The Addenbrooke's Cognitive Examination Revised as a potential screening test for elderly drivers

Inês S. Ferreira a,*, Mário R. Simões b, João Marôco c

a Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo Comportamental (CINECC), Faculty of Psychology and Educational Sciences, University of Coimbra, Coimbra, Portugal
b Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo Comportamental (CINECC), Psychological Assessment Lab., Faculty of Psychology and Educational Sciences, University of Coimbra, Coimbra, Portugal
c Instituto Superior de Psicologia Aplicada – Instituto Universitário, Lisboa, Portugal

A R T I C L E   I N F O

Article history:
Received 31 October 2011
Received in revised form 20 March 2012
Accepted 29 March 2012

Keywords:
Addenbrooke’s Cognitive Examination Revised
Fitness to drive
Prediction
On-road assessment
Older driver

A B S T R A C T

Considerable research has shown that neuropsychological tests are predictive of real-world driving ability. The Mini-Mental State Examination (MMSE) is a brief cognitive test that has been commonly used in the assessment of older drivers. However, this test has inherent problems that limit its validity to evaluate cognitive abilities related to driving and to screen for driving impairments in non-demented people. Therefore, it is useful to test new screening instruments that may predict potential unsafe drivers who require an in-depth neuropsychological assessment in a specialised centre. To date, the utility of the Addenbrooke’s Cognitive Examination Revised (ACE-R) as an indicator of driving ability has not been established. In the current study, fifty older drivers (mean age = 73.1 years) who were referred for a psychological assessment, the protocol of which included the ACE-R, underwent an on-road driving test. Using linear discriminant analyses, the results highlighted the higher classification accuracy of the ACE-R compared to the MMSE score, particularly for detecting unsafe drivers. Measures of visuospatial and executive functions, which are not incorporated in the MMSE score, had an incremental value in the prediction of driving ability. This emerging brief cognitive test may warrant additional study for use in the fitness to drive assessment of older adults.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Traffic safety for older drivers is a growing public health issue. The demographics of aging and life expectancy give rise to an increasing number of older licensed drivers on the road, as the personal automobile constitutes the primary mode of transport for the emerging elderly drivers (Organisation for Economic Co-operation and Development, 2001). Older drivers have a high rate of traffic fatality, caused both by crash involvement and injury severity (European Road Safety Observatory, 2006). Cognitive decline is more prevalent with aging and certain medical conditions (such as dementia), and the relation between cognitive decline and greater crash risk and unsafe driving performance is well documented (Carr and Ott, 2010).

Investigations have shown that neuropsychological tests measure abilities that are relevant for everyday/real-world functioning, including driving an automobile (Morgan and Heaton, 2009). For instance, two meta-analyses of the literature, one of persons with dementia (Reger et al., 2004) and one of older adults without a dementia diagnosis (Mathias and Lucas, 2009), provided support for the utility of tests of visuospatial abilities and visual attention, respectively, in predicting real-world driving ability. Tests of executive functions have also shown promise in predicting on-road (Ott et al., 2008; Whelihan et al., 2005) and crash involvement (Daigneault et al., 2002), although executive functioning has not been examined as extensively as other neuropsychological domains. Nevertheless, neuropsychological test scores do not correlate entirely with functional outcomes (Sadek and van Gorp, 2010), and current investigations attempt to conclude the most effective assessment methods for detecting driving deficits (Morgan and Heaton, 2009).

The Mini-Mental State Examination (MMSE; Folstein et al., 1975) is a cognitive screening test that has been widely used as a potential predictor of driving ability (cf., Reger et al., 2004), although it was not designed for this purpose. The investigations have produced conflicting results. A report from the American Medical Association (Iversen et al., 2010) concluded that an MMSE score lower than 25 points could be useful in identifying...
patients at increased risk of unsafe driving; however, due to conflicting data, any correlation with driving performance remains unclear.

Several serious limitations have been identified in respect to the MMSE as a means of screening for driving impairments. These limitations include the poor sensitivity for those scoring at the mid-range (Marcotte and Scott, 2009) and the weaker association with driving measures at the higher end of the scale (score greater than 26) (Fitten et al., 1995). In addition, the MMSE is based almost entirely on verbal cognitive function at the expense of non-dominant hemisphere skills and executive functions (Bak and Mioshi, 2007), which are the most related to various driving measures (cf., Amick et al., 2007; Daigneault et al., 2002; Lincoln et al., 2010; Reger et al., 2004; Whelihan et al., 2005). Moreover, the assessment of attention (verbal modality) is quite superficial. In fact, when additional measures of attention, visuospatial and executive functions are considered along with MMSE scores, the prediction accuracy of driving ability increases (Ott et al., 2005).

Due to these limitations, a new brief cognitive instrument to screen for driving-related cognitive impairment in the elderly is warranted. Such an instrument would be particularly useful for primary care physicians who are responsible for decisions regarding their patients’ fitness to drive. As driving fitness relies on a functional evaluation of multiple domains (cognitive, motor, perceptual, and psychiatric), no single test is sufficient to determine whether an individual is unsafe to drive (Rizzo, 2011). However, a cognitive screening test could be achieved to identify potential unsafe drivers (Wolfe and Clark, 2012) who require further neuropsychological (and, if available, on-road) assessment at a specialised centre.

