ORIGINAL INVESTIGATION

Predicting psychopharmacological drug effects on actual driving performance (SDLP) from psychometric tests measuring driving-related skills

Joris C. Verster · Thomas Roth

Abstract

Rationale There are various methods to examine driving ability. Comparisons between these methods and their relationship with actual on-road driving is often not determined.

Objective The objective of this study was to determine whether laboratory tests measuring driving-related skills could adequately predict on-the-road driving performance during normal traffic.

Methods Ninety-six healthy volunteers performed a standardized on-the-road driving test. Subjects were instructed to drive with a constant speed and steady lateral position within the right traffic lane. Standard deviation of lateral position (SDLP), i.e., the weaving of the car, was determined. The subjects also performed a psychometric test battery including the DSST, Sternberg memory scanning test, a tracking test, and a divided attention test. Difference scores from placebo for parameters of the psychometric tests and SDLP were computed and correlated with each other. A stepwise linear regression analysis determined the predictive validity of the laboratory test battery to SDLP.

Results Stepwise regression analyses revealed that the combination of five parameters, hard tracking, tracking and reaction time of the divided attention test, and reaction time and percentage of errors of the Sternberg memory scanning test, together had a predictive validity of 33.4%.

Conclusion The psychometric tests in this test battery showed insufficient predictive validity to replace the on-the-road driving test during normal traffic.

Keywords Driving · SDLP · Psychometric tests · Predictive validity

J. C. Verster (✉)
Utrecht Institute for Pharmaceutical Sciences,
Division of Pharmacology, Utrecht University,
Universiteitsweg 99,
3584 CG Utrecht, The Netherlands
e-mail: j.c.verster@uu.nl

T. Roth
Sleep Disorders and Research Center, Henry Ford Health System,
Detroit, MI, USA

Introduction

The use of central nervous system (CNS) drugs such as anxiolytics and hypnotics may significantly increase the risk of having a traffic accident and related injury (Orriols et al. 2009). The main reason for the increased risk of traffic accidents is that the use of CNS drugs also can have adverse effects such as sedation and reduced alertness, which impair driving ability. Many tests are used to determine whether drugs may impair performance and to judge whether patients using these drugs are fit for driving.

To determine whether a test is suitable to examine fitness for driving, it is important to take note of different models that describe driving behavior, in particular information processing and motivational models. Motivational models focus on the risks drivers are willing to take during driving (Wilde 1982; Näätänen and Summala 1976; Fuller 1984) and state that driving behavior is determined by a conscious cost–benefit analysis between the motives and goals of the trip and safety risks (Ranney 1994). Drivers adapt their driving behavior to their personal or environmental needs. For example, drivers generally adjust their speed under bad weather conditions but take high risks by speeding in case of an emergency. In contrast to motivational models, information processing models analyze driving skills and abilities at a functional level.

Like many behaviors, driving a car is a mix of automatic and controlled behaviors [Schneider and Shiffrin 1977; Shiffrin and Schneider 1977). Rasmussen (1987) explains driving behavior distinguishing three levels of cognitive control: skill-based, rule-based, and knowledge-based behavior (see Fig. 1). Skill-based behavior comprises automatic and effortless routine driving (e.g., changing gear). In contrast, rule-based and knowledge-based behaviors are controlled actions to deal with changing driving circumstances. Rule-based behavior follows prescribed rules (e.g., passing a slower vehicle). With increased driving experience, and if the rules show to be effective, rule-based behavior becomes automatic (skill-based) behavior. If rule-based behavior is not effective, conscious problem solving (knowledge-based behavior) is necessary to master the new situation. When the driver becomes familiar with the new driving situation, behavior starts following rules (rule-based behavior) or may become more or less automatic (skill-based level). A similar model by Michon (1985) explained driving at the strategic (navigation), maneuvering (tactical), and operational level. Performance at the strategic level is predominantly memory-driven, controlled processing, and concerns trip-planning and achievement of goals. Performance at the maneuvering level is environmental/data-driven, controlled processing, and includes normal driving procedures such as passing other cars or reacting to other traffic. At the operational level, behavior is automatic and concerns immediate vehicle control, such as changing gear. Decisions made at the strategic level take minutes, those made the maneuvering level are made within seconds, and at the operational level, decisions are made within 1 s. The relationship between the models of Rasmussen and Michon is shown in Fig. 1.

Thus, depending on the relationship between uncertainly and experience of the driver, driving behaviors can switch

between the three levels of driving behavior and comprise automatic (fast and effortless) or controlled processing (slow and effortful). Various methodologies are applied to determine whether or not a patient is fit for driving when using a CNS drug, ranging from subjective assessments by driving instructors and psychometric testing to driving simulator tests or actual driving tests on public highways.

