Original Paper

Can a Novel Web-Based Computer Test Predict Poor Simulated Driving Performance? A Pilot Study With Healthy and Cognitive-Impaired Participants

Tobias Nef[1,2], PhD; René M Müri[1,3], MD; Rahel Bieri[1]; Michael Jäger[1]; Nora Bethencourt[1]; Ioannis Tarnanas[1]; Urs P Mosimann[1,4], MD, PhD

[1]Gerontechnology and Rehabilitation Group, University of Bern, Bern, Switzerland

[2]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

[3]Division of Cognitive and Restorative Neurology, Department of Neurology, University Hospital Inselspital, University of Bern, Bern, Switzerland

[4]Department of Old Age Psychiatry, University Hospital of Psychiatry, University of Bern, Bern, Switzerland

Corresponding Author:
René M Müri, MD
Gerontechnology and Rehabilitation Group
University of Bern
Murtenstrasse 50
Bern, 3010
Switzerland
Phone: 41 31 632 88 17
Fax: 41 31 632 89 44
Email: Rene.muri@insel.ch

Abstract

Background: Driving a car is a complex instrumental activity of daily living and driving performance is very sensitive to cognitive impairment. The assessment of driving-relevant cognition in older drivers is challenging and requires reliable and valid tests with good sensitivity and specificity to predict safe driving. Driving simulators can be used to test fitness to drive. Several studies have found strong correlation between driving simulator performance and on-the-road driving. However, access to driving simulators is restricted to specialists and simulators are too expensive, large, and complex to allow easy access to older drivers or physicians advising them. An easily accessible, Web-based, cognitive screening test could offer a solution to this problem. The World Wide Web allows easy dissemination of the test software and implementation of the scoring algorithm on a central server, allowing generation of a dynamically growing database with normative values and ensures that all users have access to the same up-to-date normative values.

Objective: In this pilot study, we present the novel Web-based Bern Cognitive Screening Test (wBCST) and investigate whether it can predict poor simulated driving performance in healthy and cognitive-impaired participants.

Methods: The wBCST performance and simulated driving performance have been analyzed in 26 healthy younger and 44 healthy older participants as well as in 10 older participants with cognitive impairment. Correlations between the two tests were calculated. Also, simulated driving performance was used to group the participants into good performers (n=70) and poor performers (n=10). A receiver-operating characteristic analysis was calculated to determine sensitivity and specificity of the wBCST in predicting simulated driving performance.

Results: The mean wBCST score of the participants with poor simulated driving performance was reduced by 52%, compared to participants with good simulated driving performance (*P*<.001). The area under the receiver-operating characteristic curve was 0.80 with a 95% confidence interval 0.68-0.92.

Conclusions: When selecting a 75% test score as the cutoff, the novel test has 83% sensitivity, 70% specificity, and 81% efficiency, which are good values for a screening test. Overall, in this pilot study, the novel Web-based computer test appears to be a promising tool for supporting clinicians in fitness-to-drive assessments of older drivers. The Web-based distribution and scoring on a central computer will facilitate further evaluation of the novel test setup. We expect that in the near future, Web-based computer tests will become a valid and reliable tool for clinicians, for example, when assessing fitness to drive in older drivers.

XSL•FO
RenderX

Introduction

Cognition and Driving

Driving a car is a very challenging instrumental activity of daily living that requires the integration of high-level cognition, vision, and motor function [1]. These three domains are usually evaluated when assessing medical fitness to drive in older drivers, as they are commonly affected by age-related diseases [2]. In this article, we focus on the assessment of driving-relevant cognition in older drivers.

Since driving is a complex activity, driving performance is very sensitive to cognitive impairment [3], which is commonly the result of age-related neurodegenerative disorders (eg, Alzheimer's disease and other causes of dementia) [4]. The prevalence of neurodegenerative disorders doubles every five years after the age of 65 years [5]. Therefore, health professionals need easy access to screening tests in order to assess fitness to drive. Due to the ageing population in the Western world and increasing numbers of older drivers, identifying drivers at risk without unnecessarily restricting others is a challenging but important task [6]. This task requires reliable and valid cognitive screening tests with good sensitivity and specificity to identify at-risk drivers.

Testing Driving Performance

On-the-road testing (ORT) has been suggested as being a reasonable proxy measure for naturalistic driving in older adults with a range of cognitive impairments [7]. It is the gold standard for measuring driving performance and several authors suggest using it to assess fitness to drive in older drivers [7-10]. Despite its advantages, ORT has limitations: it is time consuming [11] and may have adverse effects that could lead to dangerous driving situations [12]. In addition, researchers cannot control for environmental conditions such as light, weather, traffic, and pedestrians [13].

That is why more recently, driving simulators (DS) have been recommended as a proxy measure for naturalistic driving and they have been introduced to assess fitness to drive of older drivers with and without cognitive impairment [14,15]. Simulators have the advantage of being intrinsically safe, providing excellent controllability, reproducibility, and standardization. Furthermore, they can be installed in specialist centers, their use is less time consuming, and they require fewer organizational demands than ORT. Several studies demonstrate the validity of DS as a proxy for naturalistic driving [13,16-19]. Disadvantages of DS are that they are expensive and large, both of which reduce their accessibility to primary care physicians and older drivers. Furthermore, there is a lack of standard test protocols and cutoff values. Finally, simulator sickness is a rather common side effect, especially for older female drivers. This interferes with DS driving performance [20].

Web-Based Computer Testing of Driving Performance

In a recent study, Rockwood et al [21] used a Web-based dataset to determine the level of cognitive impairment of older Internet users. They concluded that online tracking of people with cognitive impairment can be used to stage dementia. For review, see [22]. Also, when measuring driving performance, some of the limitations of DS can be resolved by the introduction of computer tests. By using standard personal computers with cheap off-the-shelf interface components similar to those that are used in computer gaming, cheap and easily accessible computer tests can be implemented [23]. Compared to DS, they can be easily integrated into the physician's office. Web-based computer tests use the World Wide Web to distribute the software that can run on local client computers in the physician's office. Moreover, with Web-based computer tests, the scoring of user performance can be conducted on a central server computer. This allows for generation of a dynamically growing database with normative values and ensures that all users have the same up-to-date normative values. This concept has been successfully introduced by Mills et al [24] for the application of a Web-based computer test to assess driving performance under alcohol and drug influence. The authors mention central scoring algorithms and a central database with normative values as the main advantage of Web-based computer tests. That is why we hypothesized that Web-based computer tests would also be helpful to assess fitness to drive of older drivers and we have developed the Web-based Bern Cognitive Screening Test (wBCST) to assess driving-relevant cognitive performance. The novel wBCST is based on a previously developed computer test [25].

