ARE DRIVING SIMULATORS SUITABLE TO MEASURE THE DRIVING COMPETENCE OF ELDERLY DRIVERS?

Ramona Kenntner-Mabiala[1], C. Maag[1], Yvonne Kaussner[1], S. Hoffmann[1], Markus Schumacher[2]

[1]WIVW, Institute for Traffic Sciences, Robert Bosch Str. 4, 97209 Veitshöchheim, Germany
[2]Federal Highway Research Institute, Brüderstr. 53, 51427 Bergisch Gladbach, Germany
Email: kenntner@wivw.de

Summary: The present project aimed to develop and validate a methodology for driving simulators to assess and diagnose driving ability of elderly drivers. A driving simulation course has been developed which covered a representative selection of driving tasks of moderate difficulty as well as scenarios which are particularly difficult for elderly drivers. Driving errors were semi-automatically registered and classified by a tablet PC application. Based on the registered driving errors, the driving competence of each driver was rated on an 11-point fitness-to-drive (FtD) rating scale by specifically trained raters. The driving course was validated on the basis of a 60-minute standardized driving test in real traffic. By including similar driving tasks, it was ensured that it was structurally comparable to the simulated course. 30 elderly drivers (> 70 years) and 30 control drivers (25-50 years) were assessed in the simulation and in real traffic. During the driving tests, more driving errors were registered for the elderly drivers than for the controls both in the simulator and in real traffic. FtD-ratings and total number of driving errors during the driving tests in the simulation correlated up to $r = .80$ with the FtD-ratings of the driving tests in real traffic. ROC-Analyses revealed at Sensitivity-Specificity Ratio of 85.71 : 82.61 at best. Overall, driving simulation was well accepted by the subjects. The findings of the study confirm the validity of driving simulation as a tool to diagnose driving ability and argue for its introduction as a diagnostic tool.

INTRODUCTION

In the near future, an increasing number of older drivers will be on the roads due to demographic changes and the increasing use of the driver's license. With age, impairments develop that can influence how safe elderly people drive (Anstey et al., 2015). However, particularly in rural areas, driving is a very important factor for the quality of life of elderlies with respect to social connection, independence and participation (Karthaus & Falkenstein, 2016). Therefore, elderly drivers who are driving safely should continue to drive as long as possible.

Modern driving simulators are considered to be a promising alternative to classical methods of fitness-to-drive (FtD) evaluations as they combine the advantages of psychometric testing procedures and driving assessments in real traffic: Modern traffic can be simulated in a realistic way and all clients can be evaluated under identical conditions. Besides, a simulation allows for a safe, specific and replicable triggering of critical events or events that are specific for the respective research question. Numerous variables can be precisely measured (e.g. lane keeping or reaction times). The driver can practice compensatory strategies which can be captured almost as exact as in real traffic.

The aim of the present study was to evaluate if the fitness-to-drive of elderly drivers can be measured similarly well in a driving simulator as in real traffic.

METHODS

Subjects and design

30 elderly (>70 years) and 30 control subjects (25-50 years) took part in the study. A 2x3 experimental plan with the dependent factor 'method' (three levels: real traffic vs. high-fidelity simulator vs. compact simulator) and the group factor 'age' was realized. All subjects were active drivers holding a valid driving license. For their participation, subjects got 100 €. All participants gave written informed consent. For the elderly drivers, the MMSE (Mini Mental Status Examination, Folstein et al. 1975) was conducted before the driving test. One subject was excluded due to the suspicion of suffering from dementia.

Prior to inclusion, a standardized simulator familiarization with a duration of about 2 hours was absolved by each subject according to Hoffmann & Buld (2006). The aim of this familiarization program is to avoid simulator sickness and to establish a naturalistic driving behavior regarding, speeding, braking, turning and estimation of distances.