Cognitive and psychomotor tests are often used to assess driving-related skills and abilities. Generally, no complex equipment is required, and the tests are easy to conduct, relatively cheap, and testing can be done under controlled standardized conditions. Tests are often of short duration, and a variety of skills and abilities can be examined. This makes their use cost-effective and time-efficient. Of vital importance is that the tests measure a valid psychological construct (Parrott 1991a, b, c), are sensitive to the effects of CNS drugs (Hindmarch 1980), and have a clear relationship with driving. Unfortunately, this often is not the case. For example, finger tapping and the Critical Flicker Fusion Test (CFFT) are often included in test batteries, although their relationship to driving or any other real-life task is unclear. The presumed rationale for including these tests is that they are used in other research and showed to be sensitive to drug-induced impairment. Guidelines and recommendations on the methodology and tests to determine driving ability have been published on behalf of the International Council on Alcohol, Drugs and Traffic Safety (ICADTS) (Vermeeren et al. 1993; Tailloires Report 2007; Walsh et al. 2008). There is consensus that tests should be (1) standardized, (2) valid and reliable, (3) be able to differentiate between dose-dependent drug effects, and (4) provide information on skills and abilities that are important during driving such as attention, alertness, vigilance, and psychomotor performance. These tests should cover performance on all levels of driving behavior in order to fully
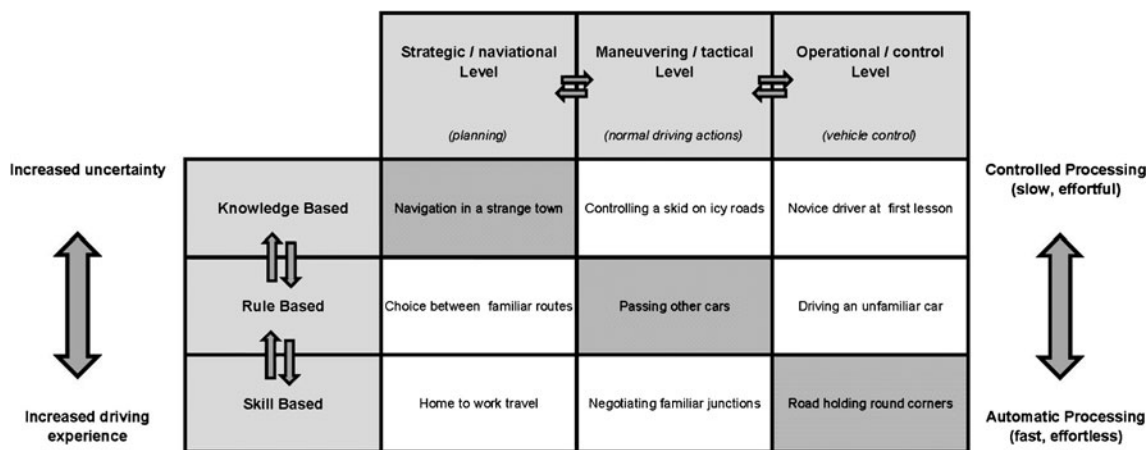

Fig. 1 A model of driving behavior according to Rasmussen and Michon. Experienced drivers operate at the *gray-scaled diagonal* of Fig. 2, whereas novice drivers operate in the *upper right corner* of the figure

understand and judge whether a CNS drug is safe when it comes to driving or to determine whether a patient is fit for driving.

Researchers have combined various psychometric tests and claimed that their test battery predicts actual driving (Fitten et al. 1995; Marottoli et al. 1998; McKnight and McKnight 1999; De Raedt and Pontjaert-Kristoffersen 2001). Their conclusion is often based on the fact that they find a significant relationship between performance on cognitive and psychomotor tests and driving performance, showing a predictive validity up to 85%. In these studies, driving performance was either judged subjectively by a driving instructor or researcher, or driving performance was measured using a driving simulator. Other researchers, however, reported that their tests had no significant predictive validity (Galski et al. 1990; Korteling and Kaptein 1996; Duchek et al. 1998; Bliokas et al. 2011; Devos et al. 2011). The tests that were included in these studies vary greatly. An explanation for differences in the relationship between impairment on cognitive or psychomotor tests versus actual driving impairment is that these tests all have a different sensitivity for drug-induced impairment. This was shown by Moskowitz and Fiorentino (2000) who summarized data on several tests and determined at which blood alcohol concentration (BAC) significant impairment was demonstrated by the majority of reviewed studies. Moskowitz and Fiorentino (2000) reported that BAC levels ranged from 0.1% to 1.0%, depending on the test that was chosen (see Fig. 2).