Cognitive impairment is a recognised driving risk factor that could be related to deficits in attention, perception, memory and executive functions (Rizzo and Kellison, 2010), specific skills that are assessed by a recent cognitive screening test, the Addenbrooke’s Cognitive Examination Revised (ACE-R; Mioshi et al., 2006). Driving impairments might be expected in various populations, as normal aging, clinical neurologic disease (namely, Alzheimer’s Disease, Parkinson’s Disease, Multiple Sclerosis, Stroke) and psychiatric conditions (namely, Major Depression, Schizophrenia) that affect a person’s cognitive functioning and may impact driving ability (Marcotte and Scott, 2009). ACE-R research has examined these conditions (e.g., Alexopoulos et al., 2010; Dudas et al., 2005; Hamilton et al., 2009; Kwak et al., 2010; Mathew et al., 2011; Pendlebury et al., 2012; Reyes et al., 2009).

The goal of the current study is to examine the potential value of the ACE-R for predicting driving ability in older drivers. There are a number of reasons to analyse this brief cognitive test in relation to on-road driving performance. Firstly, it is a more accurate means of detecting early cognitive dysfunction than the MMSE, with a number of added items to increase its sensitivity to specific dementing diseases (e.g., Alexopoulos et al., 2010; Mioshi et al., 2006; Simões et al., 2011; Torralva et al., 2011; Yoshida et al., 2012) that might adversely affect driving abilities and increase crash risk (Iverson et al., 2010). Secondly, it comprises an MMSE score and sub-scores in five cognitive domains, which can provide qualitative and discriminative data regarding the patient’s driving-related cognitive profile. Thirdly, the ACE-R may provide some discriminating value for driving ability, as it includes tasks that are similar to those previously associated with driving measures (cf., De Raedt and Ponjaert-Kristoffersen, 2001; Johansson et al., 1996; Marottoli et al., 1994). Finally, the ACE-R is a brief, inexpensive, and easily administered tool. If it is validated to predict crash risk or driving ability, it would be appealing to primary care health settings. The current research examines whether this screening tool for cognitive impairment is useful in predicting driving impairment in older adults.

2. Methods

2.1. Participants

The current investigation was conducted at the Psychological Laboratory of Mobility and Land Transports Institute (MLTI), the unique state centre that performs psychological examinations of driving fitness in Portugal. The participants’ primary care physician referred them due to doubts about their driving-related cognitive abilities during a fitness to drive assessment. The inclusion criteria of the present study were as follows: aged 65 years or older; living independently in the community; without a requirement for specific adaptive equipment for driving; auditory and/or language functioning compatible with verbal comprehension and/or expression (pre-requisite for on-road and neuropsychological testing); visual acuity in the central field of view (far-sight, with regular correction) within licensing requirements in Portugal; and at least ten years of driving experience, including driving experience in the last month. During the data collection period (November 2009 to December 2010), the MLTI examined 83 drivers aged 65 years or more, of whom 23 did not partake in the research protocol: 16 persons did not meet the inclusion criteria (eight because of the presence of motor deficits and needs for specific adaptive equipment for driving, one did not fulfill the visual acuity licensing requirements, two person were aphasic with severe deficit in verbal communication), six were not scheduled in the days of the investigation, four were not assessed due to the absence of the driving examiner and two refused to participate. Of a total of 60 drivers who completed the research protocol, 50 participants aged 65–88 years (M = 73.1, SD = 7.0) were selected to obtain a study sample with equally sized groups defined by the dependent variable (safe/unsafe) and equally distributed by the reason for referral (aging, physical, neurological or psychiatric). This sampling strategy was designed to minimise the potential negative influence of unequally sized groups on the statistical analysis (discriminant function and classification cases) and to balance the range of reasons physicians requested a further psychological assessment in a specialised centre. The participants were referred for the following reasons: aging (n = 24); physical disorders (n = 4), including rheumatoid arthritis (n = 2), prosthesis on both feet (n = 1), congenital myopathy (n = 1); neurological disorders (n = 12), including Parkinson disease (n = 6), dementia of Alzheimer type (n = 2), stroke (n = 2), cerebral angioma (n = 1) and multiple sclerosis (n = 1); and psychiatric disorders (n = 10), including depression (n = 10), remote history of alcohol abuse (n = 3), residual schizophrenia (n = 1) and bipolar disorder (n = 1). Participants had a mean education of 6.18 years (SD = 4.3; min = 4; max = 20). As to the educational level, 64% completed primary school (fourth grade), 22% middle school (ninth grade), 2% high school (twelfth grade) and 12% were university graduates. The number of school years completed by the overall Portuguese population over 64 years is as follows: 36.2% are illiterate, 46.4% complete primary education, 9.4% middle school, 3.0% high school, and 5.0% university graduation (INE, 2011). Because basic literacy (to know how to read and write) is a requisite to obtain a Portuguese driving licence, illiterate people are not represented in the study sample.