Driving course for the simulation

The course included a set of representative driving tasks of moderate difficulty on highway, rural road and in towns. Furthermore, age specific scenarios that are known to be especially difficult for elderly drivers are contained. They include particularly complex situations at intersections with diverse traffic (different vehicles, cyclists, pedestrians) and situations with different rules of priority. This testing course was programmed with the driving simulation software SILAB®. The course had a duration of about 60 min.
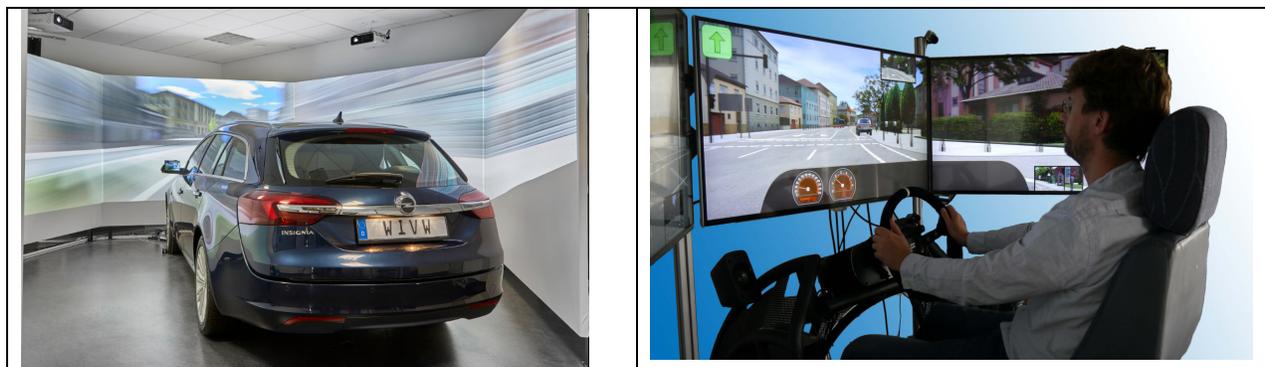

Figure 1. High-fidelity (left side) and compact simulator (right side).

To evaluate which level of hardware is needed at minimum to collect valid data, the course was implemented in two different driving simulators: a static high-fidelity simulator (figure 1, left side) with very high quality hardware components and a horizontal field of view of 300° as well as an economic compact simulator with simpler hardware and a field of view of 180° (figure 1, right side).

Driving test in real traffic

The driving test consisted of a course in and around the city of Wurzburg, Germany, that lasted 60 minutes. It comprised a representative profile of driving tasks that are structurally comparable to the simulated driving test (figure 2). All driving tests in real traffic were accompanied by a driving instructor. Test cars had double pedals that allowed interventions by the driving instructor in dangerous situations.

Description	*Real Traffic*	*Realization in SILAB®*
Straight forward at bending main road		
Dealing with slower vehicles during ban of overtaking.		

Figure 2. Examples for the structural equivalence between the courses in real and simulated environment.

Assessing fitness to drive

In simulated and in real traffic, trained raters registered and classified driving errors in a semi-automated manner with the help of the SILAB® tablet application S.A.F.E. (Standardized Application for Fitness-to-Drive-Evaluations; figure 3). Based on the registration and classification of driving errors a global assessment of the driving performance is given and a profile of the individual's ability to drive is created. As a global measure of driving fitness, the raters assessed the driving behavior using the Fitness-to-Drive-Scale (Ftd-Scale). This scale is comprised of three verbal categories (driving behavior is normal – driving behavior is noticeable – driving behavior is critical). Each of these categories is subdivided into three numerical sub-categories which allow for a more differentiated staging of driving fitness. At the upper and the lower end, the scale is complemented by the extreme categories "absolutely unfit to drive" and "fit to drive without any restrictions". In appropriate rater-trainings, investigators reached an

interrater reliability |> 0.8. After the drive, subjects rated their driving performance themselves using this scale. After the driving test in real traffic, the driving instructor also rated the driving performance of the subjects on the FtD-Scale

Figure 3. Tablet-application S.A.F.E.

According to the procedure described by Kenntner-Mabiala et al. (2015), driving errors were classified as follows:

- Tactical errors with respect to longitudinal control (speed too high, inadequate speed/acceleration/deceleration, speed too low, time headway too low/tailgating).
- Operational errors with respect to lateral control (lateral distance to objects/vehicles too low, bad lane keeping/lane departures,).
- Cognitively based tactical errors (delayed lane changing, inappropriate choice of lane, delayed/insufficient securing, driving on impermissible lanes, insufficient securing, violating right of way, overcautious securing, no/untimely blinking, ambiguous communication with other road users, inappropriate dealing with communication of other road users, navigation errors).
- Critical situations (with respect to pedestrians/cyclists, with respect to other vehicles, interventions by the driving instructor).
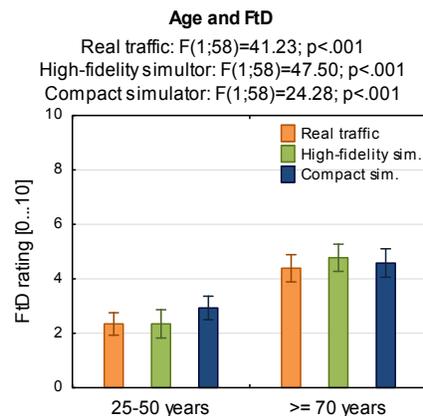
RESULTS

