EJTIR

Issue 15(2), 2015
pp. 205-225
ISSN: 1567-7141
tlo.tbm.tudelft.nl/ejtir

Comparing a driving simulator to the real road regarding distracted driving speed

Allert Knapper[1]
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, Netherlands.

Michiel Christoph
SWOV Institute for Road Safety Research, Leidschendam, Netherlands

Marjan Hagenzieker
SWOV Institute for Road Safety Research, Leidschendam, Netherlands
Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands.

Karel Brookhuis
Faculty of Technology, Policy and Management, Delft University of Technology, Delft, Netherlands;
Faculty of Behavioural and Social Sciences, University of Groningen, Netherlands.

Relative and absolute validity of a driving simulator were assessed regarding effects on mean speed and speed variation during distracting secondary tasks, and normal driving. 16 participants drove the same route four times, twice in a simulator and twice in the real world. They performed way finding tasks, using either a paper map or a route guidance system, and mobile phone conversation tasks. Furthermore, driving without secondary tasks on other road segments in the two methods was compared. As both mean speed and standard deviations of speed were not equivalent, absolute validity could not be established. However, as effects found in the experimental conditions varied in the same directions, evidence for relative validity was provided. It is concluded that driving performance regarding speed under distracting conditions may validly be researched in the driving simulator employed here.

Keywords: *Driver distraction, driving simulator, field test, mobile phone, navigation system, validity.*

1. Introduction

1.1 Driving simulators and their validity

Driving simulators provide an attractive option for studying driving behaviour. Reasons for this appeal include the fact that driving in a simulator is safe both for the driver and the experimenter (H. C. Lee, Cameron, & Lee, 2003; H. C. Lee et al., 2007), and for other traffic. Furthermore, the experimental control (Carsten & Jamson, 2011; Kaptein, Theeuwes, & Van der Horst, 1996) in terms of traffic, weather and locations, provides a scientifically sound method for studying effects on driving performance, for instance, by in-vehicle technology or roadway design changes. However, despite endeavours to develop simulators that realistically simulate driving, the road, and the surrounding environment (De Waard, Van der Hulst, Hoedemaeker, & Brookhuis, 1999a; Törnros, 1998), their validity is often criticised (e.g., Farber, 1999, but see also De Waard, Van der

1 Corresponding author. A: Jaffalaan 5, 2628 BX Delft, The Netherlands. T: +31 152 788 546.
E: A.S.knapper@tudelft.nl

Hulst, Hoedemaeker & Brookhuis' response, 1999b), for instance due to the lack of danger (Evans, 1991). Such critical stances should not be easily discarded, as it is important to verify that what is studied and found in a simulator is also applicable on the road (Mullen, Charlton, Devlin, & Bédard, 2011; Shechtman, 2010), all the more because of the large number of studies applying simulators.

Various factors can potentially affect validity of data collected in a driving simulator. One effect may stem from learning, as many drivers are not used to a driving simulator (Blana, 1997). Second, being observed may involve effects, as few drivers are used to driving while knowing to be monitored and recorded, and they may adapt their driving style to what they think the observer finds desirable (see also the Hawthorne effect, cf Jones, 1992). Third, drivers who are affected by feelings of discomfort due to simulator sickness, may cause selective drop-out (Davidse, Hagenzieker, Van Wolffelaar, & Brouwer, 2009). This is relevant when investigating, for instance, the effect on drivers' performance when using their mobile phone or the effects of taking medication or drugs on driving performance (Young, Regan, & Lee, 2009).

We therefore need to know how performance in a simulator relates to on road performance. Hence, driving simulator validation studies are needed. In fact, they have been performed all along the development of driving simulators (Blaauw, 1980; McRuer & Klein, 1976) and a number of literature reviews exist (cf Blana, 1997; Hoskins & El-Gindy, 2006; Kaptein et al., 1996; Mullen et al., 2011; Shechtman, 2010). However, no generic method is available to test whether a simulator delivers valid results. This may be due to the circumstance that validity comprises many aspects; the specific driving simulator itself, the tasks studied, the subject populations (sample similarity), research design, and even terminology (Mullen et al., 2011). Strictly speaking, this means that for each simulator, each task investigated should be validated separately (Hoskins & El-Gindy, 2006; Kaptein et al., 1996). However, that would in turn invalidate most reasons for using simulators (i.e., cost and safety).

