Accident Analysis and Prevention 43 (2011) 1348–1354

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Older drivers: On-road and off-road test results

Helena Selander[a,*], Hoe C. Lee[b], Kurt Johansson[c], Torbjörn Falkmer[b,d,e,f]

[a] School of Health Sciences, Jönköping University, Sweden
[b] School of Occupational Therapy and Social Work, Curtin Health Innovation Research Institute, Curtin University, Perth, WA, Australia
[c] Department of Neurobiology, Care Science and Society, Karolinska Institutet, Stockholm, Sweden
[d] Rehabilitation Medicine, Faculty of Health Sciences, Linköping University, Sweden
[e] School of Health Sciences, Jönköping University, Sweden
[f] School of Occupational Therapy, La Trobe University, Melbourne, VIC, Australia

ARTICLE INFO

Article history:
Received 27 August 2010
Received in revised form 4 February 2011
Accepted 6 February 2011

Keywords:
Fitness to drive
Off-road assessment
Older driver
On-road assessment

ABSTRACT

Eighty-five volunteer drivers, 65–85 years old, without cognitive impairments impacting on their driving were examined, in order to investigate driving errors characteristic for older drivers. In addition, any relationships between cognitive off-road and on-road tests results, the latter being the gold standard, were identified. Performance measurements included Trail Making Test (TMT), Nordic Stroke Driver Screening Assessment (NorSDSA), Useful Field of View (UFOV), self-rating driving performance and the two on-road protocols P-Drive and ROA. Some of the older drivers displayed questionable driving behaviour. In total, 21% of the participants failed the on-road assessment. Some of the specific errors were more serious than others. The most common driving errors embraced speed; exceeding the speed limit or not controlling the speed. Correlations with the P-Drive protocol were established for NorSDSA total score (weak), UFOV subtest 2 (weak), and UFOV subtest 3 (moderate). Correlations with the ROA protocol were established for UFOV subtest 2 (weak) and UFOV subtest 3 (weak). P-Drive and self ratings correlated weakly, whereas no correlation between self ratings and the ROA protocol was found. The results suggest that specific problems or errors seen in an older person's driving can actually be "normal driving behaviours".

1. Introduction

The total number of older drivers on the roads is rapidly increasing (Johansson et al., 1996) and at the same time, the traffic environment has gradually become more challenging (Evans, 2004).

Medical conditions, for example dementia or stroke, may compromise driving and therefore impact on a person's fitness to drive, i.e., the medical and functional requirements for driving. Cognitive assessments may contribute to determine a client's fitness to drive, but there are no specified guidelines stipulating which assessment tools to use, nor any defined cut-off scores (Swedish Transport Agency, 2010).

Several approaches have been taken to identify unsafe drivers with cognitive impairments (Brown et al., 2005; Mazer et al., 1998; Schanke and Sundet, 2000). Cognitive tests do provide valuable information about a client's specific abilities regarding fitness-to-drive, e.g., divided attention. Cognitive off-road tests that are used to make recommendations about the driving license status of a client with cognitive impairments would thus be expected to correlate with on-road test results. The relationships between the results of cognitive off-road tests and driving performance are, however, inconclusive (Akinwuntan et al., 2002; Marottoli et al., 1998; Stutts et al., 1998). Commonly, their criterion-related validity is poor. However, the more the off-road tests simulate driving i.e., the higher the face validity, the more clinically relevant they are considered to be (Anstey et al., 2005). Although most cognitive tests do not define cut-off scores to determine whether the client is a safe driver (Dobbs et al., 1998; Reger et al., 2004; Selander et al., 2010), they do provide the assessor with information about the client's cognitive functions that may have to be further assessed during an on-road assessment (Unsworth et al., 2005).

On-road assessment is the universal criterion measurement of driving competency or driving performance (Kay et al., 2008; Odenheimer et al., 1994). However, the on-road assessment has also been criticized for low validity and reliability (Fox et al., 1998; Galski et al., 2000; Odenheimer et al., 1994). Ideally, on-road assessments should be carried out on a fixed route and assess the driving performance based on standardized observations and scoring procedures (Di Stefano and Macdonald, 2003; Fox et al., 1998; Withaar et al., 2000). To use the same car during an on-road assessment further enhances standardization (Fox et al., 1998).

Older driver related research has mostly been conducted on impaired older drivers, without investigating how their healthy counterparts perform on the same outcome variables. The

* Corresponding author. Tel.: +46 31 454536; fax: +46 31 454490.
E-mail addresses: helena.selander@hhj.hj.se, helena@mobilitetscenter.se (H. Selander).

0001-4575/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2011.02.007

researchers have adopted either the number of errors made in specific traffic scenarios during the on-road assessment or the overall performance of participants as outcome measurements in their studies (Akinwuntan et al., 2002; Fox et al., 1997; Schanke and Sundet, 2000). However, it has not been clear whether the errors are due to impairments or they simply are developed throughout lifelong driving, contributing to possible sub-standard performances of older drivers (Dobbs et al., 1998). There is thus a need for valid evaluation methods of driving performance in this group. However, despite being the gold standard, the on-road assessment itself and its role in the decision of pass or fail have not been thoroughly studied. Hence, it is important to identify older drivers' characteristic driving errors among experienced and fit to drive persons, in order to improve on-road assessments for clients with cognitive impairments and declining competences. By exposing fit to drive older drivers to the same on-road and off-road tests that clients with cognitive impairments take, "normal" driving behaviours on a standardized on-road assessment can be revealed. In addition, their performances on cognitive tests may provide reference values. Our primary objectives were thus to investigate what types of driving errors are characteristic for older drivers without cognitive impairments affecting their fitness to drive, and to identify any relationships between off-road and on-road tests results.