In this pilot study, we investigate whether or not the novel wBCST correlates with DS performance and whether it is able to differentiate between participants with poor and good simulated driving performance. To have a broad and diverse test population, we recruited younger and older healthy participants as well as older participants with cognitive impairment for this study. Hence, this paper first describes the novel wBCST and the DS used, followed by a correlation analysis and a receiver operating characteristics (ROC) [26] analysis to calculate sensitivity and specificity of the novel test to predict driving simulator performance. The discussion and conclusion outline advantages and disadvantages of the wBCST and present possible future applications and research directions.

Methods

Participants

Thirty healthy younger adults (age 22-40 years), 60 healthy older adults (age >50 years), and 15 older (age >50 years) participants with cognitive impairment (Montreal Cognitive Assessment Score [MoCA] <26) [27] were recruited by advertisements in local newspapers and within the local memory clinic. All participants were required to have had a driver's

XSL·FO

RenderX

license for at least two years and to have been driving during the last two years. Exclusion criteria for the study were visual impairment (corrected far visual acuity <0.5 degrees, near visual acuity <0.8 degrees) or significant motor impairment (timed-up-and-go-test >12 seconds) [28]. The study was carried out in accordance with the Declaration of Helsinki and was approved by the local ethics board. Written informed consent was obtained from all participants prior to inclusion. No compensation for participation was provided. Six participants were excluded due to visual impairment. Eighteen participants stopped the DS drive due to simulator sickness (four younger, nine older, and five participants with cognitive impairment). Their data were excluded from further analysis. Due to a malfunctioning of the DS, the data of one healthy older test person were not recorded. The data of the remaining 80 participants were included in the analysis. There were 26 young (range 22-39 years, mean 29.4 years, SD 4.7 years), 44 healthy older (range 54-85 years, mean 68.4 years, SD 5.5 years), and 10 impaired older participants (range 55-87 years, mean 72 years, SD 9.6 years). Trail Making Test A (TMT-A) [29], Trail Making Test B (TMT-B) [29], the MoCA score [27], clock-drawing test (CDT) [30], and timed-up-and-go test [28] were assessed to characterize the participants.

Web-Based Computer Test

A literature review of the most important driving-relevant cognitive functions and how they are affected by cognitive impairment was conducted [6]. This analysis was used to develop a novel computer test [25], which we extended in the context of this study to the wBCST. The wBCST measures eye-hand coordination, selective attention, divided attention, executive function, distance judgment, and speed regulation. It is composed of five subtests, each measuring one of the before-mentioned cognitive competencies. The setup of the wBCST comprises a computer screen showing the test scenario (240B1CS/00 24 inc, Philips Inc), a steering wheel (Driving Force GT, Logitec Inc) with foot pedal, and a personal computer with Windows 7 (Microsoft Inc) operating system. Figure 1 shows a healthy test person taking the wBCST and screenshots of the visual representation of the five subtests. The test persons interact with the wBCST via steering wheel and foot pedal.

Subtest 1 measures selective attention with a visual scene consisting of a simplified street in the center of the screen (Figure 1d) with objects moving from the top down. A red dot moves automatically in the horizontal direction to avoid collisions with oncoming objects and with the roadside. The user is instructed to not touch the steering wheel, but to press the foot pedal whenever a visual target (blue square) appears in the periphery (Figure 1d). In subtest 2, eye-hand coordination is measured and the same street is presented. The test person must use the steering wheel to control the horizontal position of the red dot to avoid collisions with the other objects and with the roadside (Figure 1e). There is no peripheral subtask in this test. Subtest 3 measures divided attention and both tasks of subtests 1 and 2 need to be carried out simultaneously; hence, the user must steer (central task) and react to peripheral stimuli (peripheral task) (Figure 1f). Subtest 4 is designed to measure executive functions and the user must react to more complex peripheral stimuli (green triangle and blue square) while ignoring the movement in the center of the screen (Figure 1g). In subtest 5, distance judgment and speed regulation are assessed. The user can control the velocity of a red dot with the foot pedal. The task as shown in Figure 1h is to cross intersections without colliding with the horizontally moving objects. During the tests, the false positive and false negative errors of the peripheral target detection task and the number of collisions with moving objects as well as with the street border are recorded. The test duration is about three minutes per subtest. With instruction, administration of the entire wBCST takes roughly 20 minutes. For videos of the subtests, see Multimedia Appendices 1 and 2.

Figure 1. Older participant during the wBCST evaluation. A 24-inch monitor (a) is used to present the test material and participants interact with the system via a steering wheel (b) and foot pedal (c). Each subtest uses similar graphical objects as shown in the screenshots (d-h).
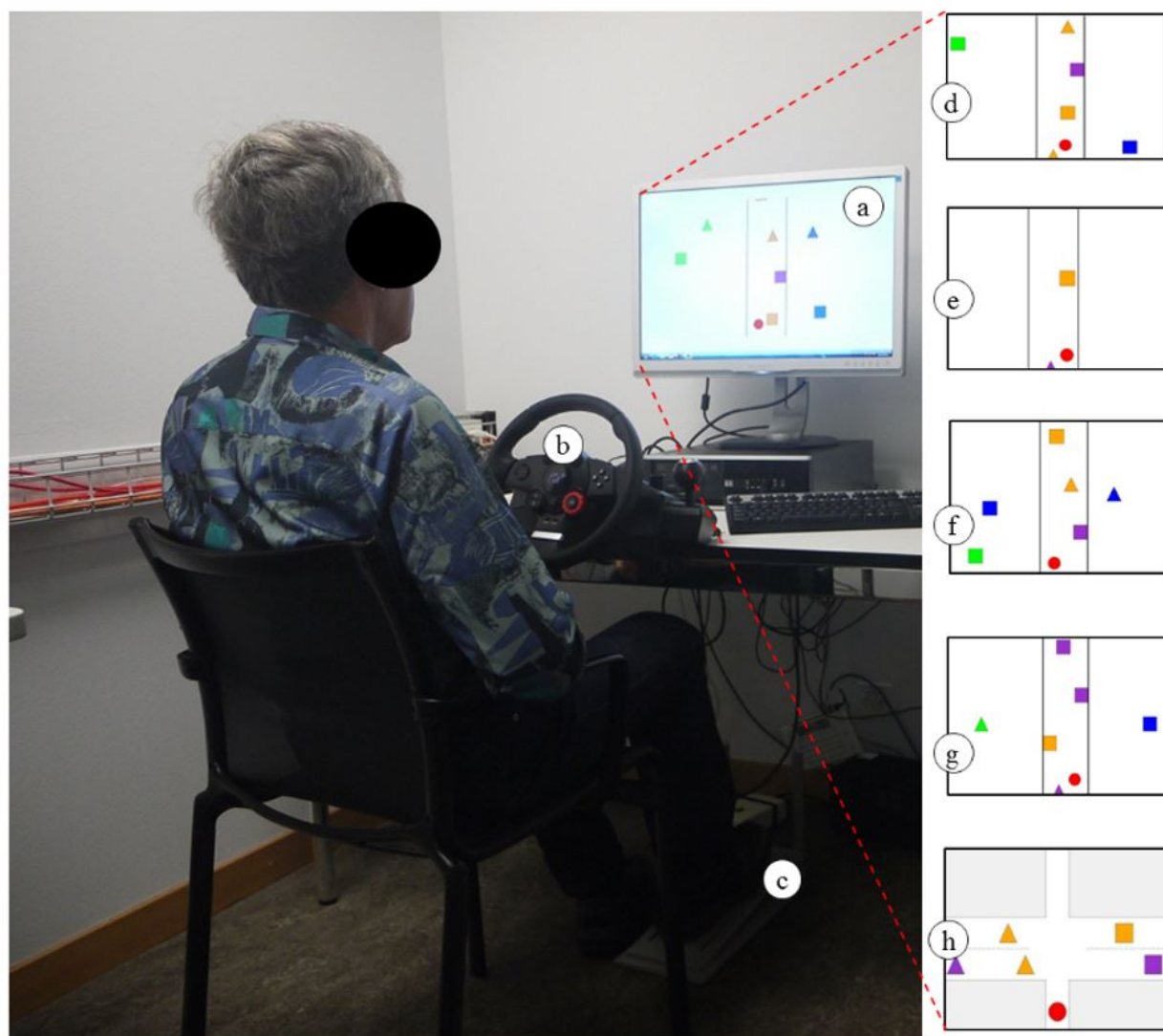