Then how should a driving simulator be validated? Blaauw (1980, 1982) distinguished between physical and behavioural validity. Physical validity comprises the extent to which a simulator itself resembles an on-road moving car in terms of similarity of controls, layout of the vehicle, and dynamics. Behavioural validity refers to how well changes in drivers' behaviour due to experimental conditions in a simulator resemble changes in real life driving, and is also referred to as predictive validity (Törnros, 1998). For the latter type of validity two directions are often distinguished, namely absolute and relative validity (Blaauw, 1982; Kaptein et al., 1996). Absolute validity is obtained in case the numerical values measured in the simulator and the comparing method are equivalent. Relative validity refers to values changing in the same direction and with comparable amplitude across methods. A hypothetical example may be that blindfolding drivers suddenly while driving may lead to braking in both simulated and real world driving (relative validity), but not with the same braking force (hence no absolute validity). Concerning the usefulness of applying a driving simulator as a method for investigating driving, Törnros (1998) indicates that relative validity may, with care, suffice for generalising to real world driving.

Driving simulator validity can be investigated by comparing driving performance during similar tasks (Blaauw, 1982). The standard for investigating validity is therefore obviously comparing it to on-road driving (Reimer, D'Ambrosio, Coughlin, Kafrissen, & Biederman, 2006; Shechtman, 2010), although some have achieved valuable results comparing simulated driving to self-report tests (Bédard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010; H.C. Lee & Lee, 2005; Reimer et al., 2006), to road crash databases (Yan, Abdel-Aty, Radwan, Wang, & Chilakapati, 2008) and to other simulators (Groeger, Carsten, Blana, & Jamson, 1999; Jamson & Jamson, 2010). Other validation studies compared their high-end simulator to both on-road driving and a lower end simulator (Reed & Green, 1999; Santos, Merat, Mouta, Brookhuis, & De Waard, 2005).

Speed

The current study compares data from a driving simulator study and data from an on-road driving study regarding driving speed during way finding, cognitive distraction due to phone conversations, and baseline driving, in an attempt to determine relative as well as absolute validity. The on-road study was part of a European project (INTERACTION). Specifically for validation purposes, the simulated route was designed to closely match the on-road study track, but the validation study was not an aim of the INTERACTION project. We first discuss results of other studies that aimed at validating speed data. Elaborate reviews including other driving task components such as lateral performance are widely available elsewhere (i.e., Hoskins & El-Gindy, 2006; Mullen et al., 2011; Shechtman, 2010).

Driving speed is an important component of the driving task for several reasons. Firstly, it influences crash severity outcomes; higher speed of impact leads to more severe injuries (Elvik, Christensen, & Amundsen, 2004; Joksch, 1993). Secondly, driving speed is associated with crash risk (Aarts & Van Schagen, 2006). Thirdly, drivers may use speed to keep in control of the driving task demands as described in the Task-Capability Interface model (Fuller, 2005), by slowing down, for instance in adverse weather (Hoogendoorn, Van Arem, Hoogendoorn, Brookhuis, & Happee, 2012). Likewise, drivers being distracted by for example interacting with in-vehicle devices often apply compensatory strategies including reducing speed (Stelling & Hagenzieker, 2012; Young & Regan, 2007).

Table 1 considers a number of studies on the topic of driving simulator validity, specifically focusing on speed related measures reported in those studies. The studies were obtained by an extensive literature search, using internet search engines (Scopus, Web of Science, Google Scholar) and other studies' reference lists. Although it is by no means exhaustive, it does provide a broad view on earlier findings.

Inspection of the table indicates that

1. Night time driving may not be simulated very well in terms of mean speed (McAvoy et al., 2007).
2. Moving base simulators may provide a slight advantage over fixed base simulators in terms of various speed measures. However, for some research questions a moving base will not be cost efficient (De Winter et al., 2009; J. D. Lee et al., 2011).
3. Driving speed in a simulator as compared to real driving may be influenced by types of curves. In easy, high radius curves, simulator drivers may adopt a higher speed (Bella, 2008) or entry speed (Bittner et al., 2002), whereas in more difficult curves, driving speed may not differ (Bella, 2008), and entry speed may be lower (Bittner et al., 2002)
4. Drivers may not be able to estimate and reproduce absolute driving speed correctly, but they are well able to distinguish between faster and slower driving (Groeger et al., 1999; Shinar & Ronen, 2007)
5. Few studies report absolute validity, whereas the majority report indications for relative validity regarding speed related measures.