2.2. Procedure

This study was approved by the Scientific Board of Faculty of Psychology and Educational Sciences of the University of Coimbra and the MLTI. All of the ethical principles of the Declaration of Helsinki were followed. Prior to participation, the subject received written information regarding the purpose of the study, and written informed consent was obtained. As the protocol assessment presented in this study is not found in a regular fitness to drive
assessment in Portugal (particularly an on-road driving test), subjects were assured that results were confidential and had no bearing on their driving license. Because 22% (40%) of the participants had an expired driving license due to the waiting list and time (2–3 months on average) between the physician’s referral and the date of assessment, an application form was also signed by the participant and an official representative of the MLTI to permit the participants to drive on a public road. The evaluations were performed individually in a single session, and the order between the cognitive and driving testing was counterbalanced in the sample. The research protocol included a multimodal assessment that collected data on demographic, clinical and driving history (through the driving process and an interview), visual function (far visual acuity, chromatic vision, and stereoscopic vision), cognitive functioning (ACE-R) and driving ability. Although vision was previously screened by physicians, visual function was examined prior to the written informed consent to ensure the selection of participants without visual problems that would affect the cognitive and driving outcomes.

The on-road driving test was carried out by an accredited professional examiner who was blind to the participants’ cognitive and personal data. The road test was conducted on a 10-km predetermined route in central Lisbon (Portugal) on the same day of the week and during an off-peak period (between 12 h and 14 h). The course began in a residential area as a familiarisation period with the experimental vehicle and driver examiner and proceeded to local and arterial roads. The route was outlined to ensure a consistently low to moderate level of traffic conditions and to enable the observation of driving behaviour in different driving situations (e.g., crossroads, roundabouts). The mean total time of the driving test was 50 min. The vehicle was fitted with manual transmission (the most familiar for all participants) and a dual-brake control system. The assessment was based on a driving measure, the Driving Behaviours Observation Grid (description in instruments section), that was completed after the road assessment. The examiner provided an initial standard instruction that omitted words such as “assessment” or “examination” to minimise potential anxiety that might produce a negative effect on driving performance. Standard verbal directional instructions were also provided during the course, and the examiner intervened physically (using the dual-brake or steering wheel) or verbally when required for safety. An interrater reliability study between the driving examiner and a research psychologist (in the front and rear passenger seat, respectively), who was previously trained on driving observation, was conducted on 24% of the study sample. To ensure that both observers were blind to the cognitive results, the interrater assessments only included observations of participants who had driven before the cognitive test.

2.3 Instruments

1. The ACE-R is a brief cognitive test that was developed to detect mild cognitive impairment and dementia. For the ACE-R original reference, see Minshi et al. (2006). It provides an MMSE score and five sub-scores that represent cognitive domains, namely attention and orientation (18 pts), memory (26 pts), letter and category verbal fluency (14 pts), language (26 pts) and visuospatial ability (16 pts). The total score is obtained by summing all of the sub-scores, ranging from 0 to 100, of which the MMSE score (30 pts) is part. Higher scores indicate better cognitive functioning. The Portuguese ACE-R version (Firmino et al., 2008) used in the present study was adapted with authorisation and advice from the authors of the original version.

2. The DBOG (Driving Behaviours Observation Grid) is a driving performance evaluation (unpublished) that was developed for a research project on the neuropsychological assessment of older drivers. Based on a literature review of studies with road tests for people with cognitive disabilities (Akinwuntan et al., 2003; De Raedt and Ponjaert-Kristoffersen, 2001; Fitten et al., 1995; Hunt et al., 1997; Janke and Eberhard, 1998; Radford, 2001), the present version was developed with the consensus of two independent experts on driving examinations. To develop a valid driving assessment for older people, the driving habits and difficulties that previous research found for this group (Ferreira et al., 2007) were also taken into account. The grid consists of 50 items or driving behaviours distributed across the following 10 categories or specific scores: (a) pre-driving checks (3 items), as mirrors, seat belt; (b) vehicle control (7 items), as accelerator, footbrake; (c) visual search (3 items), as for traffic signs, for potential hazards; (d) communication with other road users (2 items), as use of indicator, response to signals from road users; (e) driving on urban roads (15 items), as lateral position, awareness of pedestrians; (f) driving on the freeway (6 items), as speed choice, safety distance; (g) turning at junctions (4 items), as give right of way, trajectory; (h) driving on roundabouts (4 items), as position for lane choice, merging; (i) specific manoeuvres (3 items), as turn in the road, reverse parking; (j) other specific behaviours (3 items), as confidence, self-regulation according to deficits. Each driving item is scored on the following 3-point scale: 2 = correct driving behaviour, 1 = minor error (not affecting safety) and 0 = major error (severe enough to affect safety). A general score is obtained by summing all of the items, ranging from 0 to 100, in which a higher score indicates better driving performance. Based on the frequency and severity of observed behaviours during the driving task, the examiner indicates the final qualitative outcome of the participant’s driving ability as safe or unsafe. The unsafe classification is applied to drivers who commit several systematic errors with sufficient severity to affect safety (e.g., lane changes without looking, obstructs other cars at crossings, inappropriate brake/accelerator).

2.4. Statistical analyses

For demographic, medical, driving and cognitive data, a one-way analysis of variance or chi-squared ($\chi^2$) test was applied to compare the study groups. Discriminant analyses were performed to assess whether the predictor variables (independent variables) distinguished between safe and unsafe drivers (dependent variable) on a road test. Receiver operating characteristic (ROC) curves were plotted to compare the area under the curve between the ACE-R and MMSE scores. For the reliability study of the DBOG, the Pearson’s correlation coefficient was used to assess correlations between specific scores and the general score and the interrater concordance between all scores. Partial correlations were also computed between the driving measure scores and ACE-R outcomes. All analyses were conducted using SPSS Statistics (version 20.0, Chicago, IL) with the exception of the ROC analysis, which was performed in MedCalc (version 11.1, Mariakerke, Belgium).