Drive in the Virtual Reality Driving Simulator

A commercially available high-fidelity fixed-frame driving simulator (F12PI-3/A88, Foerst GmbH) with a custom-built virtual driving circuit was used to measure DS performance. A virtual scene was projected by three projectors (Ultra-Short focus LCD projectors, Sanyo) with 1024x786 pixel resolution onto three projection screens (1.8 x 1.4 m) that were installed in front of the driver. DS components utilized by the participants were steering wheel, brake and gas pedals, rear and side mirrors, and turn signals (Figure 2). The virtual driving scene consisted of a street with two lanes in each direction in a suburban environment. The test drive included two left turns at intersections: one with traffic crossing left to right and one with oncoming traffic. Furthermore, it comprised a construction area with road work on both lanes, a roundabout with an unexpected cyclist, a deer hiding behind trees and suddenly jumping onto the road, a car unexpectedly leaving its parking lot, and a child running into the street after a ball. The instruction was to respect traffic rules and to drive as if in a real-world environment.

Participants performed a familiarization run (three minutes) to get used to the handling of the simulator and a test drive (six minutes) during which data were recorded. Once the end of the track was reached, the car stopped automatically. If a participant felt uncomfortable, the DS was stopped. Primary outcome measure of the driving simulator was the number of errors, E_{DS} (ie, collisions, traffic rule violations, driving in the wrong traffic lane). Secondary outcome measures were mean speed variability, mean lateral acceleration, cumulated time spent on brake, distance to collision, and time to collision. Distance and time to collision are defined by how long (time, respective distance) the virtual car could continue on its current path with constant velocity until a collision would occur.

Figure 2. High-fidelity fixed-frame driving simulator with younger test subject. The steering wheel, cockpit, and parts of two projection screens with the virtual driving screen are shown.
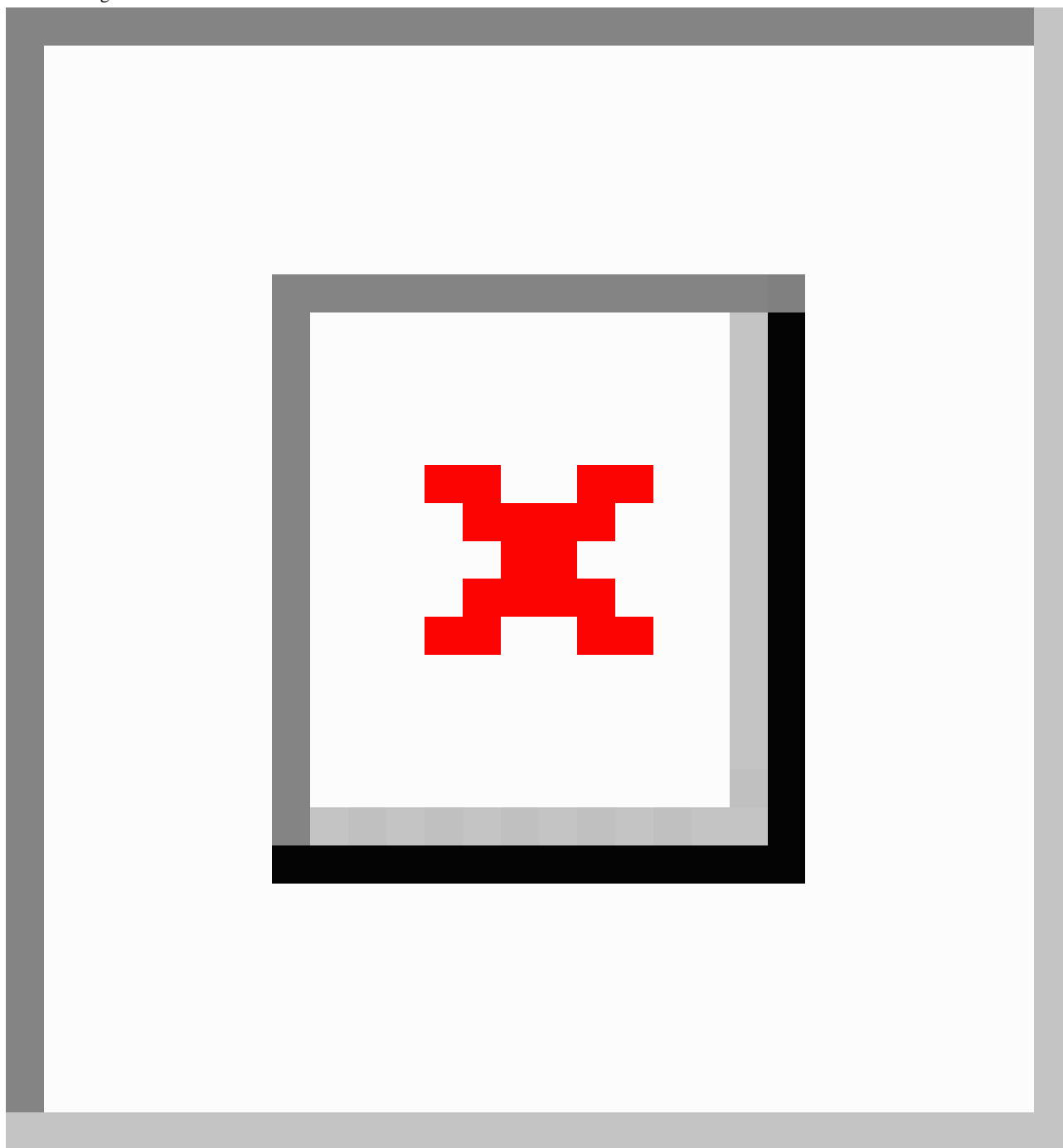
Statistical Analysis