The picture painted in Table 1 may serve to position the current study, as the employed simulator has not been subject to a formal validation study before. Moreover, it adds to the literature through comparing the effects on speed measures of performing different distracting secondary tasks while driving on the real road to while driving in the simulator.

Table 1. Speed relevant measures found in several studies investigating driving simulator validity.

Study	Dependent variable	Abs val	Rel val	Speed relation	Simulator type	Note
Harms (1996)	Mean speed driven	Yes	Yes	Insignificantly higher in the simulator	Moving base	
Törnros (1998)	Mean speed driven	No	Yes	Higher in simulator	Fixed base	Tunnel driving
Groeger, Carsten, Blana, & Jamson (1999)	Speed estimates	No	Yes	Mixed	Fixed base	
Klee, Bauer, Radwan, & Al-Deek (1999)	Mean speed driven	No	Yes	Lower in simulator	Fixed base	Relative validity not in the paper, but as reasoned by Mullen et al. (2011)
Reed & Green (1999)	SD of speed driven	No	Yes	SD of speed was higher in real driving, but larger effect for age (old subjects had higher SD speed)	Fixed base	
Bittner, Simsek, Levison, & Campbell (2002)	Speed on curve entry	No	No	Simulated speed higher in easy curves, lower in difficult curves	Moving base	
Godley, Triggs, & Fildes (2002)	Mean speed driven	No	Yes	Mixed	Moving base	
Santos, Merat, Mouta, Brookhuis, & De Waard (2005)	Mean and SD of speed driven	No	No	Speed lower in advanced simulator, similar to real driving in simple simulator. SD of speed highest in simple simulator.	Simple and advanced, both fixed base	
Brown, Dow, Marshall & Allen (2007)		No	-	Speed somewhat higher in the simulator, on average	Advanced moving base	Relative validity not specifically in the paper.
McAvoy, Schattler, & Datta (2007)	Mean speed driven	No	No	Mixed	Fixed base	Night, at work zone
Shinar & Ronen (2007)	Speed estimates and reproductions	No	Yes	Higher speed reproductions in simulator	Fixed base single screen	
Bella (2008)	Driving speed	Yes/ No	Yes	Higher where road had weakest curve	Fixed base	Real driving speed from speed gun data. Yes: 9 demanding road sections. No: 2 low demand sections
Jamson & Jamson (2010)	Mean speed driven	(Yes)	Yes	Similar	Low-cost and mid-level (fixed base)	Compared 2 types of simulators
Wang al.(2010)	Mean speed driven	No	Yes	Lower in simulator	Fixed base, car replicated	

Mayhew et al. (2011)	Subjective speeding errors	No	Yes	Similar number of speed related errors	Fixed base (both 1 and 3 screens)	Errors, speed was a subset, somewhat subjective.
Hallvig et al. (2013)	Mean speed driven	No	No	Higher speed in simulator, stable over night/day driving	Moving base	Effects of sleepiness, night driving

Note: the validity scores (yes, no, or absent) in many studies were not explicitly stated in the original publications, and therefore needed to be inferred from the data reported.

1.2 *Overview of the experiment*

In this paper we compare driving speed data obtained in a driving simulator study to speed driven in a real road study. The participants drove a route four times in total; twice on the real road and twice in a driving simulator. They drove while performing tasks of way finding, with a paper map (as opposed to driving with a route guidance system), and while having a phone conversation. We report the results of the comparison of speed parameters in terms of absolute and relative validity of driving simulators as research tools.

We address the following research questions regarding driving speed data from the field test and the driving simulator experiment.

1. To what extent are results on speed from both studies comparable with regard to the way finding conditions (using either a paper map or instructions by a navigation system) and the phone conversation conditions?
2. To what extent are the driving speed results obtained from simulated baseline driving valid in the absolute sense, when compared to driving in the real world?

The first question addresses both the issues of absolute and relative validity. Absolute validity is studied in terms of equivalence of driving speed during the specific conditions, relative validity may be extracted from the direction and amplitude of the effect. The second research question relates to absolute validity regarding driving speed.