2. Methods

2.1. Participants

The participants were recruited from the Vehicle Registration Office in Sweden. From a list, 394 randomly selected 65+ old individuals with a registered vehicle were approached by mail. Of those, 157 did not reply (non-responders) and 110 were not interested to take part in the investigation (42% men, $N = 46$), while 127 were interested to participate. Of these 127, 98 were selected on a first come-first serve basis. No data were available on the 29 who were not selected apart from their gender (59% men, $N = 17$) and that they were 65+. The 98 participants were interviewed by the first author. Eight persons did not fulfil necessary physical and cognitive fit-to-drive requirements for safe driving according to the Swedish Transport Agency guidelines (Swedish Transport Agency, 2010) and were excluded. For example, visual problems, stroke or dementia became exclusion criteria. Furthermore, an inclusion criterion was that they should still be active drivers (minimum 3000 km/year). When they were interviewed, also the presence of potential other medical conditions, e.g., heart disease, hypertension and diabetes, was checked for on a self-report basis. The remaining 90 fulfilled all inclusion criteria and agreed to participate in the study. However, five dropped out for various reasons. Hence, a total, 85 participated in the study. The participants' mean age was 72.0 (SD = 5.3; ranging from 65 to 85), 53% being male. There was no significant age difference between the sexes, viz. for males the mean age was 72.7 (SD = 5.6) and for the 40 females it was 71.2 (SD = 4.8) years. Similarly, there was no significant difference with respect to the number of years in school between the sexes, varying from 6 to 20 years (female mean = 10.9, SD = 3.3; male mean = 11.4, SD = 3.5). Of the 85 participants, 41% reported some sort of medical condition that supposedly did not affect their fitness to drive. Some reported multiple conditions, e.g., hypertension ($N = 25$), heart conditions ($N = 15$) and diabetes ($N = 5$). This group of 35 participants is henceforth labelled as DMC+ (Drivers *with* Medical Conditions). Consequently, the remaining 50 are labelled as DMC- (Drivers *without* Medical Conditions). There was no significant age difference between the two DMC-groups, for DMC+ the mean age was 73.0 (SD = 5.3) years and for the DMC- the mean age was 71.3 (SD = 5.3) years.

2.2. Procedure

The present study was approved by a local Ethical Committee in Stockholm, Sweden in accordance with Swedish law. Prior to their participation the participants received written information about the study purpose and that participation would not impinge on their driving licence. The data were collected at a driving assessment unit in Stockholm, Sweden. To guarantee that the participants fulfilled the requirements for vision, they had to undergo an examination, which included visual acuity and visual fields. They also underwent a cognitive screening with the tests TMT A & B (Trail Making Test), NorSDSA (Nordic Stroke Driver Screening Assessment), and UFOV (Useful Field of View). However, one participant did not complete the TMT B test and four participants did not complete the UFOV test. All tests are further described below. After these cognitive tests were completed, the participants filled in a self rating driver performance scale.

The driving took approximately 60 min on a fixed route (39.7 km) on public roads in a suburban district. The route is used for on-road assessments by the driving assessment unit. An occupational therapist (OT) observed the quality of the driver's behaviour, e.g., following instructions, planning, manoeuvring, lane positioning, obeying traffic rules, interaction with other road users and the attention using two scoring sheets further presented below. After each test, the OT decided whether participants passed or failed the test. The final pass/fail decision was the result of an overall impression of the participants' behaviour, based on the frequencies and severity of observed problems. The OT was blinded to their results from the cognitive tests, and whether they were drivers with or without medical conditions. A driving instructor had the safety responsibility through dual controls and gave instructions, i.e., directions to follow throughout the route. The driving instructor sat in the front passenger seat and the OT in the back seat to the right (right hand driving). Sixty-six chose to drive a manual gear shifted car, whereas the remaining 19 chose an automatic gear shifted car.

2.3. Instruments

1. The TMT (The Trail Making Test) is a cognitive test that measures visual search and sequencing, information processing speed, divided attention and flexibility (Reitan, 1986). The test consists of two subtests, A & B, completed in the shortest possible time and scored in seconds to completion.
2. The SDSA (Stroke Driver Screening Assessment) is a set of cognitive tests developed to evaluate fitness-to-drive in stroke clients (Nouri and Lincoln, 1992). The Nordic version of the SDSA, NorSDSA, was used in the present study. It has been validated with 97 stroke clients from Sweden and Norway (Lundberg et al., 2003). NorSDSA comprises of four sub tests providing six sub scores: viz. Dot Cancellation: measured in seconds to completion (maximum 15 min) and number of errors: Directions: maximum 32 points, Compass: maximum 32 points, and Road Sign Recognition, scored 0–12 after 3 and 5 min. Higher scores on Directions, Compass and Road Sign Recognition are considered better than lower. Based on results from Dot Cancellation (time and errors), Compass and Road Sign Recognition (3 min), the test provides a weighted overall score. SDSA provides clinically useful information regarding cognitive functions that are important for driving, e.g., focused and sustained attention, cognitive processing speed and the ability to attend to two visual dimensions at the same time.
3. The UFOV (Useful Field of View) is a PC-based visual and cognitive test that includes three sub tests measured in milliseconds. The first subtest measures processing speed only, while the second measures processing speed for a divided attention task and the third processing speed for a selective attention task (Edwards