Age and FtD

Real traffic: $F(1;58)=41.23$; $p<.001$
High-fidelity simultor: $F(1;58)=47.50$; $p<.001$
Compact simulator: $F(1;58)=24.28$; $p<.001$

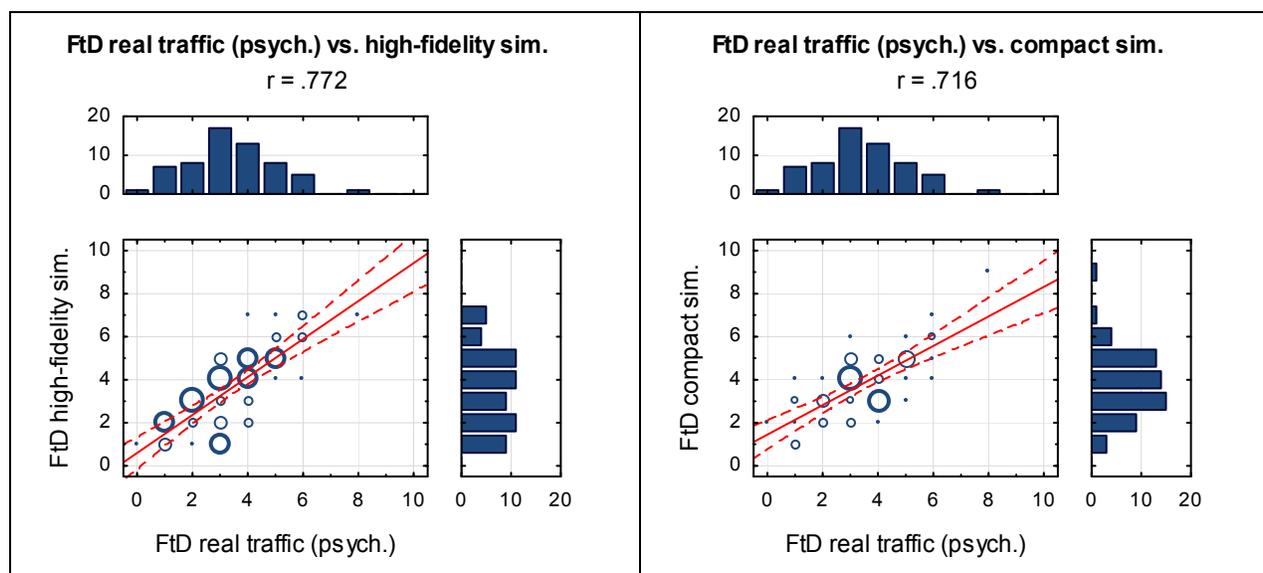**Figure 4. FtD-ratings by age group and test method assigned by the trained rater.**

Figure 5. Correlation between FtD-ratings in real traffic and Ftd-ratings in the high-fidelity simulator (left side) resp. the compact simulator (right side) given by the trained rater (psych).

Testing the relative validity of the methodological approach (see. e.g. Aksan et al., 2017), data analysis yields that on average the group of elderly drivers performs worse with regard to different driving performance parameters than the control group. This holds true for the sessions in the simulators as well as those in real traffic (figure 4). The drivers' self-ratings of their driving performance is significantly higher than that of the psychologists and the driving instructor. This effect of overestimating the own performance is more profound in the group of elderly drivers. This is also true for the sessions in the driving simulators. Consequently, the correlations regarding the driving test in real traffic between the drivers and the psychologists (r=.447) as well as between the drivers and the driving instructor (r=.552) are only medium