3. Results

3.1. Comparison of non-cognitive variables between safe and unsafe drivers

As shown in Table 1, the safe and unsafe groups were similar on the demographic (age, gender and years of schooling), reason for referral, visual function (far visual acuity and stereoscopic vision) and driving variables (years of driving experience, professional driver experience and driving incidents in previous two years). The chromatic vision, not listed in the table, was found as normal in all participants.
Table 1
Demographic, medical, and driving characteristics of the study groups.

<table>
<thead>
<tr>
<th></th>
<th>Safe (n=25)</th>
<th>Unsafe (n=25)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (M, SD)</td>
<td>72.36, 7.65</td>
<td>73.92, 6.34</td>
<td>.436</td>
</tr>
<tr>
<td>Age groups (65–74</td>
<td>75–88)</td>
<td>16/9</td>
<td>12/13</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>23/2</td>
<td>21/4</td>
<td>.384</td>
</tr>
<tr>
<td>Years of schooling (M, SD)</td>
<td>7.24, 5.09</td>
<td>5.12, 3.17</td>
<td>.084</td>
</tr>
<tr>
<td>Medical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reason for referral (age/physical/neurological/psychiatric)</td>
<td>12/2/6/5</td>
<td>12/2/6/5</td>
<td>1.000</td>
</tr>
<tr>
<td>Far visual acuity (M, SD)</td>
<td>0.67, 0.15</td>
<td>0.60, 0.14</td>
<td>.100</td>
</tr>
<tr>
<td>Stereoscopic vision (absent-reduced/average-superior)</td>
<td>9/16</td>
<td>8/17</td>
<td>.771</td>
</tr>
<tr>
<td>Driving experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of driving experience (M, SD)</td>
<td>44.56, 11.85</td>
<td>39.60, 12.83</td>
<td>.162</td>
</tr>
<tr>
<td>Professional driver experience (yes/no)</td>
<td>6/19</td>
<td>4/21</td>
<td>.480</td>
</tr>
<tr>
<td>Driving incidents in previous two years (yes/no)</td>
<td>7/18</td>
<td>5/20</td>
<td>.508</td>
</tr>
</tbody>
</table>

Note: *Differences between groups were tested by ANOVA, with the exception of gender, reason for referral, professional driver experience and driving incidents in previous two years, which were evaluated by χ² with one degree of freedom.

Table 2
Mean, standard deviation and range values of DBOG scores between study groups.

<table>
<thead>
<tr>
<th></th>
<th>Safe (n=25)</th>
<th>Unsafe (n=25)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>General score</td>
<td>87.60</td>
<td>12.47</td>
<td>59–100</td>
</tr>
<tr>
<td>Specific scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-driving checks</td>
<td>5.60</td>
<td>0.87</td>
<td>3–6</td>
</tr>
<tr>
<td>Vehicle control</td>
<td>12.00</td>
<td>2.14</td>
<td>7–14</td>
</tr>
<tr>
<td>Visual search</td>
<td>5.32</td>
<td>1.15</td>
<td>3–6</td>
</tr>
<tr>
<td>Communication with other road users</td>
<td>3.24</td>
<td>0.88</td>
<td>2–4</td>
</tr>
<tr>
<td>Driving on urban roads</td>
<td>26.80</td>
<td>3.64</td>
<td>17–30</td>
</tr>
<tr>
<td>Driving on the freeway</td>
<td>10.16</td>
<td>2.15</td>
<td>6–12</td>
</tr>
<tr>
<td>Turning at junctions</td>
<td>7.16</td>
<td>1.34</td>
<td>4–8</td>
</tr>
<tr>
<td>Driving on roundabouts</td>
<td>6.68</td>
<td>1.49</td>
<td>4–8</td>
</tr>
<tr>
<td>Specific manoeuvres</td>
<td>5.36</td>
<td>0.91</td>
<td>3–6</td>
</tr>
<tr>
<td>Specific behaviours</td>
<td>5.16</td>
<td>1.21</td>
<td>3–6</td>
</tr>
</tbody>
</table>

Note: *Differences between groups were tested by ANOVA.

The results from the on-road assessment are presented in Table 2. All scores from the DBOG were significantly different between the safe and unsafe groups, demonstrating great consistency between the objective observations registered by the examiner and the final qualitative assessment of the participant’s driving ability. Comparing the range of values on the general score, a small overlap was observed between 59 and 66 points in which two drivers (both with 59 points) were effectively classified as safe.

3.2. Comparison of ACE-R outcomes between safe and unsafe drivers

There were significant differences between groups on ACE-R outcomes (Table 3), with the ACE-R total score presenting the greatest significance, despite the overlap of scores. The visuospatial, fluency and language sub-scores were the only discriminative sub-scores; however, the attention–orientation and memory domains were marginally close to significance.