2. Method

2.1 *Participants*

In total 21 persons initially participated in the project. However, one participant was excluded because of simulator sickness, and four were not included because the instrumented vehicles' data acquisition system had not recorded both field test drives. The final sample for analysis therefore consisted of 16 paid volunteers, six females and ten males, aged 27-59 (m=37.8, SD=10). They had their drivers' licence for 3-39 years (M=15.8, SD=9.5). All participants indicated to use a navigation system and mobile phone at least once a week while driving, and to drive at least 12,000 kilometres per year. Participants signed an informed consent and before each drive they were explicitly instructed to feel free to stop participating at any time, for any reason.

2.2 *Apparatus*

The instrumented vehicles were either one of four Lancia Ypsilons or a Peugeot 207. The instrumentation consisted of four cameras and several sensors that recorded driver behaviour on each trip. The data were recorded by a computer located in the trunk. The instrumentation included a GPS device that recorded at 1Hz the GPS position, at about five meters precision. GPS data were matched to map data, and included information on speed, direction, and time. GPS speed measures have been reported to be more reliable than a car odometer driving on straight lines at a constant speed, and are therefore regarded an accurate measure for speed (Witte & Wilson, 2004). The computer started automatically when the driver side door was opened, and it would shut down automatically after about 10 minutes of inactivity. Some of the trips were not recorded due to the fact they started while the computer was still shutting down (these were excluded, see 2.1). All vehicles were equipped with a TomTom Go Live 1005 navigation system and Parrot Minikit Slim Bluetooth hands free phoning device.

According to several classifications (i.e., Kaptein et al., 1996; Young et al., 2009), the driving simulator used here may be described as mid-level; it has no moving base (which would make it high-level), nor is it only a desktop computer with a steering wheel. It does consist of a mock-up cabin including real steering wheel, car seat and controls, surrounded by three LCD screens allowing for a 180-degree view of the driving environment. The centre screen resolution is 1920x1080 (HDTV), both side screen are 1360x768 pixels. Refresh rate is 60 fps. The simulator software was developed by STSoftware © and runs on two connected personal computers. The computational vehicle model has three degrees of freedom: X, Y and vertical axis rotation. The model includes a simple combustion engine simulation, simulating a 90 hp car engine. The road contact model is based on the Pacejka 'magic formulae' (Pacejka & Bakker, 1992). The model simulates steering as a result of lateral front tire force, and allows for quite realistic steering. Friction and wind force are also included in the model. Furthermore, brake force is included as a counterforce, and depends on pedal pressure. The user controls did not provide active physical feedback. The simulator manual gearbox was used, the real vehicles were also manual. This is representative for most Dutch cars, and the participants were used to manual driving. The simulator was situated in a 20 degree Celsius air conditioned room in order to minimize potential simulator sickness (Stoner, Fisher, & Mollenhauer, 2011). The simulator is visualised in Figure 1.

Figure 1. The driving simulator

2.3 Driving environment

For the field test a route was selected in the The Hague area, depicted in Figure 2. It started in Leidschendam (X in Figure 2), and the first two kilometres were discarded from the data. These were meant for familiarising the participant with being observed (although the participants were not informed about this). The route in fact started at A in Figure 2, via B to the A4 motorway (C to D), then to E and back to Leidschendam. A to B and B to C are both 50 km/h speed limit urban areas, with B-C being the most busy one, because it consists of a heavy used exit road out of The Hague. C to D is the 100 km/h A4 motorway, D to E contains several different speed limits (50, 70, 80, 100 km/h), and may be best described as an arterial road or a ring road, as it leads from Delft to several motorways such as the A4 and the A13.

Figure 2. Route (© Google Maps), normally starting in X, in the The Hague area.

For the simulator study, the environment was replicated from the field route as accurately as possible in terms of road structure, road signs and layout. However, as the simulator was not designed to replicate reality on a micro level, components like bus lanes were omitted, traffic lights had to be moved and lowered for visibility reasons, and some intersections had a slightly different lay out. Furthermore, no street name signs or signage was included. Buildings, trees and bushes were simulated as available from the standard software database. Figure 3 includes four scenes comparing the simulator to the real world.

Other traffic was programmed to resemble typical 10:00 to 16:00 (light) traffic, which resembled field test traffic (field test drives all started between 9:30 and 15:00). In order to minimize simulator sickness, the route was cut in two parts (X-D and D-E in Figure 2), so participants could rest in between the two routes. Each of the parts started with about two kilometres of roads that allowed for the participants to practice driving in the simulator. These parts were not analysed and differed from the real road to avoid recognising the route from one method to the other.