et al., 2005). The targets on the PC-screen are presented for 16–500 ms. Lower scores (ms) indicate better performance.
4. A self-rating of the participants' driving performance was obtained on a linear scale, from 1 (poor) to 10 (excellent), where 5 was benchmarked as "the average driver".
5. P-Drive (Performance Analysis of Driving Ability) is an assessment tool for measuring driving ability (Patomella et al., 2006). It was developed for stroke clients to use in a driving simulator, but has shown to be valid and reliable also for assessing driving ability on-road (Patomella et al., 2010). Furthermore, P-Drive has proven to be a valid assessment protocol also for clients with dementia or mild cognitive impairment (Patomella et al., 2010). P-Drive was not developed on the assumption that the scores should be summed, in order to reach an overall score. The protocol consists of 27 items or driving actions, further described in Appendix A, e.g., steering, position on the road, attention to the left, heeding information sign, etc. Each driving item is scored on a scale from 1 to 4, where 4 = competent driving ability, 3 = questionable, 2 = problematic, 1 = incompetent driving ability.
6. ROA (Ryd On-road Assessment) is developed (unpublished) and clinically in use at the driving assessment unit in Stockholm. The scoring sheet comprises seven categories, i.e., speed, position, instruction, attention, indicator, traffic rules and manoeuvring, with 34 specific items. The assessment is further described in Appendix A. Errors made are graded on a 0–2 scale, where 0 implies normal driving behaviour, 1 indicates minor error, while 2 indicates a considerable risk-taking behaviour. The scores may be summed, in order to reach an overall score.

2.4. Statistical analyses

Statistical analyses were performed by SPSS® (version 17.0). All variables were tested for normal distribution with the use of the Kolmogorov Smirnov test. NorSDSA Dot cancellation (time and errors), NorSDSA Directions, UFOV subtest 1 and 2 and self ratings did not meet this requirement. Log-transformation of these data was not performed. χ^2 tests, Spearman's rank correlation tests, Mann–Whitney *U* tests and Student's *t*-tests were used with the α-level set at .05. Since the NorSDSA and UFOV tests are developed to measure one construct each, Bonferroni corrections of the α-levels for multiple testing, in order to avoid making a type I errors, were applied to the six subsections of NorSDSA (α = .008) and the three subsections of UFOV (α = .016).

3. Results

All 85 participants completed the on-road assessment with an overall pass rate of 79%, as shown in Table 1. The proportions of men and women were equally distributed between the two groups, as were those with and without any medical conditions. The choice of transmission did not differ between the two groups.

As further shown in Table 1, those who failed the on-road test were on average older and rated themselves as less good drivers than those who passed. However, both groups rated themselves better than the average driver, i.e., 5. As a matter of fact, 69% of all participants considered themselves to be better than the average driver. Another 26% rated themselves as an average driver, whereas only 5% considered themselves as worse than the average driver. With respect to sex, 47% of the women rated themselves as good as the average driver or worse, whereas among men only 18% rated themselves the same way. Furthermore, in the fail group, nearly half of them (47%) thought they were better than the average driver.

Self ratings and the on-road scores protocols were tested for possible correlations. While P-Drive and self ratings correlated weakly (rho = .24, *p* = .046), no correlation between self ratings and ROA was found. Correlation analyses were further made between all off-road scores (TMT, NorSDSA and UFOV) and self ratings, but no correlations were found.

The on-road and off-road cognitive tests results are presented in Table 2. NorSDSA Road sign recognition 5 min and UFOV subtest 3 results were both significantly better in the pass group than in the fail group. With respect to on-road tests, the pass groups presented with better results than the fail group in both protocols.

The cognitive off road tests, including the subtests, were correlated with the two on-road protocols. Correlations with the P-Drive protocol were significant for NorSDSA compass (rho = .28), NorSDSA Road sign recognition 3 min (rho = .26), NorSDSA Road sign recognition 5 min (rho = .31), NorSDSA total score (rho = .22), UFOV subtest 2 (rho = −.29) and UFOV subtest 3 (rho = −.45). The correlations with the ROA protocol were significant for UFOV subtest

Table 1
Demographic and self rating data for the failed and passed on-road groups.

	Failed on-road (*N* = 18)	Passed on-road (*N* = 67)	Test and *p*-values
Age (mean, SD)	75.9, 6.3	71.0, 4.5	$t = 3.1, p = .005^*$
Sex (male/female)	8/10	37/30	$\chi2 = 0.7, p = .44$
Medical conditions	DMC-/7 DMC+	39 DMC-/28 DMC+	$\chi^2 = 0.05, p = 1.00$
Manual transmission	16	50	$\chi2 = 1.66, p = .34$
Self rating (mean, SD)	5.94, 1.4	6.60, 1.3	$z = -2.13, p = .033^*$

* Significant differences between the two groups.

Table 2
The off-road cognitive test results and on-road protocol results for the failed and passed on-road groups.

	Failed on-road (*N* = 18)	Passed on-road (*N* = 67)	Test and *p*-values
TMT A, mean time (s)	42.3, 95% CI = 36.5–48.2	37.3, 95% CI = 34.2–40.5	$t = 1.47, p = .15$
TMT B[a], mean time (s)	98.1, 95% CI = 70.4–125.7	87.6, 95% CI = 78.2–96.9	$t = .92, p = .36$
NorSDSA, Dot cancellation, median time (s)	396, [365, 518]	389, [347, 425]	$z = -1.28, p = .20$
Median number of errors	7, [2, 12]	7, [3, 13]	$z = -.16, p = .88$
NorSDSA, Directions, median	32, [28, 32]	32, [32, 32]	$z = -2.05, p = .04$
NorSDSA, Compass, mean	23.3, 95% CI = 20.7–25.9	26.2, 95% CI = 24.7–27.	$t = -1.89, p = .06$
NorSDSA, Road sign recognition, 3 min, mean	5.5, 95% CI = 4.7–6.3	6.7, 95% CI = 6.1–7.3	$t = -1.96, p = .05$
NorSDSA, Road sign recognition, 5 min, mean	6.9, 95% CI = 6.2–7.7	8.5, 95% CI = 7.9–9.0	$t = -2.61, p = .002^*$
NorSDSA, total score mean	1.47, 95% CI = 1.06–1.89	1.93, 95% CI = 1.65–2.21	$t = -1.58, p = .12$
UFOV[a], subtest 1, median	16, [16, 23]	16, [16, 20]	$z = -1.26, p = .21$
UFOV[a], subtest 2, median	53, [26, 126]	26, [16, 70]	$z = -2.30, p = .02$
UFOV[a], subtest 3, mean	340, 95% CI = 280–400	195, 95% CI = 169–220	$t = 5.14, p < .001^*$
P-Drive, median	78, [74, 81]	94, [90, 96]	$z = -6.09, p < .001^*$
ROA, median	86, [62, 111]	49, [35, 59]	$z = -4.82, p < .001^*$