3.3. Discriminant analyses with ACE-R variables to predict driving ability

Different predictor variables were analysed independently with an enter method to highlight their relative contribution in the discriminant model including leave-one-out classification (Table 4). All Wilks’ lambdas were significant at p < .05.

The first two models (1 and 2) analysed the contribution of an ACE-R score or MMSE score with selected non-cognitive variables (age, gender, years of schooling, far visual acuity and years of driving experience) in predicting on-road driving ability. The only

Table 3
Mean, standard deviation and range values of ACE-R outcomes between study groups.

<table>
<thead>
<tr>
<th></th>
<th>Safe (n=25)</th>
<th>Unsafe (n=25)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>Scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE-R</td>
<td>78.28</td>
<td>10.95</td>
<td>47–94</td>
</tr>
<tr>
<td>MMSE</td>
<td>26.48</td>
<td>2.80</td>
<td>18–30</td>
</tr>
<tr>
<td>Sub-scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention–orientation</td>
<td>16.96</td>
<td>1.67</td>
<td>12–18</td>
</tr>
<tr>
<td>Memory</td>
<td>16.96</td>
<td>5.78</td>
<td>3–26</td>
</tr>
<tr>
<td>Fluency</td>
<td>8.28</td>
<td>2.97</td>
<td>3–14</td>
</tr>
<tr>
<td>Language</td>
<td>22.44</td>
<td>3.00</td>
<td>15–26</td>
</tr>
<tr>
<td>Visuospatial</td>
<td>13.64</td>
<td>2.34</td>
<td>9–16</td>
</tr>
</tbody>
</table>

Note: ACE-R = Addenbrooke’s Cognitive Examination Revised; MMSE = Mini-Mental State Examination.

* Differences between groups are tested by ANOVA.
The discriminative variable in these models was the cognitive score, and the variable age had the lowest contribution to the classification functions. The next two analyses (models 3 and 4) without the non-cognitive variables demonstrated superior classification accuracy, with a lower percentage of classification errors. The accuracy of predictive models of driving ability with an ACE-R score or MMSE score decreased when non-cognitive variables were accounted for in the discriminant function analyses. Rather than contributing to the discrimination between groups, these variables seemed to produce random errors or noise.

The ACE-R score model demonstrated higher classification accuracy (70%; 35 out of 50 drivers correctly classified), with 72% sensitivity (18/25 unsafe correctly classified), 68% specificity (17/25 safe correctly classified), 69.2% positive predictive value and 70.8% negative predictive value. The MMSE score model exhibited minor sensitivity (56%; 14/25 unsafe correctly classified), high specificity (76%; 19/25 safe correctly classified), and a final accuracy of 66% (33/50 correctly classified), producing positive and negative predictive values of 70% and 63.3%, respectively.

Concerning the ACE-R cognitive domains (models 5 and 6), the visuospatial, fluency, and language sub-scores were the only sub-scores that significantly discriminated the two groups, with a higher classification accuracy (64%) than the predictive model with five sub-scores (62%), including the attention--orientation and memory domains. The final model (model 7) that included eight discriminating tasks between the study groups (retrograde memory, phonemic verbal fluency, naming 10 pictures, pictures comprehension, pentagons, cube, clock and delayed memory) correctly classified 58% of the drivers.

To understand the contribution of each ACE-R discriminating task in accurately predicting safe and unsafe drivers, discriminant analyses were made independently (Table 5). The phonemic verbal fluency task represented the most efficient model, with an accuracy of 74% (37/50 drivers correctly classified). The language tasks (naming ten pictures and pictures comprehension) and a visuospatial task (pentagons) were the most efficient in detecting safe drivers (specificity ranging from 80 to 92) but were less robust in detecting unsafe drivers (sensitivity ranging from 48 to 56). Other visuospatial tasks (copying a cube and clock drawing) and memory tasks (retrograde and delayed memory) also efficiently predicted safe and unsafe drivers.

### 3.4. Comparison of ROC curves between the ACE-R score and MMSE score

The area under the ROC curve of the ACE-R score was .75. This score is significantly higher than an AUC of .05 for no discrimination (z = 3.46; p < .001) and represents acceptable discriminative power (0.7 ≤ ROC < 0.8). The area under the ROC curve of the MMSE score was .66, which is also higher than an AUC of .05 for no discrimination (z = 2.05; p = .040), but with low discrimination (0.5 < ROC < 0.7). The AUC of the ACE-R score was considerably larger than the AUC of the MMSE score, although these differences did not reach significance (z = 1.88; p = .093) (Fig. 1).

### 3.5. Reliability of DBOG

The homogeneity of the DBOG was calculated with correlations between the DBOG specific scores and the DBOG general score, which ranged between 0.516 (p < .086) and 0.963 (p < .000). With the exception of “pre-driving checks”, all of the specific scores had a strong correlation (r > 0.7) with the general score.

In regards to the results of the DBOG’s interrater reliability study, the dichotomous qualitative outcome (safe/unsafe) of the participants’ driving ability was 100% identical between observers. The correlations between quantitative scores were: pre-driving checks (r = 0.06, p = .847), vehicle control (r = 0.79, p = .002), visual search (r = 0.89, p = .000), communication with other road users (r = 0.87, p = .000), driving on urban roads (r = 0.94, p = .000), driving on the freeway (r = 0.83, p = .001), turning at junctions (r = 0.93, p = .000), driving on roundabouts (r = 0.90, p = .000), specific manoeuvres (r = 0.98, p = .000), specific behaviours (r = 0.91, p = .000), and general score (r = 0.97, p = .000).

