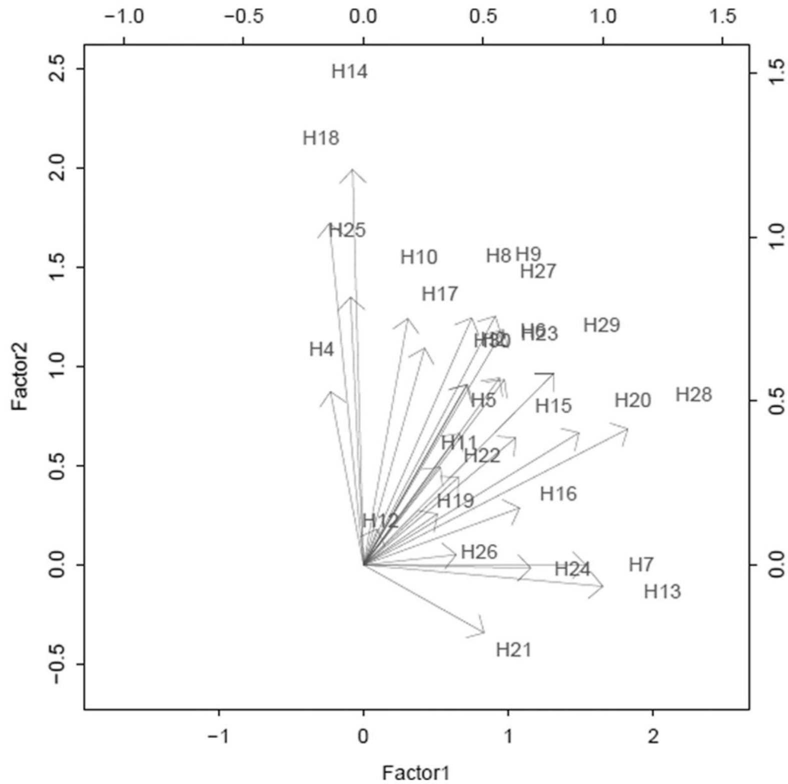

Figure 2.
Biplot of two-factor solution for the Composite Driving Assessment Scale (CDAS). CDAS items measuring performance on naturalistic driving from home are numbered in the same order as in Table 3.

TABLE 1

Rhode Island Road Test Scale Used to Evaluate On-Road Driving Performance

1. Awareness of signs	0	1	2
2. Light awareness	0	1	2
3. Approaches intersection at appropriate speed	0	1	2
4. Brakes smoothly and accurately	0	1	2
5. Comes to a complete stop	0	1	2
6. Aware of traffic situations	0	1	2
7. Responds appropriately to signal	0	1	2
8. Scans	0	1	2
9. Uses mirrors	0	1	2
10. Positions for turn	0	1	2
11. Yields right of way	0	1	2
12. Proceeds timely	0	1	2
13. Accelerates smoothly and accurately	0	1	2
14. Signal use	0	1	2
15. Attends to task	0	1	2
16. Awareness of pedestrians	0	1	2
17. Spatial awareness	0	1	2
18. Lane selection	0	1	2
19. Lane keeping	0	1	2
20. Awareness of environment	0	1	2
21. Speed control	0	1	2
22. Steering control	0	1	2
23. Uses mirrors for lane change	0	1	2
24. Checks blind spots	0	1	2
25. Changes lanes without verbal cue	0	1	2
26. Scans while in motion	0	1	2
27. Shifts for park	0	1	2
28. Turns vehicle engine off	0	1	2

Note. Key: 0 = *unimpaired*, 1 = *mildly impaired*, 2 = *moderately to severely impaired*.