* Significant differences between the two groups. Bonferroni corrections: NorSDSA subsections – α = .008 and UFOV subsections – α = .016. Brackets denotes [25th percentile, 75th percentile].
[a] *Note*: TMT B; missing data for one participant in the fail group, UFOV; missing data for four participants in the pass group.

2 (rho = .27) and UFOV subtest 3 (rho = .36). All correlations are presented in Table 3.

For both groups, six P-Drive items had a mean score of ≤3, viz. heeding signs, attending to the left, attending to the right, follow speed regulation, giving right of way, and speed control high pace. Fig. 1 shows these items for both the fail and the pass groups.

In seven items within the ROA-protocol the most frequent errors were made, viz. obeying speed limit, change gear, do not use indicator, blind spot to the left, attention to the left, attention to the right, speed – too fast. Fig. 2 shows these items for both the fail and the pass groups. When further scrutinising the overall ROA results, age was found to correlate negatively with them (rho = .348, p = .001). However, this correlation was weak.

Although no valid pass/fail cut-off values were given by the originator of the two off-road tests that showed different outcomes between the fail and the pass group, i.e., NorSDSA Road sign recognition 5 min and UFOV subtest 3, we elaborated with different tentative cut-off values, based on the mean values presented in Table 2, to establish the optimal cut-off value for each of them, given our results. The same analysis was applied on the two on-road protocols, P-Drive and ROA, that also displayed different outcomes between the two groups.

Table 3
Correlations between on-road protocols and off-road tests.

	P-Drive	ROA
TMT A	rho = −.17, p = .13	rho = .15, p = .18
TMT B	rho = −.20, p = .68	rho = .17, p = .14
NorSDSA, Dot cancellation (time)	rho = −.16, p = .14	rho = .16, p = .16
NorSDSA, Dot cancellation (errors)	rho = .03, p = .80	rho = .14, p = .22
NorSDSA, Directions	rho = .19, p = .08	rho = .15, p = .17
NorSDSA, Compass	rho = .28, p = .01*	rho = −.17, p = .13
NorSDSA, Road sign recognition, 3 min	rho = .26, p = .02*	rho = −.08, p = .49
NorSDSA, Road sign recognition, 5 min	rho = .31, p = .004*	rho = −.17, p = .14
NorSDSA, total score	rho = .22, p = .047*	rho = −.15, p = .18
UFOV, subtest1	rho = −.15, p = .17	rho = .20, p = .07
UFOV, subtest 2	rho = −.29, p = .008*	rho = .27, p = .02*
UFOV, subtest 3	rho = −.45, p < .001*	rho = .36, p < .001*

* Significant correlations between the protocols and off-road tests.

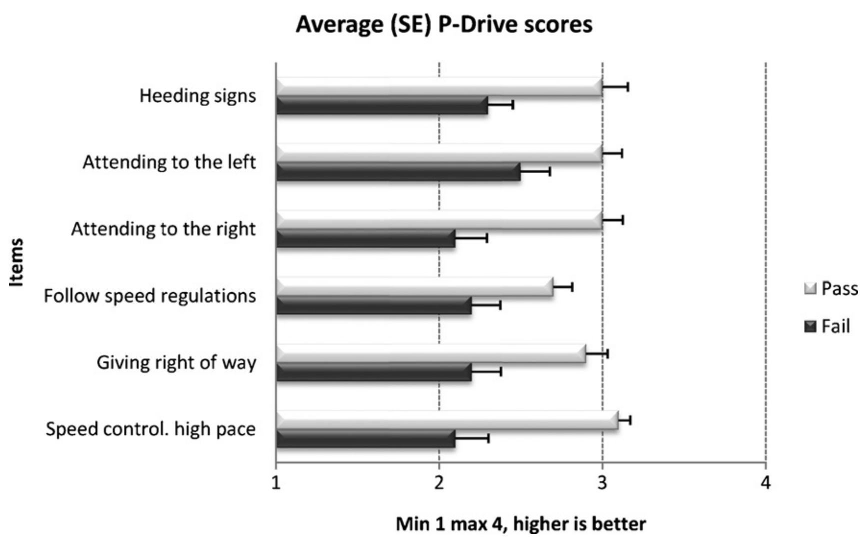

Fig. 1. Mean score of P-Drive items (1 = incompetent driving ability, 4 = competent driving ability). Error bars indicate standard error of the mean (SE).