Figure 3. Several simulated environments alongside their approximate real-world counterparts (© Google Maps; Simulator pictures and real life screenshots have different viewing angles).

2.4 *Experimental design and analyses*

All 16 participants included in this study drove both in the simulator and in the field test. Some differences between both studies pertain to the order of conditions. The field test order normally was the same, whereas the simulated conditions were counterbalanced. In order to control for potential order effects between the two methods, four participants first drove the two simulator drives, the others first completed the field test. Full counterbalancing was not possible due to the practical constraints of the INTERACTION project. There were two exceptions to the normal field test order of conditions: Two of the drives, on the participant's request, started at Delft University (between D and E), but the route then still consisted of the same roads and tasks were carried out at the same sections.

2.5 Procedure and materials

Participants drove an instrumented vehicle instead of their own vehicle for five to six weeks, as part of the INTERACTION project. Regarding the field test, participants performed the first drive after having used the instrumented vehicle for at least one week, so that they were used to driving it. They were invited to come to the SWOV (Institute for Road Safety Research), in Leidschendam, where they were briefed. In the briefing, they were informed that they would have a 42 kilometre drive, with two observers in the vehicle, one in the front seat, and one in the rear seat, and that they would be asked to perform certain secondary tasks while driving (see 2.6). It was emphasised that if for whatever reason the participant felt uncomfortable to perform a certain task while driving, he or she would not have to perform it. The participants were not informed that the front seat observer recorded interactions with other road users and special events, whereas the rear seat observer performed standardised observations such as correctness of speed, lateral performance, distance to other road users, according to the Wiener Fahrprobe methodology (Chaloupka & Risser, 1995). Participants were requested to drive as they normally would do through the entire drive. Also, they were allowed to talk during driving, which would be a sign that they felt at ease, but the observers did try to limit conversation. After returning, the observers interviewed the driver, to discuss specific traffic events that sometimes had occurred during the drive.

The second field test drive was normally planned together with the participant handing in the vehicle, which was about four weeks after the first drive. As in the first drive, participants were briefed first, then drove the route together with two observers while performing some in-vehicle tasks, and had an interview afterwards. Two second drives were postponed due to adverse weather conditions and/or illness.

The two simulator drives were separated by at least seven days. During each visit, participants had to drive two routes (in Figure 2: A to C and D to E), in different orders. During the break between the two routes, participants were requested to complete some questionnaires (results not reported here), and after finishing the second route, participants signed a receipt for a €30 gift voucher.

2.6 In-vehicle tasks

During both the field study and the driving simulator study, participants performed several secondary tasks. Way finding was performed on the A to B sections in both studies, the phone conversations on the D to E segments. B to D in the field test contained no specific tasks, whereas in the simulator, participants programmed a navigation system (B to C) and texted on a mobile (C to D) on this stretch. The trips in which no task was performed (baseline driving) on B to D served as comparison to field test driving. Table 2 provides an overview of the tasks.

Table 2. Overview of the road sections (see Figure 2) and their respective in-vehicle tasks.

Section	Description	Road type (speed limit)	Field test		Driving Simulator		Results in section
			Drive 1	Drive 2	Drive 1	Drive 2	
A-B	Way finding	Urban (50)	Route guidance	Map	Route guidance	Map	3.1
B-C	Baselines	Urban (50)	Baseline 1	Baseline 2	Baseline	Destination entry	3.3
C-D	Baselines	Motorway (100)	Baseline 1	Baseline 2	Baseline	Texting	3.3
D-E	Phone conversation	Ring road (50, 70, 80, 100)	Baseline	Phone conversation	Baseline	Phone conversation	3.2

Note: Each line represents one comparison of driving speed between field test and simulated driving. Parts of the road where no specific task was performed are referred to as baselines.

Way finding

In the first field test drive, on the A to B section (see Figure 4), participants were told to follow the navigation system's route guidance, and, when arriving at B (Figure 4), to follow instructions by the observers for the remainder of the route. On the second drive, A to B consisted of following a paper map route. The paper map was printed from Google Maps, and consisted of both the map and the written instructions. Participants were given a few minutes to study the map before the drive. In both drives, in case participants (almost) took a wrong turn, the observers redirected them and marked one way finding error.