The results of the wBCST were ranked and scored. Thus, the score, S_i, of each subtest was calculated as $S_i = (N - rank(E_i)) / (N-1)$ with N being the total number of test samples and E_i being the number of errors in the subtest. With this formula, 1 corresponds to the best possible test results and 0 the worst. Subtest 3 reveals two results, S_{3p} for the peripheral task and S_{3c} for the central task. The overall test result, S_{wBCST}, was calculated as the mean value of S_1 to S_5. The same formula was also used to score the performance in the DS—namely, the score, S_{DS}, that was calculated out of E_{DS}, as well as the scores for speed

variability, lateral acceleration, time on brake, distance to collision, and time to collision.

Pearson product-moment correlations, step-wise regressions, and associated tests of significance were calculated across S_{DS}, S_{wBCST}, and the other secondary measures. The statistical significance of the correlation was computed by transforming the correlation matrix to create a t statistic having n-2 degrees of freedom where n was the number of observations.

In a second step, the number of errors E_{DS} in the driving simulator was used as a classification criterion to divide the participants into two groups. As suggested by others [31], the mean value plus 1 SD was used as cutoff for one group with

good DS performance and another group with poor DS performance. The mean values of the DS and the wBCST for both groups were calculated and the significance of the differences was calculated using a non-parametric Mann-Whitney U-test. Hypotheses were one-sided. Furthermore, a ROC analysis was conducted to determine the sensitivity and specificity of the wBCST when predicting DS performance. A Matlab (The MathWorks Inc) algorithm based on the work of Hanley et al [32] and implemented by Cardillo et al [33] was used to calculate the ROC curve and the area under the ROC curve (AUC). The Wilcoxon test was used to calculate whether the difference from a random classifier was significant. Also, the sensitivity, specificity, and efficiency (fraction of subjects that are correctly classified) were computed for different cutoff values.

Results

User Statistics

The data of 80 participants were included in the data analysis. Ten were assigned, based on the number of errors in the DS using mean value + 1 SD cutoff, to the poor DS performance group ($E_{DS} \geq 4$) and 70 were assigned to the good DS performance group ($E_{DS} < 4$). With this criterion, all younger participants, 86% (38/44) of the healthy older, and 60% (6/10) of the older participants with cognitive impairment were classified into the good DS group. The mean test performance in MoCA, TMT-A, TMT-B, CDT, and timed-up-and-go-test of the participants with good DS results was better compared to the other group. These differences were statistically significant.

Table 1. User statistics of all participants and the two subgroups.

	All Participants, n=80	Participants with good DS[a] performance, n=70	Participants with poor DS performance, n=10	Significance of group difference
Young (<40 years), n	26	26	0	
Older (>50 years), without cognitive impairment (MoCA[b]≥26), n	44	38	6	
Older (>50 years), with cognitive impairment (MoCA<26), n	10	6	4	
MoCA, score (SD)	28.2 (2.9)	28.5 (2.8)	26.2 (3.1)	$U_{70,10}$=179.00, P=.004
TMT[c] A, seconds (SD)	28.0 (14.4)	26.9 (14.2)	35.7 (14.1)	$U_{68,9}$=535.00, P<.001
TMT B, seconds (SD) [30]	71.6 (52.5)	68.8 (53.2)	91.2 (45.7)	$U_{68,9}$=479.00, P=.003
CDT, score (SD)	6.2 (1.9)	6.5 (1.6)	4.3 (3.1)	$U_{70,10}$=170.00, P<.001
Timed-up-and-go-test, seconds (SD)	6.9 (3.2)	6.8 (2.9)	7.5 (4.7)	$U_{63,8}$=420.50, P=.001

[a]DS: driving simulator

[b]MoCA: Montreal Cognitive Assessment

[c]TMT: trail making test (A and B)

[d]CDT: clock drawing test

Correlation Analysis of wBCST and DS Performance

Figure 3 shows the results of the correlation analysis with the correlation coefficient r and the associated P values. The wBCST score, S_{wBCST}, correlates with the DS score, S_{DS}, with r=0.32 (P=.004). The individual results of the wBCST subtest, $S_{1...6}$, correlate with the overall score, S_{wBCST}, with r values varying between 0.68-0.83 with P<.001. In the DS, the overall score S_{DS} correlates with the speed variation with r=0.38 (P<.001), the lateral acceleration with r=0.40 (P<.001), the time spent on the brake pedal r=0.13 (not significant, P=.24), the distance to collision with r=0.19 (not significant, P=.19), and the time to collision r=0.31 (not significant, P=.23).

Sensitivity and Specificity of the wBCST in Predicting DS Performance

The ranked normalized wBCST performances of the two groups are represented in Figure 4. The group with poor DS performance performed less well in all tests. The group differences are significant for all subscores and highly significant (P<.001) for the overall S_{wBCST} score, subtest 2 (S_2), and subtest 5 (S_5).

The ROC curve for using S_{wBCST} to predict DS performance is shown in Figure 5. The AUC=0.80 is significantly better (P<.001) than a random classifier. The 95% confidence interval of the AUC is 0.68-0.92. A selection of possible cutoff scores and the corresponding sensitivity, specificity, and efficiency values is presented in Table 2.

XSL•FO
RenderX

Figure 3. Correlation analysis of the wBCST and the DS. The table shows the Pearson product-moment correlation coefficient *r* and the associated *P* values in parentheses. Correlations with *P*<.05 are indicated in bold font.