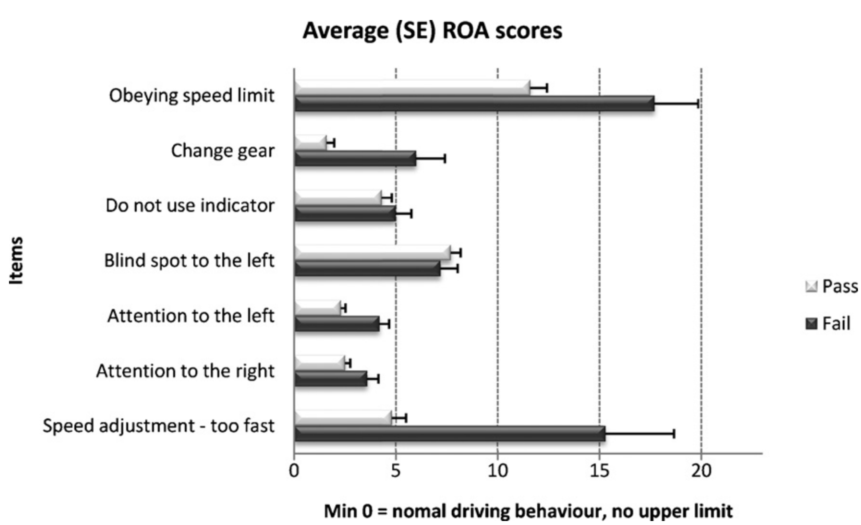

Fig. 2. Mean score of ROA items (0 = normal driving performance, 1 = minor error and 2 = considerable risk-taking performance. The errors are summed up.). Error bars indicate standard error of the mean (SE). "Obeying speed limit" refers to actual speed limit, whereas "Speed adjustment – too fast" refers to controlling and adjustment of speed.

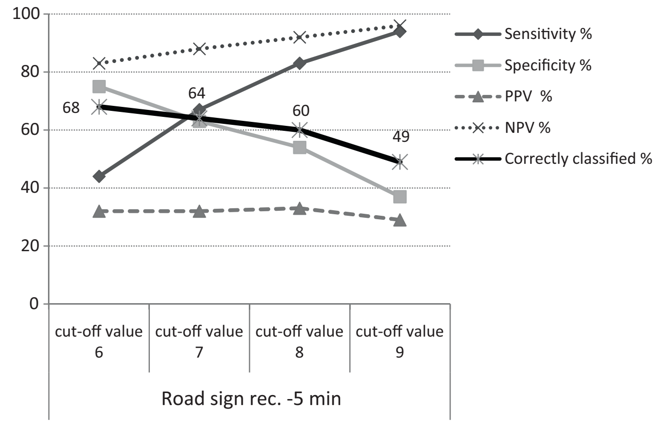

Fig. 3. Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and Correctly classified, all presented as percentage-values, for different cut-off values for NorSDSA Road sign recognition 5 min. The actual percentage figures for correctly classified are displayed.

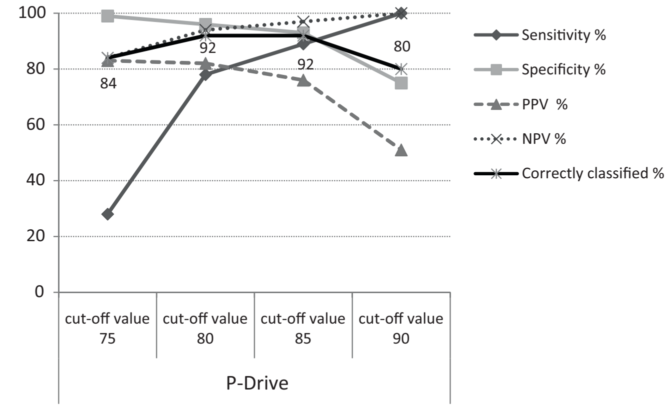

Fig. 5. Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and Correctly classified, all presented as percentage-values, for different cut-off values for P-Drive. The actual percentage figures for correctly classified are displayed.

As shown in Fig. 3, the combined sensitivity/specificity curve crosses at 7, suggesting that as the optimal cut-off value for the NorSDSA Road sign recognition 5 min. Interestingly, the percentage of correctly classified was higher for a cut-off value at 6, but then the sensitivity becomes low.

As shown in Fig. 4, the combined sensitivity/specificity curve crosses at 250, suggesting that as the optimal cut-off value for the UFOV subtest 3. The percentage of correctly classified was higher for a cut-off value at 350, but then the sensitivity becomes low.

As shown in Fig. 5, the combined sensitivity/specificity curve crosses at 85, suggesting that as the optimal cut-off value for the P-Drive protocol. The percentage of correctly classified peaks at the cut-off value 85.

As shown in Fig. 6, the combined sensitivity/specificity curve crosses at 60, suggesting that as the optimal cut-off value for the ROA protocol. The percentage of correctly classified was higher for a cut-off value at 70, but then the sensitivity becomes low.

4. Discussion

To use standardized methods for assessing fitness to drive is important for valid and reliable outcomes. On-road evaluations are often used in different ways, e.g., with different protocols, evaluators and scoring. However, on-road assessments still have high face validity and are generally seen as the gold standard despite such weaknesses (Odenheimer et al., 1994). In the present study, the same standardized route, evaluator and scoring protocol routines as used in the driving assessment clinic were applied. The results show that some of the older drivers, without cognitive impairments impacting on their driving, displayed questionable driving behaviour. These findings indicate that we should be aware of the fact that some specific problems or errors can actually be “normal driving behaviours” and not due to cognitive impairments in the driver. When the outcome of an on-road assessment has consequences for the driver’s license status, this awareness is important for the assessor to keep in mind when assessing that driver’s particular fitness to drive.