range. In contrast, the correlation between the trained rater and the driving instructor was distinctly higher (r=.759).The correlation of results achieved in the high-fidelity simulator and in the compact simulator is very high. The correlation of results achieved in the simulator and in real traffic is slightly higher for the high-fidelity simulator (r=.772, figure 5, left side) than for the compact simulator (r=.716; figure 5, right side).
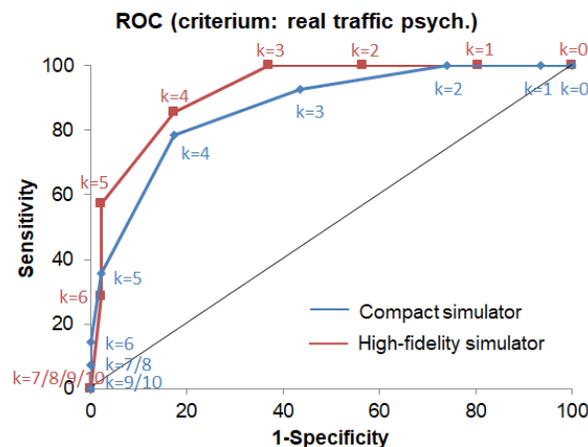

Figure 6. Comparison of ROC-curves of the high-fidelity simulator and the compact simulator. k reflects the criterion up to which the driving test is passed.

The ROC curve (Receiver Operating Characteristics) is a diagram that shows the rate of correctly classified positives against the rate of false positives. The curve illustrates the relationship between sensitivity (relation of the number of persons who failed the driving test in simulation to all persons who failed the driving test in real traffic) and false alarms (1-specificity: relation of the number of persons who failed the driving test in simulation to all persons who passed the driving test in real traffic). Therefore, the curve reflects the diagnostic quality of the driving assessment in the simulators. By modifying the criterion for passing the driving test in the simulator, the ratio of sensitivity and specificity is affected. Figure 6 compares the ROC of the high-fidelity simulator with the ROC of the compact simulator. The aim is to achieve high sensitivity while maintaining high specificity. The discriminative quality of the compact simulator is somewhat lower compared to the high-fidelity simulator. The best trade-off between sensitivity and specificity is given for the criterion k=4. That means that driving tests up to a FtD rating of 4 are rated as „passed", FtD ratings from 5 result in failing the test.

CONCLUSIONS

To sum it up, the results of the validation study show that the described driving assessment is appropriate for assessing fitness-to-drive of elderly drivers. This applies for both configurations of driving simulators, i.e. the high-fidelity simulator as well as the compact simulator. Besides the sound diagnostic quality and prediction accuracy of the driving assessment in the simulator, the interviews with the participating drivers show that the driving test in the simulator is highly accepted. The described method of assessing fitness-to-drive using a driving simulator is more predictive for the driving competence in real traffic compared to traditional psychometric testing

(e.g. Jongen et al. 2015). Regarding the minimum requirements on simulator configuration, a quite cost-efficient compact simulator is sufficient. Driving tests in a compact simulator assure a nearly equivalent predictive quality as in a high-fidelity simulator.

ACKNOWLEDGMENTS

REFERENCES

Aksan, N., Hacker, S.D., Sager, L., Dawson, J., Anderson, S., & Rizzo, M. (2017). Correspondence between Simulator and On-Road Drive Performance: Implications for Assesssment of Driving Safety. Geriatrics, 1, 8.

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

Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975). Mini-Mental State (a practical method for grading the state of patients for the clinician). Journal of the Psychiatric Research, 12, 189-198.

Hoffmann, S. & Buld, S. (2006). Darstellung und Evaluation eines Trainings zum Fahren in der Fahrsimulation. Integrierte Sicherheit und Fahrerassistenzsysteme, VDI-Berichte, 1960, 113-132.

Jongen, s., Perrier, J., Vuurman, E.F., Ramaekers, J.G, & Vermeeren, A. (2015). Sensitivity and Validity of Psychometric Tests for Assessing Driving Impairment Effects of Sleep Deprivation. PLos One, 10, e0117045.

Karthaus, M. & Falkenstein, M. (2016). Functional Changes and Driving Performance in Older Drivers: Assessment and Interventions. Geriatrics, 1, 1-18.

Kenntner-Mabiala, R., Kaussner, Y., Jagiellowicz-Kaufmann, M., Hoffmann S. & Krüger, H.-P. (2015). Driving performance under alcohol in simulated representative driving tasks: an alcohol calibration study for impairments related to medicinal drugs. Journal of Clinical Psychopharmacology, 35(2), 134-142.