TABLE 2

Rhode Island Road Test Design

Total Score	**0–960**
Global score	Pass; marginal; fail
# traffic situations	34
# driving behaviors	28
# driving behavior ratings	480
Point grades per rating	0, 1, 2
# turns	22
Left	12
Right	10
Lights	16
Stops	16
Signs	14
Course length (miles)	6.5
Time to complete (minutes)	45
Evaluators	Front seat instructor–rater and scorer
Environment	Mixed residential and commercial urban

TABLE 3

Composite Driving Assessment Scale (CDAS) Used to Evaluate Naturalistic Driving

Discrete Events			
1. Follows directions from other occupant	0	1	2
2. Responds appropriately to traffic signs, lights, and road markings	0	1	2
3. Demonstrates appropriate response during emergency situations	0	1	2
4. Starting	0	1	2
5. Backing	0	1	2
6. Parking (horizontal)	0	1	2
7. Uses turn signal	0	1	2
8. Left turn	0	1	2
9. Right turn	0	1	2
10. Proceeds without hesitation	0	1	2
11. Accelerates smoothly, accurately, and timely	0	1	2
12. Stops completely at sign	0	1	2
13. Checks blind spots	0	1	2
14. Centers vehicle in travel lane	0	1	2
15. Brakes smoothly with adequate time and space	0	1	2
16. Demonstrates appropriate steering recovery	0	1	2
17. Speed control	0	1	2
18. Lane keeping	0	1	2
19. Awareness of pedestrians		0	1
Global Assessments			
20. Scans timely and comprehensively	0	1	2
21. Uses mirrors properly	0	1	2
22. Maintains safe distance from other vehicles while in motion	0	1	2
23. Able to judge spatial positioning for turns and parking	0	1	2
24. Demonstrates proper use of vehicle equipment	0	1	2
25. Attention	0	1	2
26. Attitude	0	1	2
27. Awareness of how driving affects others	0	1	2
28. Anticipation of traffic situations	0	1	2
29. Reaction time	0	1	2
30. Response to other vehicle's brake lights, signals, etc.	0	1	2
Total		_________	(/60)

Note. Key: 0 = *unimpaired*, 1 = *mildly impaired*, 2 = *moderately to severely impaired.*

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 4

Demographic Characteristics and Driving Performance Scores for All Study Participants

	Normal Participants (n = 38)		**Cognitively Impaired Participants (n = 42)**		**All Participants (N = 80)**	
	M	***SD***	***M***	***SD***	***M***	***SD***
Gender (% male)	60.5		47.6		53.8	
Age (years)**	68.9	7.2	76.1	6.0	73.1	7.3
Driving experience (years)*	52.2	6.8	56.3	9.2	54.3	8.4
Education (years)**	16.4	4.0	13.8	3.7	15.1	4.0
MMSE**	29.5	0.7	25.0	3.1	27.1	3.2
Road test error score**	0.03	0.02	0.08	0.06	0.06	0.05
Home CDAS error score**	0.10	0.08	0.20	0.13	0.15	0.12

Note. Values are *M* and *SD* or %. MMSE = Mini-Mental State Examination; CDAS = Composite Driving Assessment Scale.

* $p < .05$.

** $p < .005$.

TABLE 5

Demographic Characteristics and Driving Performance Scores of Study Participants Who Also Had Their Road Test Video Recorded

	Normal Participants (*n* = 28)		Cognitively Impaired Participants (*n* = 19)		All Participants (*N* = 47)	
	M	***SD***	***M***	***SD***	***M***	***SD***
Gender (% male)*	67.9		53.8		55.1	
Age (years)**	68.8	7.1	75.9	5.8	71.7	7.5
Driving experience (years)*	50.5	6.1	55.7	10.3	52.6	8.4
Education (years)*	16.3	3.4	13.6	3.9	15.2	3.9
MMSE**	29.6	0.7	24.7	3.8	27.6	3.4
Road test error score**	0.04	0.03	0.08	0.06	0.06	.05
Road test CDAS error score*	0.11	0.07	0.16	0.06	0.13	0.07
Home CDAS error score*	0.11	0.08	0.18	0.11	0.14	0.10

Note. Values are *M* and *SD* or %. MMSE = Mini-Mental State Examination; CDAS = Composite Driving Assessment Scale.

* $p < .05$.

** $p < .005$.