Despite the fact that the participants in our study reported themselves as fit to drive, every fifth failed the on-road assessment. Other researchers have also found a great number of unimpaired drivers who failed the on-road test (Dobbs et al., 1998; Lundqvist et al., 2000; Soderstrom et al., 2006). In the Lundqvist et al. study 20% of their controls failed, whereas 50% of the controls failed in a study by Soderstrom and colleagues (Lundqvist et al., 2000; Soderstrom et al., 2006). Those who failed in our study most frequently had problems with controlling speed and obeying speed limits. This phenomenon is surprising, since older drivers generally are believed to drive more slowly and self-regulate their driving (Persson, 1993), and it should be subjected to future research. The participants frequently exceeded the speed limit on the route, which may represent “normal driving behaviour”, for example on

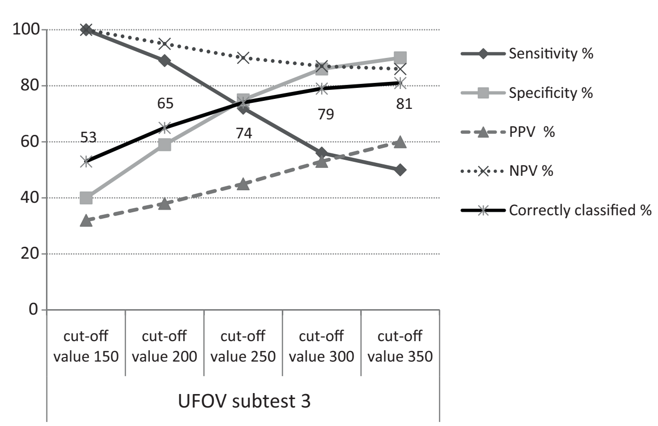

Fig. 4. Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and Correctly classified, all presented as percentage-values, for different cut-off values for UFOV subtest 3. The actual percentage figures for correctly classified are displayed.

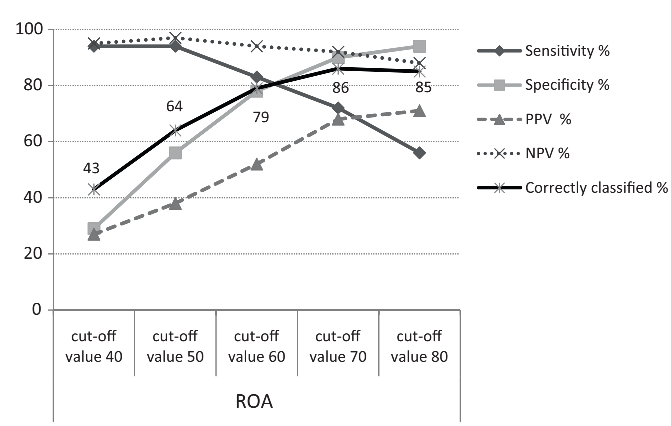

Fig. 6. Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and Correctly classified, all presented as percentage-values, for different cut-off values for ROA. The actual percentage figures for correctly classified are displayed.

highways. However, older drivers avoid challenging situations (Ball et al., 1998), but they often drive in high risk environments in urban areas (Di Stefano and Macdonald, 2003). In these areas they are dependent on, for example, controlling and safely adjusting their speed and attention to other road users. Our participants had difficulties maintaining the speed or controlling it. They often drove too fast for a specific traffic situation or condition. That is a more serious error and risk-taking behaviour.

Furthermore, the speeding problems would possibly also have affected the manoeuvring skills, e.g., gear changing. Although the on-road outcome was not affected by transmission type (manual vs. automatic gear shifting), changing gear was the fourth most common driving error in the ROA protocol, particularly for the fail-group. However, some of these errors are probably not dangerous for others, like gear changing and using indicators, and can represent normal driving behaviour or "bad habits". Gear changing is often seen as an automatic task and should not be a problem for an experienced driver (Shinar et al., 1998), but for the older driver the execution of motor skills can become less automatic than they used to (Brouwer and Ponds, 1994), which could affect gear changing.

Lack of attention is a serious error when driving demands interaction with other road users. Errors concerning attention were observed and scored frequently, both to the right and to the left. Supported by previous research (Di Stefano and Macdonald, 2003; Kay et al., 2008; Wood et al., 2009), failure to check the blind spot was also a common driving error. A question that arises from our findings is: do the speeding problems cause some of the other driving errors? For an older driver a higher speed would make it more difficult to maintain the visual attention needed for safe driving. Declines in visual attention have previously been associated with an increased crash risk in older drivers (Ball et al., 1993; Owsley et al., 1998). Our results displayed some driving behaviours in older drivers that were questionable and may cause problems. Some of the specific errors were more serious than others. Thus, there may be a need for re-training of older drivers or to find other solutions to help them maintain safe driving, like gear changing and speed control.

The present study comprised drivers from 65 to 85 years old. Those who failed the on-road test were on average older than those who passed. The on-road performance, as measured by one of the on-road protocols, weakly correlated with age. However, from this cross-sectional study, comprising drivers with such a large age span, no conclusions should be drawn upon these facts. The findings may simply reflect a cohort effect.

A discrepancy between the participants' subjective driving skills and actual driving behaviour was found, a finding further supported by other studies confirming that drivers generally have overly positive beliefs in their own driving skills (Freund et al., 2005; Lundqvist et al., 1997; Marottoli and Richardson, 1998; Patomella et al., 2008). As a matter of fact, self ratings tend to be severely biased in favour of the person's own driving performance, as confirmed by our results. The overestimation of their driving skills may imply that they do not modify their driving sufficiently, e.g., by avoiding difficult situations. In addition, this finding sends the message that self-reported driving skills should be viewed with care.

What other implications do our results suggest? Off-road tests may in clinical practice identify clients that are far from fit to drive and for safety reasons should not even be tested on road. Offering a high degree of standardization, off-road tests could also guide the on-road evaluator to assess skills that do need extra attention during the on-road assessment (Mazer et al., 1998; Unsworth et al., 2005). However, the outcomes of only two sub tests in our off-road test battery that were actually significantly different between the fail group and the pass group suggests limited benefits from the off-road tests used for the target group of the present study.

On-road tests have good face validity and are seen as the most accurate measurement of driving competence (Hunt et al., 1993), despite lack of standardization and data on reliability and validity (Withaar et al., 2000). However, clinical on-road assessments are not always used as a complement to off-road fitness-to-drive assessments (Larsson et al., 2007). To enhance validity, the on-road assessment should include a variety of situations that drivers normally encounter on a daily basis. However, constructing realistic measurements of driving performance has in previous research been proven difficult (Anstey et al., 2005). For example, an on-road assessment may never be completely standardized regarding other road users.