A study by Myers et al. (2000) illustrates that the choice of tests is of great importance. In their study, a laboratory test battery comprising visual screening, a reaction time task, a split-attention task, a visual organization test, and a verbal and symbolic sign recognition test did not adequately predict on-road driving performance of patients referred for evaluation of their driving skills. However, applying the Useful Field Of View test (UFOV, a three-part test, measuring speed of visual processing, ability to divide attention, and selective attention) a predictive validity of 86% was found.
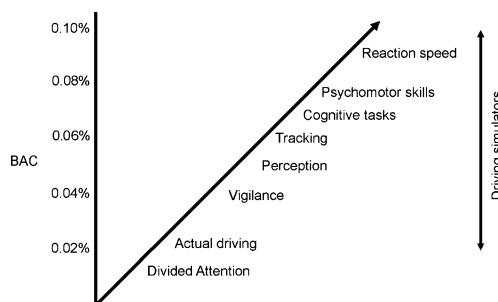

Fig. 2 Blood alcohol concentration (BAC) at which more than half of tests showed significant impairment (Moskowitz and Fiorentino 2000)

Two major limitations of the methodology to assess actual driving performance prevent a fair judgment of the test batteries used in the studies discussed above. First, subjective assessments of driving performance often lack standardization, and reliability and validity are not determined (Fox et al. 1998). Also, simple ratings of driving performance (e.g., visual analog scales or passing/failing judgments) do not sufficiently differentiate between performance levels. A recent comparison also showed that drivers themselves also poorly predict their own driving performance (Verster and Roth 2011a). Second, it can also be questioned if assessments in a driving simulator can adequately predict actual driving. Although today, very advanced driving simulators are available, it remains a test in an artificial environment, the test may be experienced as a game, and in contrast to on-road driving the test poses no real safety risks. Therefore, it is likely that motivation to perform a driving simulator test is different from on-road driving (Verster and Roth 2011a).

As stated by O'Hanlon (1988), a realistic driving test is needed to assess the safe use of drugs by drivers. In their effort to mimic actual driving circumstances, researchers also used closed roads to test driving performance (e.g., Betts and Birtle 1982). Closed road tests often comprise testing specific driving behaviors, such as maneuvering the vehicle along a circuit of cones, gap judgment, parking, or measuring break reaction time (e.g., Tashiro et al. 2005). Although it may be interesting to examine what happens in these special circumstances, the obtained data provides little information on how drivers will behave in normal traffic. Another major disadvantage of this approach is that other traffic is not present during the test. Since interacting with other drivers is an essential element of participating in traffic, this greatly limits the ecological validity of these tests.

Taken together, the subjective nature of the on-road assessments, the use of driving simulator assessments, and the questionable relevance to driving of some of the included tests, probably accounts for the inconclusive results of studies that aimed at developing a suitable test battery to replace actual driving tests.

Currently, the on-the-road driving test in normal traffic is regarded as the gold standard to examine driving ability (O'Hanlon et al. 1982; Verster and Roth 2011b). This 100-km driving test is performed on a public highway in normal traffic. Participants are instructed to drive with a steady lateral position within the right traffic lane and a constant speed (95 km/h). The primary outcome measure is the Standard Deviation of Lateral Position (SDLP), i.e., the weaving of the car. A disadvantage of the test is that it is time consuming and relatively expensive and requires an instrumented vehicle and trained personnel. Therefore, it would be more efficient to have an easy to use psychometric test battery that could predict actual driving. Cognitive and psychomotor tests can provide valuable information about specific impair-

ment caused by a drug. If these tests, alone or in combination, can predict actual driving performance, they could serve as a time-and cost-effective safe alternative for on-the-road driving tests.

Ramaekers (2003) correlated changes in driving performance (SDLP) and performance in seven psychometric tests (tracking, divided attention, reaction time, vigilance, Critical Flicker Fusion test, tapping, and memory). Analyzing data from two studies (N=32 in total), he found that correlations were at best relatively modest (r=0.2 to 0.4). Correlation was highest between SDLP and the tracking test.

In the current study, data from a larger sample of healthy volunteers was analyzed, who participated in three studies examining the effects of various CNS drugs on both on-the-road driving and a number of psychometric test. The objective was to determine to what extend performance on the psychometric tests predicts on-the-road driving performance (SDLP), alone or in combination.

Methods

Data from three studies was used to compose the current data set (Verster et al. 2002a,b, 2003a,b, 2006). The Medical Ethics Committee of the University Medical Center Utrecht approved the study protocols, and subjects were treated according to ICH guidelines for Good Clinical Practice and the Declaration of Helsinki and its amendments.