Correlation coefficient r (p value)	Web-based Computer Test (WbCT)							Driving Simulator (DS)					
	WbCT Score	Subtest 1 (peripheral)	Subtest 2 (central)	Subtest 3 (peripheral)	Subtest 3(central)	Subtest 4	Subtest 5	Total Errors DS	Speed variation	Lateral acceleration	Time on brakes	Distance to collision	Time to collision
Age	**-0.78**	**-0.5**	**-0.64**	**-0.61**	**-0.64**	**-0.51**	**-0.74**	-0.18	0.11	-0.09	0.04	**0.01**	0.16
	(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	-0.1	-0.32	-0.44	-0.24	**-0.04**	-0.16
TMT A	**-0.6**	**-0.51**	**-0.53**	**-0.41**	**-0.52**	**-0.46**	**-0.46**	-0.21	**0.22**	-0.13	0.17	-0.05	0.17
	(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	-0.06	**-0.05**	-0.25	-0.12	-0.32	-0.13
TMT B	**-0.6**	**-0.5**	**-0.47**	**-0.45**	**-0.5**	**-0.53**	**-0.4**	**-0.24**	0.18	-0.17	0.19	0.03	0.06
	(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**-0.03**	-0.11	-0.13	-0.08	-0.19	-0.43
MoCA	**0.62**	**0.49**	**0.46**	**0.45**	**0.52**	**0.58**	**0.46**	**0.31**	-0.11	0.19	-0.16	-0.07	-0.11
	(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**-0.005**	-0.34	-0.09	-0.15	-0.46	-0.34
CDT	0.26	0.23	0.22	0.21	0.21	**0.25**	0.07	0.28	0.16	0.14	0.16	-0.06	0.11
	-0.02	-0.04	-0.05	-0.06	-0.07	-0.02	-0.44	-0.01	-0.15	-0.22	-0.15	-0.4	-0.33
Timed up and go	0.05	**0.13**	-0.01	0.11	-0.03	0.08	-0.09	0.07	-0.07	0.07	-0.07	0.19	0.13
	-0.33	**-0.25**	-0.05	-0.35	-0.23	-0.5	-0.43	-0.48	-0.47	-0.45	-0.45	-0.09	-0.25
WbCT Score		**0.68**	**0.81**	**0.83**	**0.84**	**0.67**	**0.8**	**0.32**	-0.16	0.11	-0.06	-0.07	-0.18
		(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**-0.004**	-0.16	-0.35	-0.42	-0.48	-0.11
Subtest 1 (peripheral)			**0.5**	**0.47**	**0.51**	**0.45**	**0.48**	**0.24**	-0.08	**0.26**	0.08	-0.04	-0.12
			(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	**(<.001)**	**-0.03**	-0.47	**-0.02**	-0.48	-0.28	-0.3
Subtest 2 (central)				**0.47**	**0.85**	**0.48**	**0.65**	0.15	-0.21	-0.04	-0.04	-0.14	**-0.25**
				(<.001)	**(<.001)**	**(<.001)**	**(<.001)**	-0.18	-0.06	-0.29	-0.26	-0.21	**-0.03**
Subtest 3 (peripheral)					**0.52**	**0.42**	**0.58**	**0.25**	-0.2	0.06	-0.12	-0.06	-0.1
					(<.001)	**(<.001)**	**(<.001)**	**-0.03**	-0.07	-0.39	-0.28	-0.38	-0.39
Subtest 3 (central)						**0.45**	**0.76**	**0.24**	-0.21	**0.1**	-0.07	-0.18	**-0.26**
						(<.001)	**(<.001)**	**-0.03**	-0.06	**-0.007**	-0.44	-0.1	**-0.02**
Subtest 4							**0.37**	0.29	0.05	0.12	**0.01**	0.07	-0.04
							(<.001)	-0.01	-0.34	-0.28	**-0.04**	-0.48	-0.3
Subtest 5								0.28	-0.02	0.16	-0.01	-0.04	-0.12
								-0.01	-0.16	-0.16	-0.1	-0.3	-0.29
Total Errors DS									**0.38**	**0.4**	0.13	0.19	0.13
									(<.001)	**(<.001)**	-0.24	-0.09	-0.23
Speed variation										**0.43**	**0.46**	**0.22**	**0.53**
										(<.001)	**(<.001)**	**-0.05**	**(<.001)**
Lateral acceleration											0.16	**0.3**	**0.32**
											-0.16	**-0.007**	**-0.003**
Time on brakes												0.26	**0.3**
												-0.02	**-0.007**
Distance to collision													**0.5**
													(<.001)

Table 2. Receiver operating characteristic (ROC) curve analysis: possible cutoff values and corresponding sensitivity, specificity, and efficiency.

Cutoff	Sensitivity	Specificity	Efficiency
0.88	0.94	0.30	0.86
0.84	0.91	0.30	0.84
0.83	0.89	0.30	0.81
0.81	0.86	0.30	0.79
0.77	0.84	0.50	0.80
0.75	0.83	0.70	0.81
0.73	0.80	0.70	0.79
0.68	0.77	0.70	0.76
0.66	0.74	0.70	0.74
0.65	0.71	0.70	0.71
0.64	0.69	0.70	0.69
0.63	0.69	0.90	0.71
0.61	0.66	0.90	0.69
0.57	0.63	0.90	0.66
0.56	0.60	0.90	0.64
0.53	0.57	0.90	0.61
0.52	0.54	0.90	0.59
0.51	0.53	0.90	0.58