In the simulator, both conditions were counterbalanced across the two drives. Participants were requested to either follow the simulator's simple navigation system's instructions (arrows and spoken instructions) or use the paper map. In the simulator, the paper map consisted of a top view of the environment as it was built in the driving simulator (see Figure 4), containing all intersecting streets, roundabouts and tunnels. Participants were requested to drive from the red arrow to the red star. Similar to the field test, (potential) way finding errors led the experimenter to redirect the participant and register the error. Both maps included landmarks such as intersections, tunnels, and bends, that could aid the driver in finding the way (cf. May & Ross, 2006), and the routes were not particularly difficult.

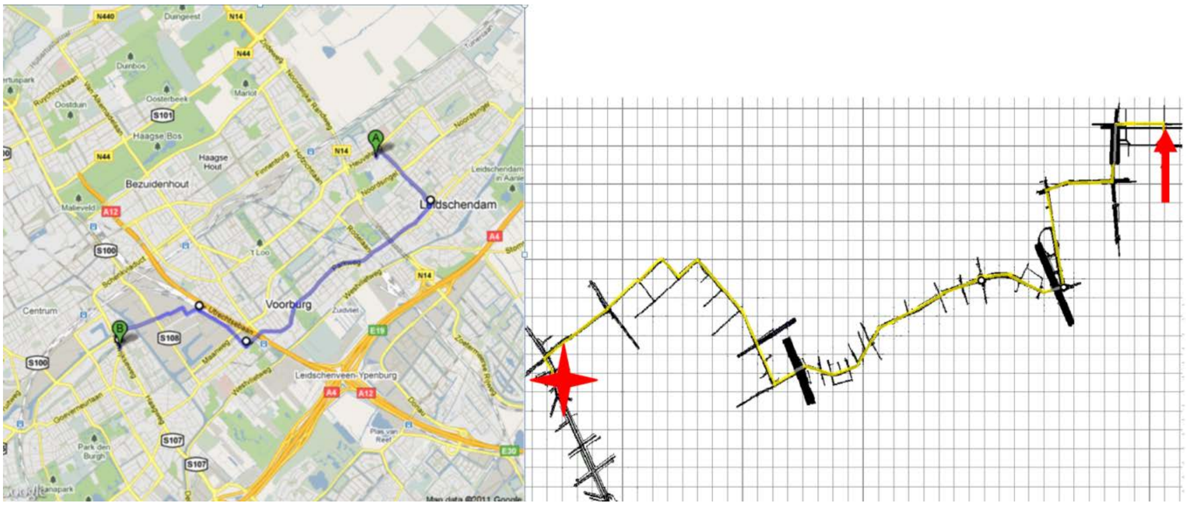

Figure 4. Two paper maps used in this study. The left map (© Google Maps) was used in the field test, and participants were instructed to drive from A to B, whereas the right map was used in the driving simulator, and participants drove from the red arrow (top right) to the red asterisk (left; both signs are enlarged for printing clarity). Note that the first two kilometre differ, which was meant for avoiding participants recognising the route immediately so they would not be engaged in way finding but in remembering the route by heart.

Phone conversation

The phone 'conversation' consisted of a questionnaire containing a total of eight blocks of five questions (four blocks were used in the field test and four in the simulator). Each block included five questions of different categories: list as many (e.g., rivers) as possible, true or not: (i.e., 100 grams of caviar is more expensive than 100 grams of tuna), repeat a sentence (i.e., the home team was playing well until the third quarter of the match), answer the following question ("If I say Jack stole Ann's ball who is the thief?"), and describe (i.e., a friend). The questions were translated to Dutch, and had been based on the Rosenbaum Verbal Cognitive Test Battery, as used by Waugh et al. (2000, see also Rakauskas, Ward, Bernat, Chadwallade, & De Waard, 2005), and on questions used by Perreira et al. (2008). Answers were rated by the remote experimenter.

The hands free phone conversation in the field test was always part of the first of the two drives. During the briefing session, a set of example questions had been read to the participant. During the drive, after having reached the end of the A4 motorway (C-D), the participant was asked whether he or she was ready to have a phone conversation with another experimenter. If so, the rear seat experimenter would instruct a remote colleague by phone to start making the phone call. The phone conversation consisted of a short introduction, including an instruction to watch the road and traffic during answering the questions. The four blocks were administered in two sets. After the first set of two blocks of questions, a second phone call followed about one minute later. Then, approximately one minute later, the next two blocks were administered. During the second drive, no instructions were provided, other than to drive normally.