Participants

A total of 96 healthy male and female volunteers completed the three studies. Mean (SD) age was 23.9 (2.2) years old. Written informed consent was obtained before their inclusion. Subjects were medically screened, used no concomitant medication other than acetaminophen and oral contraceptives, and they had no history of alcohol or drug abuse. Before the start and at the end of the studies, blood chemistry and hematology and urinalysis were determined, and a 12-lead ECG was recorded. All assessments were within normal limits. To confirm compliance, at all visits, subjects were tested on the presence of alcohol and drugs of abuse (amphetamines, barbiturates, cannabinoids, benzodiazepines, cocaine, and opioids). In addition, female subjects underwent β-HCG pregnancy tests. None of the subjects were positive on any of these tests. Subjects possessed a valid driver's license and drove more than 5,000 km/year during each of the past 3 years. A thorough discussion on the inclusion and exclusion criteria of participants and description of the study designs can be found elsewhere (Verster et al. 2002a, b, 2003a, b, 2006). In each study, subjects performed an on-road driving test and a psychometric test battery. The design of the studies was set up in a balanced manner: half of the subjects first performed the driving test, whereas the other half first performed the psychometric test battery.

Treatments

A variety of treatments were tested in the three studies, including alcohol, hypnotics, anxiolytics, analgesics, and antihistamine drugs. Study 1 (part 1) examined the residual effects of zaleplon (10 and 20 mg) and zolpidem (10 and 20 mg), 4–6 h after middle of the night administration (Verster et al. 2002a). The results were compared to a single dose of alcohol (a blood alcohol concentration of 0.05%) and alcohol-placebo (study 1, part 2). Study 2 examined the acute effects bromfenac (25 and 50 mg), oxycodone with paracetamol (5/325 and 10/650 mg), and alprazolam (1 mg) (Verster et al. 2002b, 2006). Study 3 examined the acute (day 1) and subchronic (day 4) effects of levocetirizine (5 mg) and diphenhydramine (50 mg) (Verster et al. 2003a, b). In all studies, results were compared with those obtained in a placebo condition.

The on-the-road driving test

Standardized 100-km driving tests (O' Hanlon et al. 1982; Verster and Roth 2011a) were performed on a primary highway during normal traffic, between the cities of Utrecht and Arnhem. A camera, mounted on the roof of the test vehicle, measured the vehicle's lateral position relative to the road delineation. Participants were instructed to drive with a steady lateral position within the right traffic lane while maintaining a constant speed of 95 km/h (60 mph). The amount of weaving of the car, measured by the standard deviation of the lateral position (SDLP, centimeter), is the primary outcome parameter. The standard deviation of speed (kilometer per hour) is a secondary parameter. Duration of the driving test was approximately 75 min.

Patients were allowed to deviate from the instructions to overtake a slower-moving vehicle in the same traffic lane. The vehicle's speed and lateral position were continuously recorded. A licensed driving instructor who had access to dual controls sat in the right front seat, guarding the subject's safety. Tests could be terminated if the driving instructor or the subject felt it was unsafe to continue. Before disclosure of treatment blinding, the data were edited off-line to remove data that were disturbed by extraneous events (e.g., overtaking maneuvers and traffic jam), and SDLP values were calculated.

Psychometric tests

All tests were computerized and developed from ERTS (Beringer 1992). Subjects were seated in a soundproof test

room, in which the luminosity was kept constant during the study. The test battery consisted of validated and reliable standardized tests measuring different aspects of memory functioning (Sternberg memory scanning test), psychomotor performance (tracking test), attention (divided attention test), and information processing (DSST). Quality of performance was assessed by recording the percentage of errors on the tests. Before the start of the actual studies, participants were trained on the tests to attain baseline performance and to become familiar with test procedures.

Sternberg memory scanning test

This test is designed to measure various aspects of working memory. Working memory is especially important at the strategic/navigational level of driving behavior (e.g., remembering which side-way should be taken).

After learning a memory set of one to five digits (ranging from 0 to 9), a single digit (or probe) was presented on the computer screen. Subjects were instructed to indicate by button press whether the probe was part of the memory set (presented, right hand button) or not (not presented, left hand button). A total of 100 different memory sets were presented. Mean reaction time (RT, millisecond) and percent errors are the parameters of interest. Time on task is ±12 min.

Continuous tracking test

This test is designed to measure psychomotor coordination. Tracking skills are especially important at the operational/control level of driving behavior (e.g., keeping the car in a steady position within the lane).

Subjects were instructed to keep an unstable bar in the middle of a horizontal plane by counteracting or reverse its movements with the aid of a computer mouse. If the unstable bar hit the edges of the plane, they had to start again. An easy version ($\lambda=150$) and a hard version ($\lambda=300$) of the tracking test were conducted (lambda is a measure of instability of the moving bar: the higher lambda, the more difficult the task). The parameter for tracking accuracy is the root mean square (R.M.S.) of the deviation of the unstable bar. Time on task is ±8 min.