Figure 4. Web-based Bern Cognitive Screening Test (wBCST) performance of the group with good simulated driving performance (n=70) and group with poor simulator performance (n=10). All values are normalized and ranked. The score S_wBCST is the mean value of the subscores S (1…5). Subtest 3 is represented with two entries, S_{3p} for the peripheral recognition task and S_{3c} for the central steering task. Bars indicate the standard error.
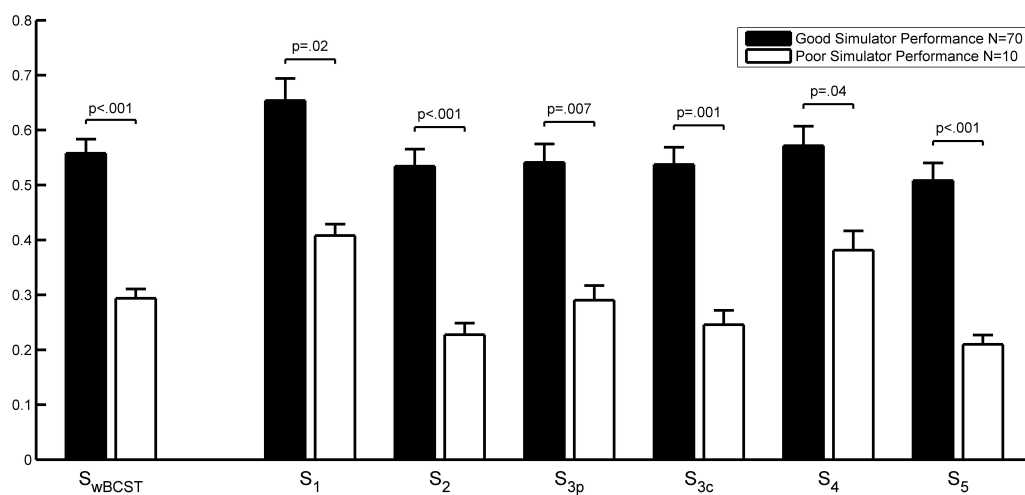
Figure 5. Receiver operating characteristics (ROC) curve for using S_wBCST to predict SD driving performance, respectively to differentiate between subjects with good and poor DS performance. The thin gray line is the empirical curve, the solid black line is the smoothed (Gaussian-based) curve, and the dotted diagonal line indicates no discrimination (50% chance).
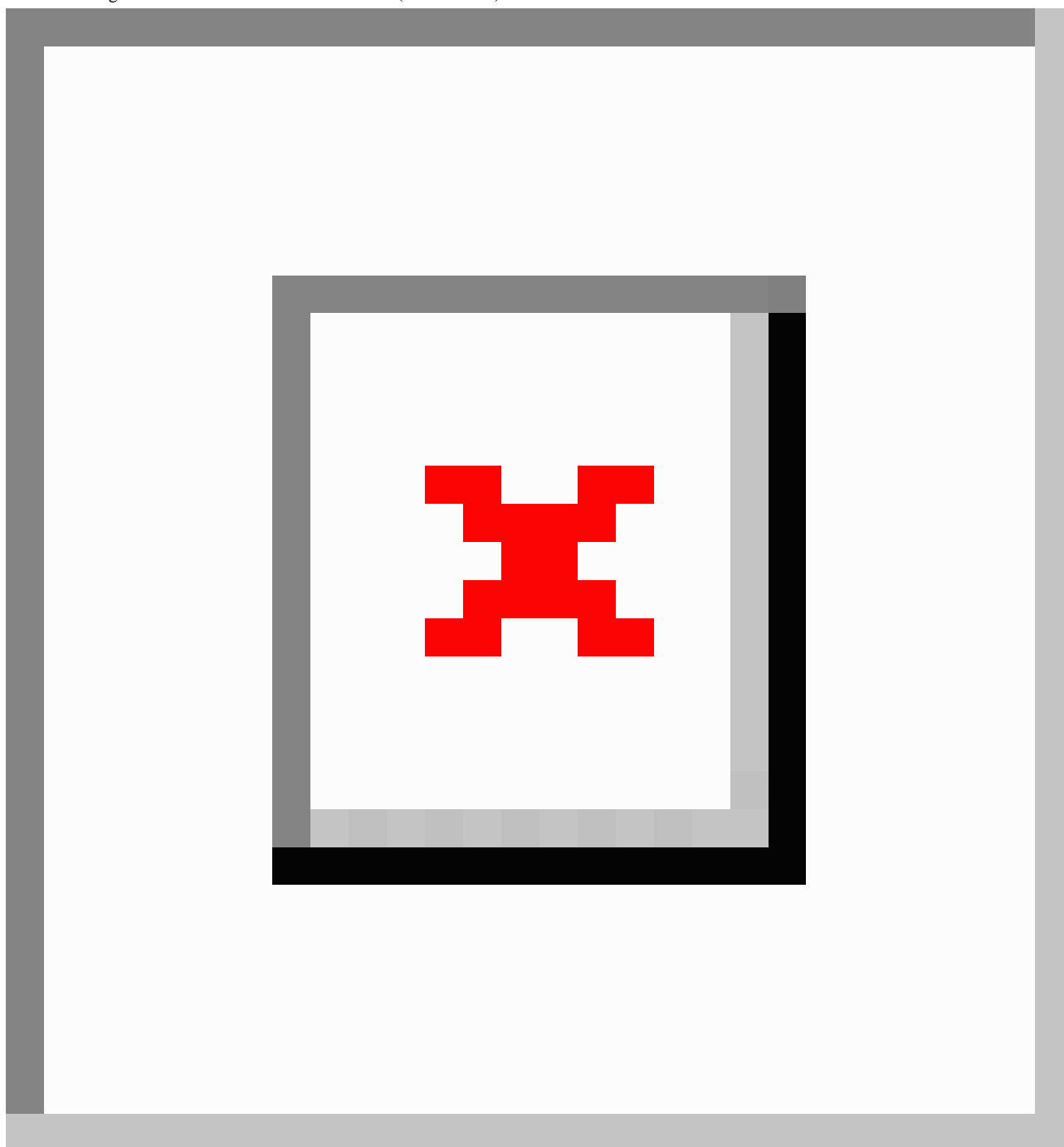
Discussion

Principal Results

When using the criteria proposed by Cohen [34], the correlation between S_{DS} and the total score S_{wBCST} has a medium effect size ($r > 0.3$) and is significant with $P = .004$, which supports the hypothesis that the wBCST correlates with DS performance. Very good correlations with large effect sizes ($r > 0.5$) were found among the five subtests of the wBCST. This is an important result for the novel test since it indicates good consistency among the wBCST subtests. It could be explained by the uniform visual stimulation material that is used in the subtests (Figure 1). The correlation of the main outcome measure, S_{DS}, and the secondary outcome measures is medium and statistically significant for the speed variation and the lateral acceleration, but only small and not significant for the time spent on brakes, distance to collision, and time to collision. This was also reported by others [35]. However, it is worthwhile to emphasize here that the novel tests show a good correlation of S_{wBCST} with existing paper-pencil-based cognitive screening tools, such as the MoCA score, the TMT-A, and the TMT-B, but not with the CDT.

The group differences in the wBCST test and its subscores are all significant, which is a prerequisite for using the wBCST to

predict DS performance. The AUC of the ROC curve is slightly larger than 0.80, which is generally considered a good test [33,36]. The ROC curve (Figure 5) and the corresponding table with sensitivity, specificity, and efficiency values allows for selecting a cutoff score that leads to the desired test properties. As mentioned in the introduction, the test should have both a high sensitivity and specificity. When considering Table 2, an appropriate cutoff with a high efficiency could be 0.75, which would lead to 83% sensitivity, 70% specificity, and 81% efficiency. These values are far from perfect, but are within what is to be expected for a test measuring multifaceted characteristics of the cognitive ability to drive safely [37-43].