![Fig. 1. ROC curves of the ACE-R and MMSE scores as classification models.](image)
Table 6
Partial correlations between on-road driving and ACE-R outcomes (after controlling for visual acuity).

<table>
<thead>
<tr>
<th></th>
<th>ACE-R</th>
<th>MMSE</th>
<th>Attention/orientation</th>
<th>Memory</th>
<th>Fluency</th>
<th>Language</th>
<th>Visuospatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-driving checks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle controls</td>
<td>.44*</td>
<td>.33*</td>
<td>.39*</td>
<td>.32*</td>
<td>.40**</td>
<td>.33*</td>
<td>.33*</td>
</tr>
<tr>
<td>Visual search</td>
<td>.51**</td>
<td>.36</td>
<td>.34*</td>
<td>.40**</td>
<td>.46*</td>
<td>.40*</td>
<td>.37</td>
</tr>
<tr>
<td>Communication with other road users</td>
<td>.39</td>
<td>.32</td>
<td>.32*</td>
<td>.39*</td>
<td>.32*</td>
<td>.30*</td>
<td>.36</td>
</tr>
<tr>
<td>Driving on urban roads</td>
<td>.47**</td>
<td>.31</td>
<td>.32*</td>
<td>.37*</td>
<td>.44*</td>
<td>.34*</td>
<td>.35*</td>
</tr>
<tr>
<td>Driving on the freeway</td>
<td>.40**</td>
<td>.34</td>
<td>.35*</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>Turning at junctions</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving on roundabouts</td>
<td>.38*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific manoeuvres</td>
<td>.38*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific behaviours</td>
<td>.34*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General score</td>
<td>.45</td>
<td>.30</td>
<td>.31*</td>
<td>.34</td>
<td>.42**</td>
<td>.33*</td>
<td>.36</td>
</tr>
</tbody>
</table>

* p < .05.
** p < .01.

3.6. Associations between DBOG and ACE-R outcomes

To determine which road test variables from the DBOG were related to the ACE-R outcomes, partial correlations were computed with visual acuity as a covariate, as this variable was correlated with a number of driving scores (r ranged from .15 to .35, p < .05). The “pre-driving checks” was the only specific score without an association with cognitive outcomes. The ACE-R score and fluency and visuospatial sub-scores were associated with all other driving scores. Comparatively, the MMSE score and the attention–orientation sub-score were only related with three specific scores, vehicle control, visual search, and driving on urban roads, which were associated with all cognitive domains. The memory sub-score was related to “communication with other road users” and “driving on the freeway”, and language domain was associated with “driving on roundabouts”. All of the correlations were moderate (Table 6).

4. Discussion

The present study aimed to analyse the potential value of a brief cognitive test, the ACE-R, in predicting driving ability in older drivers. The study groups, drivers classified as safe or unsafe on a real driving test, had similar demographic and specific clinical variables. This allowed an approach to the neuropsychological functions assessed by the ACE-R with the greatest influence on driving performance. The discriminative variables between groups were the cognitive and driving results, which may represent the first sign that this brief cognitive test is useful in predicting on-road driving ability.

A major finding was the higher classification accuracy of the ACE-R total score in relation to the MMSE score (included in the ACE-R), particularly in detecting unsafe drivers. The overall accuracy of the ACE-R was 70%, with balanced sensitivity (72%) and specificity (68%). Comparatively, the overall accuracy of the MMSE score was not only lower (66%), but also presented a major trade-off between sensitivity (56%) and specificity (76%), producing a greater percentage of false negatives (unsafe incorrectly identified as safe drivers). Thus, the MMSE score was less useful for detecting unsafe drivers. These results suggest that MMSE scores may be inappropriate for the screening of driving impairments in a primary care sample of older people. The results from the ROC analyses demonstrated the following trend: the ACE-R had greater discriminant ability in the prediction of safe/unsafe drivers than the MMSE, although the differences did not reach statistical significance. The lack of statistical significance may be due to the lower power of the AUC comparison test given the limited sample size. The ACE-R score seems to predict driving ability better than the MMSE score. However, of note, the ACE-R score does not present sufficiently high levels of sensibility and specificity (conventionally both greater than 80%) to be used as a screening test to predict unsafe drivers.

The visuospatial, fluency and language sub-scores were the only sub-scores that were predictive of driving ability, while attention–orientation and memory domains were marginally close to statistical significance. Considerable research has shown that visuospatial ability measures are strongly correlated with driving ability (Reger et al., 2004). Concerning the visuospatial tasks included on the ACE-R, a poor pentagon design (copying of overlapping pentagons, from the MMSE) was predictive of self-reported driving difficulties (Gallo et al., 1999) and automobile crashes and moving violations (Marottoli et al., 1994), copying a cube was a strong predictor of crashes (Johansson et al., 1998), and clock drawing was predictive of real (De Raedt and Ponjaert-Kristoffersen, 2001) and simulated driving tasks (Freund et al., 2008) in active older drivers.