Divided attention test

Divided attention is especially important at the maneuvering/tactical level of driving behavior (e.g., passing another car while changing gear).

Simultaneously, the easy tracking test (right hand) and a memory scanning test (fixed memory set size of four digits, left hand) were performed. Subsequently, digits (ranging from 0 to 9) were presented on the computer screen, and participants had to indicate by button press whether the

digit was part of the memory set or not. Parameters are the R.M.S., the mean reaction time (milliseconds) and percent errors. Time on task is ±8.5 min.

Digit symbol substitution test

This test is designed to measure information processing, an ability that is important at all levels of driving behavior.

The digit symbol substitution test (DSST) is a paper-and-pencil test in which subjects have to match accompanying symbols and digits. Subjects were instructed to complete as many pairs as possible within 90 s. Different versions of the tests were used every test day. The number of correct completed pairs was the variable of interest.

Statistical analysis

For each treatment, difference scores from placebo were calculated. This was done for each dependent variable. The data of the three studies were combined into a single dataset to allow comparisons between changes in the primary parameter of the driving test (SDLP) and difference scores of the psychometric test variables. Pearson r correlations were computed (significance was established if $p<0.05$, two-tailed). The analyses were done for all data together, as well as for the individual drug treatments. Finally, a step-forward regression analysis was conducted to determine if a combination of test variables adequately predicts driving performance (SDLP).

Results

Data from $N=604$ driving tests ($N=96$ participants) and corresponding psychometric tests was available for the statistical analysis. For the difference scores from placebo, a total of $N=431$ comparisons could be made. Tables 1 and 2 show the correlations and their significance between difference scores from placebo for SDLP and parameters of the psychometric tests.

Best correlations are found for tracking and the divided attention test. However, the predictive validity of the parameter that correlates best with SDLP (tracking in the divided attention test) is only 22%. Data for individual drugs show comparable results, but not consistently. The relative small number of subjects for individual drugs may have had an impact on the observed correlations. Also, no clear dose–response relationship was observed. For example, performance on the tracking tests and tracking in the divided attention test correlated significant with SDLP of the low dose of bromfenac (25 mg) but not with SDLP of the high dose (50 mg). To determine whether a combination of test parameters would show a higher predictive validity to changes in

Table 1 Difference scores from placebo for the tracking test and divided attention test

Treatment	N	ΔSDLP	Tracking test		Divided attention test		
			Easy	Hard	Tracking	RT	Errors (%)
Zolpidem (20 mg)	30	+10.41[a]	0.694**	0.602**	0.660**	0.599**	0.445*
Alprazolam (1 mg)	18	+9.62[a]	0.447	0.875**	0.693**	0.372	−0.025
Zolpidem (10 mg)	30	+3.35[a]	0.245	0.387*	0.250	0.348	0.282
Diphenhydramine (50 mg, day 1)	48	+2.74[a]	0.252	0.285*	0.255	0.012	0.238
Oxycodone/paracetamol (10/650 mg)	18	+1.87	0.246	0.630**	0.496*	0.262	0.280
Diphenhydramine (50 mg, day 4)	48	+1.56[a]	−0.133	0.031	0.224	−0.166	0.014
Alcohol (BAC=0.05%)	29	+1.03[a]	0.185	0.072	0.065	0.298	0.295
Bromfenac (50 mg)	18	+0.51	0.225	0.241	0.092	0.056	0.337
Levocetirizine (5 mg, day 4)	48	+0.46	0.167	0.330*	0.397**	0.187	0.286*
Zaleplon (10 mg)	30	−0.45	0.229	0.201	0.442*	0.265	0.031
Zaleplon (20 mg)	30	+0.43	0.039	0.292	0.079	0.289	−0.176
Levocetirizine (5 mg, day 1)	48	+0.22	0.029	0.459**	0.229	0.263	0.025
Oxycodone/paracetamol (5/325 mg)	18	−0.65	0.598**	0.719**	0.536*	0.286	0.257
Bromfenac (25 mg)	18	−0.66	0.545*	0.415	0.539*	0.109	−0.050
Overall	491	+2.09	0.412**	0.444**	0.470**	0.420**	0.219**
			N=431	N=431	N=431	N=431	N=431

RT reaction time, SDLP standard deviation of lateral position, N number of subjects

*$p<0.05$, **$p<0.01$; significance of correlation

[a] SDLP differed significantly from placebo

SDLP, stepwise regression analyses were conducted. The DSST was excluded after a first analysis because the test did not contribute to the predictive validity. Table 3 shows the results of the regression analysis including the remaining test parameters.