In the present study, the overall test result, S_{wBCST}, is calculated simply as the mean of the subscores, $S_{1...5}$. It could be that one subtest is more informative than another and, in this case, the subscores should be weighted differently. The sample size of this study is too small to determine the optimal weighting parameters, but this is an interesting question for a future study.

Overall, the novel wBCST was very well accepted by the test population and there were no drop-outs in the wBCST. This is in contrast to the DS with 19 drop-outs (19.1%) due to simulator sickness (18 drop-outs) and technical problems (1 drop-out). Since it is Web-based, the distribution of the test software to different test-sites will be fairly easy and with the central scoring algorithm, data consistency among the different centers can be ensured, which will facilitate validation studies in larger populations.

The online instructions of the test procedure take about 5 minutes. We did not observe difficulties of the test persons to understand the task, except two cognitively impaired test subjects wanted to go through the instructions twice. We concluded that the instructions are clear, but participants should have the option to repeat the instructions.

The steering wheel and the foot pedal are fairly cheap accessories, but to further improve accessibility of the novel test it would be beneficial if the test could also be used with keyboard and mouse. We observed that steering wheel and foot pedal seem to increase the face-validity of the tests, but we would expect that the measured cognitive functions should be independent from the input modality. This could be investigated in a future study.

Limitations

This pilot study has some limitations that need to be mentioned. One is the large number of drop-outs in the DS that might create a selection bias, since it cannot be excluded that people with poor driving performance might be more prone to simulator sickness. Although the current literature on simulator sickness suggests other contributing factors (eg, age, gender), this cannot be excluded [20]. Another limitation is that there are no published and accepted cutoff values for the DS, which makes the selection of the cutoff for the group differentiation difficult. Compared to the DS evaluation, the administration of the wBCST takes 20 minutes, which is quite long. In future studies, one could investigate whether or not test duration could be shortened. Furthermore, the test-retest and inter-rater reliability should be assessed in a future study.

Conclusions

In this pilot study, the novel wBCST looks like a promising test to support clinicians in fitness-to-drive assessments of older drivers. The Web-based distribution and the scoring on a central computer will facilitate further evaluation of the novel test setup. In its current form, the program requires local installation on a client computer in the physician's office. This is currently not considered a disadvantage, but, of course, one could transfer the test program to run within a Web browser, which would not require local installation. The hardware requirements of the wBCST are very minor and include an office-type personal computer with Windows 7 operating system and a simple steering wheel (eg, Driving Force GT, Logitec Inc). Overall, when considering disadvantages of DS (costs, simulator sickness, space requirements), the authors believe that in many clinical environments the wBCST is better suited to support physicians in fitness-to-drive assessments than a DS. That is why we expect that in the near future, Web-based computer tests will become a valuable and usable tool for fitness-to-drive assessment in older drivers.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screencast with animated images from the experimental setup (.wmv File, 44 MB).

[WMV File (Windows Media Video), 43MB-Multimedia Appendix 1]

Multimedia Appendix 2

Screencast with animated images from the experimental setup (.mp4 File, 92 MB).

XSL•FO

RenderX

[MP4 File (MP4 Video), 90MB-Multimedia Appendix 2]

References

1. Sherman FT. Driving: the ultimate IADL. Geriatrics 2006 Oct;61(10):9-10. [Medline: 17034267]
2. Marshall S, Demmings EM, Woolnough A, Salim D, Man-Son-Hing M. Determining fitness to drive in older persons: a survey of medical and surgical specialists. Can Geriatr J 2012 Dec;15(4):101-119 [FREE Full text] [doi: 10.5770/cgj.15.30] [Medline: 23259024]
3. Mosimann UP, Bächli-Biétry J, Boll J, Bopp-Kistler I, Donati F, Kressig RW, et al. [Consensus recommendations for the assessment of fitness to drive in cognitively impaired patients]. Praxis (Bern 1994) 2012 Mar 28;101(7):451-464. [doi: 10.1024/1661-8157/a000893] [Medline: 22454307]
4. Chi MJ, Lee CY, Wu SC. The prevalence of chronic conditions and medical expenditures of the elderly by chronic condition indicator (CCI). Arch Gerontol Geriatr 2011;52(3):284-289. [doi: 10.1016/j.archger.2010.04.017] [Medline: 20452688]
5. Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. Lancet 2005 Dec 17;366(9503):2112-2117 [FREE Full text] [doi: 10.1016/S0140-6736(05)67889-0] [Medline: 16360788]
6. Wagner JT, Müri RM, Nef T, Mosimann UP. Cognition and driving in older persons. Swiss Med Wkly 2011;140:w13136 [FREE Full text] [doi: 10.4414/smw.2011.13136] [Medline: 21240690]
7. Davis JD, Papandonatos GD, Miller LA, Hewitt SD, Festa EK, Heindel WC, et al. Road test and naturalistic driving performance in healthy and cognitively impaired older adults: does environment matter? J Am Geriatr Soc 2012 Nov;60(11):2056-2062. [doi: 10.1111/j.1532-5415.2012.04206.x] [Medline: 23110378]
8. Withaar FK, Brouwer WH, van Zomeren AH. Fitness to drive in older drivers with cognitive impairment. J Int Neuropsychol Soc 2000 May;6(4):480-490. [Medline: 10902417]
9. Ott BR, Heindel WC, Papandonatos GD, Festa EK, Davis JD, Daiello LA, et al. A longitudinal study of drivers with Alzheimer disease. Neurology 2008 Apr 1;70(14):1171-1178. [doi: 10.1212/01.wnl.0000294469.27156.30] [Medline: 18216302]
10. Duchek JM, Carr DB, Hunt L, Roe CM, Xiong C, Shah K, et al. Longitudinal driving performance in early-stage dementia of the Alzheimer type. J Am Geriatr Soc 2003 Oct;51(10):1342-1347. [Medline: 14511152]
11. Di Stefano M, Macdonald W. On-the-road evaluation of driving performance. In: Driver rehabilitation and community mobility: principles and practice. St. Louis, Mo: Elsevier Mosby; 2006:255-274.
12. Lee HC, Lee AH, Cameron D. Validation of a driving simulator by measuring the visual attention skill of older adult drivers. Am J Occup Ther 2003;57(3):324-328. [Medline: 12785671]
13. Shechtman O, Classen S, Awadzi K, Mann W. Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. Traffic Inj Prev 2009 Aug;10(4):379-385. [doi: 10.1080/15389580902894989] [Medline: 19593717]
14. Frittelli C, Borghetti D, Iudice G, Bonanni E, Maestri M, Tognoni G, et al. Effects of Alzheimer's disease and mild cognitive impairment on driving ability: a controlled clinical study by simulated driving test. Int J Geriatr Psychiatry 2009 Mar;24(3):232-238. [doi: 10.1002/gps.2095] [Medline: 18615781]
15. de Simone V, Kaplan L, Patronas N, Wassermann EM, Grafman J. Driving abilities in frontotemporal dementia patients. Dement Geriatr Cogn Disord 2007;23(1):1-7. [doi: 10.1159/000096317] [Medline: 17047327]
16. Bella F. Driving simulator for speed research on two-lane rural roads. Accid Anal Prev 2008 May;40(3):1078-1087. [doi: 10.1016/j.aap.2007.10.015] [Medline: 18460376]
17. Bieliauskas LA. Neuropsychological assessment of geriatric driving competence. Brain Inj 2005 Mar;19(3):221-226. [Medline: 15832896]
18. Johnson MJ, Chahal T, Stinchcombe A, Mullen N, Weaver B, Bédard M. Physiological responses to simulated and on-road driving. Int J Psychophysiol 2011 Sep;81(3):203-208. [doi: 10.1016/j.ijpsycho.2011.06.012] [Medline: 21726587]
19. Yan X, Abdel-Aty M, Radwan E, Wang X, Chilakapati P. Validating a driving simulator using surrogate safety measures. Accid Anal Prev 2008 Jan;40(1):274-288. [doi: 10.1016/j.aap.2007.06.007] [Medline: 18215559]
20. Classen S, Bewernitz M, Shechtman O. Driving simulator sickness: an evidence-based review of the literature. Am J Occup Ther 2011;65(2):179-188. [Medline: 21476365]
21. Rockwood K, Richard M, Leibman C, Mucha L, Mitnitski A. Staging dementia from symptom profiles on a care partner website. J Med Internet Res 2013;15(8):e145 [FREE Full text] [doi: 10.2196/jmir.2461] [Medline: 23924608]
22. Eghdam A, Scholl J, Bartfai A, Koch S. Information and communication technology to support self-management of patients with mild acquired cognitive impairments: systematic review. J Med Internet Res 2012;14(6):e159 [FREE Full text] [doi: 10.2196/jmir.2275] [Medline: 23165152]
23. Molnar FJ, Patel A, Marshall SC, Man-Son-Hing M, Wilson KG. Clinical utility of office-based cognitive predictors of fitness to drive in persons with dementia: A systematic review. J Am Geriatr Soc 2006 Dec;54(12):1809-1824. [doi: 10.1111/j.1532-5415.2006.00967.x] [Medline: 17198485]
24. Mills KC, Spruill SE, Walker JM, Lamson M. A clinical trial demonstration of a web-based test for alcohol and drug effects. J Stud Alcohol Drugs 2009 Mar;70(2):308-313. [Medline: 19261244]