The present study comprised a rather small sample and there were no measurements of inter-rater reliability. Instead, the on-road performance was based on the OT's subjective impression of the older drivers' overall driving behaviour. The pass/fail outcome was not based on the total number of driving errors, which can be seen as another limitation concerning validity of the present study. Yet another limitation of the present study was the drivers' self reported health statues. Drivers with cognitive impairments could be present in both DMC-groups, which is likely to introduce unknown bias. Moreover, the representativeness of the sample is unclear, since the characteristics of those who did not participate remain unknown.

5. Conclusions

To use standardized methods for assessing fitness to drive is important for valid and reliable outcomes. The results show that some of the older drivers, without cognitive impairments impacting on their driving, displayed questionable driving behaviours. These findings indicate that we should be aware of the fact that some specific errors can actually be "normal" driving behaviours or "bad habits" habituated in years of driving. The current study also identified two cognitive subtests, which only weakly correlated to the result of the on-road assessment.

Acknowledgements

We thank Charlotta Ryd and Catarina Lundberg for the data collection, and the staff at Traffic Medicine Centre for their support. This study was funded by the Swedish Transport Agency.

Appendix A. The two scoring protocols used in the study

A.1. P-Drive (Performance Analysis of Driving Ability)

Items			
	7. Reversing the car	14. Yielding	21. Heeding warn/prohibition signs
1. Steering	8. Following instructions	15. Obeying stop regulation	22. Heeding regulation sign
2. Changing gears	9. Finding the way	16. Follow speed regulation	23. Heeding information sign
3. Using pedals	10. Position on the road	17. Attending straight ahead	24. Attention to fellow road users
4. Controlling speed, low pace	11. Keeping distance	18. Attending to the right	25. Reacting to fellow road users
5. Controlling speed, high pace	12. Organising	19. Attending to the left	26. Focusing
6. Using indicators	13. Giving right-of-way	20. Attending to mirrors	27. Solving problems

A.2. ROA (Ryd On-road Assessment)

1. Speed	2. Position	3. Attention	4. Indicator	5. Manouvering	6. Instructions	7. Traffic rules
A. Too fast for the situation	A. To the right	A. To the right	A. Does not use indicator	A. Pedals	A. Repeating needed	A. Give right of way
B. Too slow for the situation	B. To the left	B. To the left	B. Wrong direction	B. Steering	B. Reminding needed	B. Yield to traffic
C. Slow/late braking	C. Close to the vehicle in front	C. Ahead	C. Too late	C. Change gear	C. Drive the wrong way	C. Obligation to stop
D. Brake without reason	D. Sway between lanes	D. To the rear incl. rear-view mirror	D. Too early	D. Manage controls to the left		D. Exceeding speed limit
		E. Blind spot, to the right	E. Does not switch off	E. Manage controls to the right		E. Rules regarding buses
		F. Blind spot, to the left		F. Reverse		F. Crossing a solid lane line

References

Akinwuntan, A.E., Feys, H., DeWeerdt, W., Pauwels, J., Baten, G., Strypstein, E., 2002. Determinants of driving after stroke. Arch. Phys. Med. Rehabil. 83, 334–341.

Anstey, K.J., Wood, J., Lord, S., Walker, J.G., 2005. Cognitive, sensory and physical factors enabling driving safety in older adults. Clin. Psychol. Rev. 25 (1), 45–65.

Ball, K., Owsley, C., Sloane, M.E., Roenker, D.L., Bruni, J.R., 1993. Visual attention problems as a predictor of vehicle crashes in older drivers. Invest. Ophthalmol. Vis. Sci. 34 (11), 3110–3123.

Ball, K., Owsley, C., Stalvey, B., Roenker, D.L., Sloane, M.E., Graves, M., 1998. Driving avoidance and functional impairment in older drivers. Accid. Anal. Prev. 30 (3), 313–322.

Brouwer, W.H., Ponds, R.W.H.M., 1994. Driving competence in older persons. Disabil. Rehabil. 16 (3), 149–161.

Brown, L.B., Stern, R.A., Cahn-Weiner, D.A., Rogers, B., Messer, M.A., Lannon, M.C., et al., 2005. Driving scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. Arch. Clin. Neuropsychol. 20 (2), 209–215.

Di Stefano, M., Macdonald, W., 2003. Assessment of older drivers: relationships among on-road errors, medical conditions and test outcome. J. Safety Res. 34 (4), 415–429.

Dobbs, A.R., Heller, R.B., Schopflocher, D., 1998. A comparative approach to identify unsafe older drivers. Accid. Anal. Prev. 30 (3), 363–370.

Edwards, J.D., Vance, D.E., Wadley, V.G., Cissell, G.M., Roenker, D.L., Ball, K.K., 2005. Reliability and validity of useful field of view test scores as administered by personal computer. J. Clin. Exp. Neuropsychol. 27 (5), 529–543.

Evans, L., 2004. Traffic Safety, first ed. Science Serving Society, Bloomfield Hills.

Fox, G.K., Bowden, S.C., Bashford, G.M., Smith, D.S., 1997. Alzheimer's disease and driving: prediction and assessment of driving performance. J. Am. Geriatr. Soc. 45 (8), 949–953.

Fox, G.K., Bowden, S.C., Smith, D.S., 1998. On-road assessment of driving competence after brain impairment: review of current practice and recommendations for a standardized examination. Arch. Phys. Med. Rehabil. 79 (10), 1288–1296.

Freund, B., Colgrove, L.A., Burke, B.L., McLeod, R., 2005. Self-rated driving performance among elderly drivers referred for driving evaluation. Accid. Anal. Prev. 37 (4), 613–618.