The combination of five parameters, hard tracking, tracking, and reaction time of the divided attention test, and reaction time and percentage of errors of the Sternberg memory scanning test, together had a predictive validity of 33.4%. The parameters easy tracking and percentage of

Table 2 Difference scores from placebo for parameters of the Sternberg memory scanning test and DSST

Treatment	N	ΔSDLP	Sternberg memory scanning		DSST
			RT	Errors (%)	
Zolpidem (20 mg)	30	+10.41[a]	0.713**	0.344	−0.623**
Alprazolam (1 mg)	18	+9.62[a]	0.510*	0.391	N/D
Zolpidem (10 mg)	30	+3.35[a]	0.333	0.256	0.034
Diphenhydramine (50 mg, day 1)	48	+2.74[a]	0.156	0.412**	N/D
Oxycodone/paracetamol (10/650 mg)	18	+1.87	0.375	0.276	N/D
Diphenhydramine (50 mg, day 4)	48	+1.56[a]	0.050	−0.093	N/D
Alcohol (BAC=0.05%)	29	+1.03[a]	0.261	0.050	−0.070
Bromfenac (50 mg)	18	+0.51	−0.071	0.247	N/D
Levocetirizine (5 mg, day 4)	48	+0.46	0.055	−0.225	N/D
Zaleplon (10 mg)	30	−0.45	0.308	−0.300	−0.267
Zaleplon (20 mg)	30	+0.43	0.326	−0.106	−0.148
Levocetirizine (5 mg, day 1)	48	+0.22	0.046	−0.155	N/D
Oxycodone/paracetamol (5/325 mg)	18	−0.65	0.349	0.313	N/D
Bromfenac (25 mg)	18	−0.66	0.193	0.113	N/D
Overall	491	+2.09	0.435**	0.150**	−0.396**
			N=431	N=431	N=149

RT reaction time, SDLP standard deviation of lateral position, DSST digit symbol substitution test, N number of subjects, N/D test was not included in test battery

*$p<0.05$, **$p<0.01$; significance of correlation

[a] SDLP differed significantly from placebo

Table 3 Results from the stepwise regression analysis to determine which test parameters predict changes in SDLP

Model	Included variables	R	R^2	Adjusted R^2	Significance
1	Tracking divided attention test	0.470	0.221	0.219	$p=0.0001$
2	Tracking divided attention test RT Sternberg memory scanning test	0.531	0.282	0.279	$p=0.0001$
3	Tracking divided attention test RT Sternberg memory scanning test Tracking (hard condition)	0.562	0.316	0.311	$p=0.0001$
4	Tracking divided attention test RT Sternberg memory scanning test Tracking (hard condition) Error percent Sternberg memory scanning test	0.578	0.334	0.327	$p=0.001$
5	Tracking divided attention test RT Sternberg memory scanning test Tracking (hard condition) Error percent Sternberg memory scanning test RT divided attention test	0.585	0.342	0.334	$p=0.020$

RT reaction time

errors in the divided attention test did not significantly contribute to the predictive validity of the model.

Discussion

The objective of the presented analyses was to determine to what extent performance on the psychometric tests predicts on-the-road driving performance. The data show that our laboratory test, composed of a variety of skills and abilities related to driving has insufficient power to predict on-road driving performance, as expressed in SDLP. Although some skills such as tracking performance correlate well with road tracking (SDLP), the tests measuring tracking skills are not suitable to replace on-road testing. Tables 1 and 2 show that impairment found on the road (i.e., significant SDLP increment relative to placebo) is not consistently seen for parameters of the psychometric tests. Also, sometimes skills related to driving are unimpaired when tested in isolation, whereas the on-road driving assessment shows significant impairment. Taken together, the results of the analyses do not advocate replacing on-road driving assessments by psychometric tests measuring related skills and abilities.

The studies that were used for the analyses included a limited number of psychometric tests. These tests were carefully selected, based on the models of driving performance described in the introduction of this paper, and the quality of the tests. It was aimed to include tests that had a clear rationale and measured psychological constructs that are relevant to driving. For example, the Sternberg memory scanning test is based on a sound theory of short-term memory and information processing theory, and a comprehensive background literature is available on the test and its

parameters (Sternberg 1966, 1975). One of the fundamental propositions of information processing theory is that the capacity of the cognitive systems is limited (Schneider and Shiffrin 1977; Shiffrin and Schneider 1977). Dual tasks, such as the divided attention test included in the present test battery, are used to investigate how subjects allocate resources among the tasks and how this affects performance. In the divided attention test, attention is divided among two components that are vital to driving a car. That is, tracking places demands on motor-related resources, whereas memory scanning places demands on working memory. It is therefore not surprising that its parameters, tested individual or in the divided attention test, correlate significant with SDLP, and do so better than parameters from tests that are less clearly related to driving.