25.   Bieri R, Gruber N, Nef T, Müri R, Mosimann UP. A novel computer test to assess driving-relevant cognitive functions - a pilot study. International Psychogeriatrics 2013 Dec 01:1 (forthcoming).

26.   Swets JA. Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers. In: Signal detection theory and ROC analysis in psychology and diagnostics: collected papers. Mahwah, N.J: Lawrence Erlbaum Associates; 1996.

27.   Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

28.   Bischoff HA, Stähelin HB, Monsch AU, Iversen MD, Weyh A, von Dechend M, et al. Identifying a cut-off point for normal mobility: a comparison of the timed 'up and go' test in community-dwelling and institutionalised elderly women. Age Ageing 2003 May;32(3):315-320 [FREE Full text] [Medline: 12720619]

29.   Reitan RM. The relation of the trail making test to organic brain damage. J Consult Psychol 1955 Oct;19(5):393-394. [Medline: 13263471]

30.   Shulman KI. Clock-drawing: is it the ideal cognitive screening test? Int J Geriatr Psychiatry 2000 Jun;15(6):548-561. [Medline: 10861923]

31.   Trustram Eve C, de Jager CA. Piloting and validation of a novel self-administered online cognitive screening tool in normal older persons: the Cognitive Function Test. Int J Geriatr Psychiatry 2013 Jun 11:1. [doi: 10.1002/gps.3993] [Medline: 23754255]

32.   Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982 Apr;143(1):29-36. [Medline: 7063747]

33.   Cardillo G. Matlabcentral. 2008 May 16. ROC curve: compute a Receiver Operating Characteristics curve URL: http://www.mathworks.com/matlabcentral/fileexchange/19950-roc-curve [accessed 2013-10-14] [WebCite Cache ID 6KNSXZ9HO]

34.   Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]

35.   de Winter JC, de Groot S, Mulder M, Wieringa PA, Dankelman J, Mulder JA. Relationships between driving simulator performance and driving test results. Ergonomics 2009 Feb;52(2):137-153. [doi: 10.1080/00140130802277521] [Medline: 18972239]

36.   Swets JA. Measuring the accuracy of diagnostic systems. Science 1988 Jun 03;240(4857):1285-1293. [doi: 10.1126/science.3287615] [Medline: 3287615]

37.   Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001 May 8;56(9):1133-1142. [Medline: 11342677]

38.   Smith T, Gildeh N, Holmes C. The Montreal Cognitive Assessment: validity and utility in a memory clinic setting. Can J Psychiatry 2007 May;52(5):329-332. [Medline: 17542384]

39.   Barberger-Gateau P, Commenges D, Gagnon M, Letenneur L, Sauvel C, Dartigues JF. Instrumental activities of daily living as a screening tool for cognitive impairment and dementia in elderly community dwellers. J Am Geriatr Soc 1992 Nov;40(11):1129-1134. [Medline: 1401698]

40.   Janke MK. Assessing older drivers. Journal of Safety Research 2001 Mar;32(1):43-74. [doi: 10.1016/S0022-4375(00)00048-7]

41.   Ball K. Attentional problems and older drivers. Alzheimer Dis Assoc Disord 1997 Jun;11 Suppl 1:42-47. [Medline: 9194968]

42.   De Raedt R, Ponjaert-Kristoffersen I. Predicting at-fault car accidents of older drivers. Accid Anal Prev 2001 Nov;33(6):809-819. [Medline: 11579983]

43.   Janke MK, Eberhard JW. Assessing medically impaired older drivers in a licensing agency setting. Accid Anal Prev 1998 May;30(3):347-361. [Medline: 9663294]

Abbreviations

AUC: area under the curve
CDT: clock drawing test
DS: driving simulator
MoCA: Montreal Cognitive Assessment
ORT: on-the-road testing
ROC: receiver operating characteristic curve
TMT A/B: trail making tests A/B
wBCST: Web-based Bern Cognitive Screening Test