Galski, T., Ehle, H.T., McDonald, M.A., Mackevich, J., 2000. Evaluating fitness to drive after cerebral injury: basic issues and recommendations for medical and legal communities. J. Head Trauma Rehabil. 15 (3), 895–908.

Hunt, L., Morris, J.C., Edwards, D., Wilson, B.S., 1993. Driving performance in persons with mild senile dementia of the Alzheimer type. J. Am. Geriatr. Soc. 41 (7), 747–752.

Johansson, K., Bronge, L., Lundberg, C., Persson, A., Seideman, M., Viitanen, M., 1996. Can a physician recognize an older driver with increased crash risk potential? J. Am. Geriatr. Soc. 44 (10), 1198–1204.

Kay, L., Bundy, A., Clemson, L., Jolly, N., 2008. Validity and reliability of the on-road driving assessment with senior drivers. Accid. Anal. Prev. 40 (2), 751–759.

Larsson, H., Lundberg, C., Falkmer, T., Johansson, K., 2007. A Swedish survey of occupational therapists' involvement and performance in driving assessments. Scand. J. Occup. Ther. 14, 215–220.

Lundberg, C., Caneman, G., Samuelsson, S.M., Hakamies-Blomqvist, L., Almkvist, O., 2003. The assessment of fitness to drive after a stroke: the Nordic Stroke Driver Screening Assessment. Scand. J. Psychol. 44 (1), 23–30.

Lundqvist, A., Alinder, J., Alm, H., Gerdle, B., Levander, S., Rönnberg, J., 1997. Neuropsychological aspects of driving after brain lesion: simulator study and on-road driving. Appl. Neuropsychol. 4 (4), 220–230.

Lundqvist, A., Gerdle, B., Rönnberg, J., 2000. Neuropsyological aspects of driving after a stroke – in the simulator and on the road. Appl. Cognit. Psychol. 14, 135–148.

Marottoli, R.A., Richardson, E.D., 1998. Confidence in, and self-rating of, driving ability among older drivers. Accid. Anal. Prev. 30 (3), 331–336.

Marottoli, R.A., Richardson, E.D., Stowe, M.H., Miller, E.G., Brass, L.M., Cooney, L.M., 1998. Development of a test battery to identify older drivers at risk for self-reported adverse driving events. J. Am. Geriatr. Soc. 46 (5), 562–568.

Mazer, B.L., Korner-Bitensky, N.A., Sofer, S., 1998. Predicting ability to drive after stroke. Arch. Phys. Med. Rehabil. 79 (7), 743–750.

Nouri, F.M., Lincoln, N.B., 1992. Validation of a cognitive assessment: predicting driving performance after stroke. Clin. Rehabil. 6, 275–281.

Odenheimer, G.L., Beaudet, M., Jette, A.M., Albert, M.S., Grande, L., Minaker, K.L., 1994. Performance-based driving evaluation of the elderly driver: safety, reliability, and validity. J. Gerontol. 49 (4), M153–M159.

Owsley, C., Ball, K., McGwin Jr., G., Sloane, M.E., Roenker, D.L., White, M.F., et al., 1998. Visual processing impairment and risk of motor vehicle crash among older adults. JAMA 279 (14), 1083–1088.

Patomella, A.H., Kottorp, A., Tham, K., 2008. Awareness of driving disability in people with stroke tested in a simulator. Scand. J. Occup. Ther. 15 (3), 184–192.

Patomella, A.H., Tham, K., Johansson, K., Kottorp, A., 2010. P-Drive on-road: internal scale validity and reliability of an assessment of on-road driving performance in people with neurological disorders. Scand. J. Occup. Ther. 17 (1), 86–93.

Patomella, A.H., Tham, K., Kottorp, A., 2006. P-drive: assessment of driving performance after stroke. J. Rehabil. Med. 38 (5), 273–279.

Persson, D., 1993. The elderly driver: deciding when to stop. Gerontologist 33 (1), 88–91.

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

Reitan, R., 1986. Trail Making Test. Manual for Administration and Scoring Tuscon. Neuropsychological Laboratory, AZ.

Schanke, A.K., Sundet, K., 2000. Comprehensive driving assessment: neuropsychological testing and on-road evaluation of brain injured patients. Scand. J. Psychol. 41 (2), 113–121.

Selander, H., Johansson, K., Lundberg, C., Falkmer, T., 2010. The Nordic stroke driver screening assessment as predictor for the outcome of an on-road test. Scand. J. Occup. Ther. 17 (1), 10–17.

Shinar, D., Meir, M., Ben-Shoham, I., 1998. How automatic is manual gear shifting? Hum. Factors 40 (4), 647–654.

Soderstrom, S.T., Pettersson, R.P., Leppert, J., 2006. Prediction of driving ability after stroke and the effect of behind-the-wheel training. Scand. J. Psychol. 47 (5), 419–429.

Stutts, J.C., Stewart, J.R., Martell, C., 1998. Cognitive test performance and crash risk in an older driver population. Accid. Anal. Prev. 30 (3), 337–346.

Swedish Transport Agency, 2010. Code of Statutes 2010:125. Swedish Transport Agency, Norrköping.

Unsworth, C.A., Lovell, R.K., Terrington, N.S., Thomas, S.A., 2005. Review of tests contributing to the occupational therapy off-road driver assessment. Aust. Occup. Ther. J. 52, 57–74.

Withaar, F.K., Brouwer, W.H., van Zomeren, A.H., 2000. Fitness to drive in older drivers with cognitive impairment. J. Int. Neuropsychol. Soc. 6 (4), 480–490.

Wood, J.M., Anstey, K.J., Lacherez, P.F., Kerr, G.K., Mallon, K., Lord, S.R., 2009. The on-road difficulties of older drivers and their relationship with self-reported motor vehicle crashes. J. Am. Geriatr. Soc. 57 (11), 2062–2069.