This comparison of psychometric test results with SDLP, measured on the road in normal traffic, confirms previous findings by Ramaekers (2003) that results from psychometric tests poorly predict SDLP. The current analysis also found that tracking tests showed the best correlation with SDLP. Combining the test parameters increases the predictive validity only by about 10%, relative to individual tests.

In the past, Volkerts et al. (1992) compared performance on the on-the-road driving test with measurements in a driving simulator. The TS2 driving simulator test consisted of a number of subsequent curve-following maneuvers (road tracking), while simultaneously, subjects had to react to visual signs by button press. The number of correct maneuvers and reaction time were the two outcome measures. Treatments (lormetazepam 1 mg, oxazepam 50 mg, and placebo) were administered and subjects performed an on-road driving test and simulator driving test the morning following night 1 and night 2. On both test

days, SDLP was significantly increased after both oxazepam and lormetazepam in the morning session, but not in the afternoon. In contrast, in the driving simulator, the drugs did not significantly impair performance, and no significant correlation was reported between its parameters and SDLP. This study showed that driving simulator results poorly predicted SDLP. It would be interested to see if SDLP measured in more sophisticated driving simulators is related to on-road assessment of the same parameter. Future studies should also examine other psychometric tests and their relationship to actual driving.

The results of our analyses do not mean that psychometric tests are useless or can be omitted when evaluating potential adverse effects of CNS drugs. On the contrary, it is of importance that these tests are conducted because they provide supportive and or additional information to that obtained by on-road assessments. Whereas SDLP can be regarded as a measure of overall vehicle control, it does not give information on the specific skills and abilities that led to performance impairment. Information on which skills and abilities are more or less impaired can be obtained by using a suitable psychometric test battery. In addition, psychometric tests enable testing skills and abilities related to risk taking and avoidance that due to safely limitations cannot be conducted in normal traffic.

References

Beringer J (1992) ERTS: Experimentelles Run Time System (3.17). Institut für Psychologie, Darmstadt

Betts TA, Birtle J (1982) Effects of two hypnotics on actual driving performance next morning. Br Med J 285:852

Bliokas VV, Taylor JE, Leung J, Deane FP (2011) Neuropsychological assessment of fitness to drive following acquired cognitive impairment. Brain Inj 25:471–487

De Raedt R, Pontjaert-Kristoffersen I (2001) Short cognitive/neuropsychological test battery for first-tier fitness-to-drive assessment of older adults. Clin Neuropsychol 15:329–336

Devos H, Akinwuntan AE, Nieuwboer A, Truijen S, Tant M, De Weerdt W (2011) Screening for fitness to drive after stroke: a systematic review and meta-analysis. Neurology 76:747–756

Duchek JM, Hunt L, Ball K, Buckles V, Morris JC (1998) Attention and driving performance in Alzheimer's disease. J Gerontology 53B:P130–P141

Fitten LJ, Perryman KM, Wilkinson CJ, Little RJ, Burns MM, Pachana N, Mervis JR, Malmgren R, Siembieda DW, Ganzell S (1995) Alzheimer and vascular dementias and driving. JAMA 273:1360–1365

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

Fuller R (1984) A conceptualization of driving behavior as threat avoidance. Erogonomics 27:1139–1155

Galski T, Ehle HT, Bruno RL (1990) An assessment of measures to predict the outcome of driving evaluations in patients with cerebral damage. Am J Occup Ther 44:708–713

Hindmarch I (1980) Psychomotor function and psychoactive drugs. Br J Clin Pharm 10:189–209

Korteling JE, Kaptein NA (1996) Neuropsychological driving fitness tests for brain-damaged subjects. Arch Phys Med Rehabil 77:138–146

Marottoli RA, Richardson ED, Stowe MH, Miller EG, Brass LM, Cooney LM (1998) Development of a test battery to identify older drivers at risk for self-reported adverse driving events. JAGS 48:562–568

McKnight AJ, McKnight AS (1999) Multivariate analysis of age-related driver ability and performance deficits. Accid Anal Prev 31:445–454

Michon JA (1985) A critical view of driver behavior models. What do we know, what should we do? In: Evans L, Schwing RC (eds) Human behavior and traffic safety. Plenum Press, New York

Moskowitz H, Fiorentino D (2000) A review of the literature on the effects of low doses of alcohol on driving-related skills. National Highway Traffic Safety Administration, Washington, Report DOT HS 809 028

Myers RS, Kalina TD, Roth DL, Goode KT (2000) Relation of useful field of view and other screening tests to on-road driving performance. Percept Mot Ski 91:279–290

Näätänen R, Summala H (1976) Road user behavior and traffic accidents. Elsevier, Amsterdam