Previous research on driving abilities was less conclusive with respect to the utility of verbal fluency tasks, mainly conceptualised as measures of executive functions and language (Lezak et al., 2004; Strauss et al., 2006). For instance, a category fluency task was predictive of simulated driving performance in patients with Alzheimer’s disease (AD) (Rebok et al., 1994), and the Controlled Oral Word Association Test (COWA) was a predictor of on-road driving performance in patients with early AD (Uc et al., 2004). However, the Word Fluency Test (Duchek et al., 1998; Hunt et al., 1993) for phonemic association and Generative Naming (Wheihan et al., 2005) for semantic association were not related to on-road driving performance in patients with early cognitive decline or AD. Of note, the verbal fluency tasks were primarily studied in drivers with AD, and the present study found that the fluency sub-score of the ACE-R was useful in predicting driving ability among a clinical sample of older adults referred from primary care settings, with an unspecified (“age”) or different diagnosis.

The language domain of the ACE-R was also predictive of driving ability. Although language functions have been shown to be irrelevant for driving behaviour (e.g., Golper et al., 1980) and have not been recommended as part of a neuropsychological battery for driving (McKenna, 1998), it is reasonable to expect that language deficits may be indicative of a global cognitive decline that compromises driving performance. For instance, the Boston Naming Test has shown positive associations with on-road driving performance in dementia cases (Duchek et al., 1998; Hunt et al., 1993). Similarly, the current study found that the language sub-score and other cognitive sub-scores of the ACE-R were associated with the same driving skills.
Far from expected, the attention–orientation domain, included in the MMSE score, did not sufficiently discriminate the study groups. The role of attentional skills in driving is well documented (Ball et al., 1993; Parasuraman and Nestor, 1991). However, this cognitive domain of the ACE-R does not specifically assess attention, as more than half of this score is related to orientation items. To better understand whether the ACE-R is sufficiently robust in the assessment of attention, a further study must distinguish sub-scores for orientation and attention items.

The memory domain showed the least significant differences between safe and unsafe drivers. Studies have shown a relationship between memory deficits and the increased risk of driving errors, especially in patients with a severe cognitive impairment (Hunt et al., 1993). Memory tests are quite sensitive to overall cognitive impairment, as the presence of memory deficits is a key criterion in the diagnosis of mild cognitive impairment or dementia. For instance, more than one-quarter of the ACE-R score pertains to the memory domain. The present results demonstrated associations between the memory domain and driving performance; however, these associations were less striking than those for the fluency and visuospatial domains (associated with all driving scores). Thus, the ACE-R memory score was not the most face valid cognitive domain for on-road driving. Furthermore, the presence of memory impairments may interact with impairments in other cognitive domains relevant to driving. To isolate the effects of memory impairment in driving, Anderson et al. (2007) found that drivers with severe amnesia due to bilateral hippocampal lesions were not impaired in most aspects of driving performance (knowledge of driving rules, safety procedures, and road sign meaning were normal), but presented some difficulties in following route directions on the road. A study from Grace et al. (2005) that compared neuropsychological and driving performance in AD and Parkinson Disease patients found that severe memory impairment in the AD group (i.e., a cortical dementia with impaired learning, retrieval and recognition memory) did not relate to poor driving, but memory impairment in the unsafe Parkinson disease group (i.e., a subcortical dementia with impaired retrieval but relatively preserved recognition) reflected the influence of executive dysfunction on memory. These data corroborate the present findings that a neuropsychological test with a large executive component may be more robust for driving-related functions than memory tasks.

An abbreviated form of the ACE-R with eight discriminating tasks between safe and unsafe drivers was not successful (58% of accuracy) in predicting driving ability. The predictor model with the ACE-R score was more efficient, possibly because it covers more driving-related cognitive data or information and therefore prides a reasonable compromise between the dynamics of cognitive functioning and driving performance. From the discriminating tasks between the safe and unsafe groups, only the pentagon task was part of the MMSE score. In contrast, several discriminating tasks informed the ACE-R score, which contributed not only to a greater accuracy, but also to a more balanced model in the correct classification of drivers.

We aimed to understand the ability of each discriminative task from the ACE-R to correctly classify drivers. Retrace memory, naming ten pictures, pictures comprehension and pentagons were more tailored for detecting safe drivers (all with a specificity >75 but with low sensitivity). Phonemic verbal fluency was distinctive for detecting unsafe drivers, with high sensitivity and overall accuracy. Finally, cube, clock and delayed memory were balanced tasks for classifying safe and unsafe drivers (which could be fair for a predictor model), but alone were not sufficiently accurate (accuracy ranged between 64 and 66).

Of the eight discriminating items of the ACE-R, phonemic verbal fluency was the best classifier of driving ability, particularly for detecting unsafe older drivers. It was also associated with all driving-specific scores in the present study. This is a sensitive task for frontal lobe functions (Alvarez and Emory, 2006) and hence to frontotemporal dementia, whereas category fluency is a marker of semantic memory breakdown typical in AD (e.g., Mathuranath et al., 2000). Phonemic verbal fluency is considered valuable in detecting cognitive alterations in the aged given its stability throughout the ageing process (Brickman et al., 2005), which may offer diagnostic utility in the assessment of older drivers. However, it is important to note that one item from the ACE-R does not tell the whole “story” about driving ability, as visuospatial abilities and visual attention (modalities not examined in the ACE-R) are also crucial for safe driving (Reger et al., 2004; Mathias and Lucas, 2009). Due to the cognitive complexity and dynamics intrinsic to driving, a single task is insufficient for determining whether an individual should drive (Rizzolo, 2011).