O'Hanlon JF, Haak TW, Blaauw GJ, Riemersma JBJ (1982) Diazepam impairs lateral position control in highway driving. Science 217:79–81

O'Hanlon JF (1988) Are actual driving tests necessary for evaluating drug safety? Int Clin Psychopharmacol 3(suppl 1):81–85

Orriols L, Salmi L-R, Philip P, Moore N, Delorme B, Castot A, Lagarde E (2009) The impact of medicinal drugs on traffic safety: a systematic review on epidemiological studies. Pharmacoepidemiol Drug Saf 18:647–658

Parrott AC (1991a) Performance tests in human psychopharmacology (1): test reliability and standardization. Hum Psychopharmacol 6:1–9

Parrot AC (1991) Performance tests in human psychopharmacology (2): content validity, criterion validity, and face validity. Hum Psychopharmacol 6:91–98

Parrott AC (1991b) Performance tests in human psychopharmacology (3): construct validity and test interpretation. Hum Psychopharmacol 6:197–207

Ramaekers JG (2003) Antidepressants and driver impairment: empirical evidence from a standard on-the-road test. J Clin Psychiatry 64:20–29

Ranney TA (1994) Models of driving behavior: a review of their evolution. Accid Anal Prev 26:733–750

Rasmussen J (1987) The definition of human error and a taxonomy for technical system design. In: Rasmussen J, Duncan K, Leplat J (eds) New technology and human error. Wiley, Chichester

Schneider W, Shiffrin RM (1977) Controlled and automatic human information processing: I. Detection search and attention. Psychol Rev 84:1–54

Shiffrin RM, Schneider W (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychol Rev 84:155–171

Sternberg S (1966) High speed scanning in human memory. Science 153:652–654

Sternberg S (1975) Memory scanning: new findings and current controversies. Q J Exp Psychol 27:1–32

Tailloires Report (2007) Guidelines for drugged driving research

Tashiro M, Horikawa E, Mochizuki H et al (2005) Effects of fexofenadine and hydroxyzine on brake reaction time during car driving with cellular phone use. Hum Psychopharmacol Clin Exp 20:501–509

Vermeeren A, de Gier JJ, O'Hanlon JF (1993) Methodological guidelines for experimental drugs affecting driving performance: an international expert survey. Institute for Human Psychopharmacology. University of Limburg, The Netherlands

Verster JC, Roth T (2011a) Drivers can poorly predict their own driving impairment: a comparison between measurements of subjective and objective driving quality. Psychopharmacology (in press)

Verster JC, Roth T (2011b) Standard Operation Procedures for conducting the on-the-road driving test, and measurement of the Standard Deviation of Lateral Position (SDLP). International Journal of General Medicine 4:359–371

Verster JC, Volkerts ER, Schreuder AHCML, Eijken EJE, Van Heuckelum JHG, Veldhuijzen DS, Verbaten MN, Patat A, Paty I (2002a) Residual effects of middle-of-the-night administration of zaleplon and zolpidem on driving ability, memory functions and psychomotor performance. J Clin Psychopharmacol 22:576–583

Verster JC, Volkerts ER, Verbaten MN (2002b) Effects of alprazolam on driving ability, memory functioning and psychomotor perfor-

mance: a randomized, placebo-controlled study. Neuropsychopharmacology 27:260–269

Verster JC, De Weert AM, Bijtjes SIR, Aarab M, van Oosterwijck AWAA, Eijken EJE, Verbaten MN, Volkerts ER (2003a) Driving ability after acute and subchronic administration of levocetirizine and diphenhydramine: a randomized, double blind, placebo-controlled trial. Psychopharmacology 169:84–90

Verster JC, Volkerts ER, van Oosterwijck AWAA, Aarab M, Bijtjes SIR, De Weert AM, Eijken EJE, Verbaten MN (2003b) Acute and subchronic effects of levocetirizine and diphenhydramine on memory functioning, psychomotor performance and mood. J Allergy Clin Immunol 111:623–627

Verster JC, Veldhuijzen DS, Volkerts ER (2006) Effects of an opioid (oxycodone/paracetamol) and an NSAID (bromfenac) on driving ability, memory functioning, psychomotor performance, pupil size and mood. Clin J Pain 22:499–504

Volkerts ER, Van Laar MW, Van Willigenburg APP, Plomp TA, Maes RAA (1992) A comparative study of on-the-road and simulated driving performance after nocturnal treatment with lormetazepam 1 mg and oxazepam 50 mg. Hum Psychopharmacol 7:297–309

Walsh JM, Verstraete AG, Huestis MA, Mørland J (2008) Guidelines for research on drugged driving. Addiction 103:1258–1268

Wilde GJS (1982) The theory of risk homeostasis: implications for safety and health. Risk Anal 2:209–225