Concerning the associations between the DBOG and ACE-R outcomes, the results suggested that global cognitive screening and specific cognitive domains of executive functions and visuospatial abilities were associated with all driving skills, stressing that deficits in these factors represent impaired driving ability. Comparatively, the remaining domains and the MMSE score did not appear (again) as the most face valid ACE-R outcomes for on-road driving performance.

To improve the accuracy of the ACE-R in predicting driving ability, a specific model or test ratio with a cut-off point could be studied in the future to differentiate safe and unsafe drivers. Such a study would be similar to the work developed by Mathuranath et al. (2000) to differentiate AD and frontotemporal dementia in the first ACE paper. However, the ACE-R must be studied with other complementary tests that have systematically predicted driving outcomes, such as measures of visual attention (e.g., Dot Counting, Trail Making Test – A and UFOV), executive functioning (e.g., Trail Making Test – B, Key Search and Rule Shift) and visuospatial abilities (e.g., Square Matrices Directions, Road Sign Recognition, WAIS-III Block Design) (cf., Ball et al., 1993; Ferreira et al., 2010; Lincoln et al., 2010; McKenna and Bell, 2007).

The study of predicting driving ability using neuropsychological tests implies a reliable and valid driving outcome. The consistent results between the observations registered by the examiner and the final qualitative outcome of the participants’ driving ability is suggestive of the objectivity of the DBOG (when used by a driving specialist) to assess driving performance in older drivers. However, a slight overlap in the general score of the study groups raises the question of whether scores should be summed to an overall score or classifications should be based on a final qualitative outcome. Both criteria were considered in the present study. The overall impression of the examiner was based on the DBOG as an operationalised measure of driving behaviour, without neglecting the examiner’s experience and expertise to subjectively classify a driver as safe or unsafe. Of note, an overall score may be more affected by minor errors (not affecting safety) or “bad habits” (cf., Selander et al., 2011), which should be distinguished and not considered by a trained observer for an unsafe driving classification. The results from the reliability of the DBOG also showed that this driving measure was objective and stable when used by different trained observers. However, one specific score, “pre-driving checks” (adjustment of seat, mirrors and putting on the seat belt), was not sufficiently consistent during observations. We suggest that these preliminary driving behaviours may deserve careful attention in a further investigation with this instrument.

The results of the present study provide support for the potential utility of the ACE-R as a screening instrument for older drivers. We argue that it might be simple and less time consuming for primary care physicians to administer a brief screening tool for cognitive impairment in the first stage of the fitness to drive assessment process for older people. This suggestion is particularly relevant
considering the diversity of neuropathology associated with aging, which is not consistently diagnosed or well-documented in primary health care settings and referrals (cf., McKenna and Bell, 2007). A screening tool must aid clinicians in determining whether a person should be referred for a more detailed evaluation (Wolfe and Clark, 2012). Thus, the prediction of unsafe drivers due to cognitive impairment should be a second stage of the assessment process that is conducted in specialised centres. Such an evaluation must focus on in-depth analyses of neuropsychological functions with validated tests that relate to (desirably) various driving measures, such as driving performance and accident involvement. In the specialised centres of many countries, a road test is considered as an additional and complementary source of evidence of driving safety.

Some limitations of the current study should be noted. First, the sample size is small; therefore, the present findings require further independent validation with a larger sample. In Portugal, a road driving assessment is not provided in a fitness to drive assessment, which leads to constraints in developing a full evaluation that includes on-road driving. Notwithstanding the use of the “gold standard” and validity of the external criterion, it is possible that some bias may have occurred in the classification of drivers. The presence of undiagnosed physical disorders is common in older people, and some explicit motor dysfunctions may slightly harm driving performance. Despite these limitations, no previous studies analysed the relationship between the ACE-R and real-world driving performance. The current findings support the potential value of the ACE-R in predicting driving ability, as it is more effective than an MMSE score. Cognitive domains of executive functions and visuospatial abilities were particularly related to driving performance. This emerging cognitive test may warrant additional study in relation to driving measures, but it seems as though the ACE-R could be included in revised psychological assessment protocols for determining driving fitness in older adults.

Conflicts of interest

All authors of this study declare that there are no conflicts of interest.

Acknowledgements

We would like to thank the partnership between the Faculty of Psychology and Educational Sciences from the University of Coimbra and the Mobility and Land Transports Institute for providing conditions to make this study possible. We also thank Sérgio Marques from the Automobile Club of Portugal for the time and expertise in carrying out the driving assessments.

The authors would also like to thank the anonymous referees for their comments and suggestions.

This work was supported by the Fundação para a Ciência e Tecnologia [Portuguese Foundation for Science and Technology] through of a PhD fellowship (SFRH/BD/27255/2006) awarded to the first author.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jaa.2012.03.036.

References


Rizzo, M., 2011. Impaired driving from medical conditions: a 70-year-old man trying to decide if he should continue driving. The Journal of the American Medical 305 (10), 1018–1026.