In the driving simulator, the remaining four blocks of questions were all administered to the participant in a single hand held phone call. Again, the experimenter phoned a colleague on a remote location, who immediately started the phone call. The procedure was similar to the field test, except for the one minute pause in between the two sets of two blocks.

Baseline driving

The B-C and C-D sections in the field test were driven twice without any instruction (except for the necessary route instructions). In the simulator study (see Table 2), on these sections baseline driving was compared to driving while texting and programming a navigation system. As these tasks were not performed in the field test, only the baselines served to be compared to the same sections in the field test here.

2.7 Dependent variables

In both the driving simulator (at 10 Hz) and the field test (at 1 Hz), speed was recorded. For the analyses, we chose to focus on straight road stretches where speed could be selected freely. In order to compare the standard deviation of speed as accurately as possible, the first of every ten simulator data points were included in the analyses.

Next to that, as a coarse measure of secondary task performance for the way finding condition, each (potential) clearly wrong turn (as judged by indicator use) was tallied. This was done both in the field test and in the driving simulator. Furthermore, in both studies response quality of the participant to the mobile phone conversation questionnaire was recorded. If an answer was not rated good or sufficient by the telephonist, it was marked, which resulted in a score for bad answers for each participant, with a maximum of 20 per method.

2.8 Matching data

Based on notes and video recordings, irregularities were removed from the field data, including eventualities such as open bridges, refuelling stops, road works, and traffic jams, in order to ensure comparing driving speed that was as unconstrained as possible. As a result, about 23% of GPS data rows were removed, which together account for about 7% of the total distance covered. Similarly, some "crashes" and road stretches where the phone conversation had started late in the driving simulator data were removed (approximately 9% of data rows). At first sight, the removal percentages seem high, but do include a considerable amount of stopping time. For example, one six-minute refuelling stop may already account for 2 of the 23% removal.

2.9 Statistical analyses

To analyse the differences in the two way finding conditions in both methods, and for the phone conversation induced distraction, we used a 2x2 factorial design (GLM repeated measures), examining mean speed and standard deviation of speed (SD speed). In cases where the residuals were not normally distributed or the data were ordinal (way finding errors) we applied an aligned rank transformation (ART, see Wobbrock, Findlater, Gergle, & Higgins, 2011).

If the main effect for experimental method is reported to be significant, this means that the mean scores on both conditions for each method are different, i.e., absolute validity is untenable. A significant main effect for condition implies that scores on both conditions, averaged across methods for similar conditions, are different. These scores do not reveal much information about validity, but do show whether an experimental effect was present. A non-significant interaction would indicate relative validity, as no differences would have been observed between effects in the simulator and in the field test, in other words, the same trend is found in both methods (Shechtman, Classen, Awadzi, & Mann, 2009).

As two baseline conditions were obtained in the field test, to be compared with the simulated baselines, we applied a repeated measures ANOVA, if necessary corrected for sphericity violations (based on Mauchly's test being <.05) using the Greenhouse-Geisser adjustment, both for urban driving (50 km/h speed limit) and for motorway driving (100 km/h). The results for the mobile phone questionnaire were assessed using a Wilcoxon Signed-Rank test.

3. Results

3.1 Way finding

The results concerning the way finding data are summarised in Table 3. A substantial main effect for mean speed was found for both the method, indicating higher speed in the simulator, $F_{ART}(1{,}15)=4.70$, r=.49 and the conditions, $F_{ART}(1.15)=14.69$, r = .70, indicating higher mean speed using the paper map. The interaction effect was non-significant, F_{ART}(1,15)=.74, r=.22, indicating that the methods had no different effect on participants' driving speed in the paper map or route guidance conditions.

SD speed revealed a main effect for method, F_{ART}1,15)=12.1, p<.05, r=.67, with more speed variation in the driving simulator, whereas neither the condition (navigation system versus paper map) main effect (F_{ART}(1,15)=1.52, r=.30) nor the interaction effect (F_{ART}(1,15)=.11, r=.09) showed significance.

The results did show that participants made more route errors while driving with the paper map, as compared to driving with a route guidance device, F_{ART}(1,15)=26.21, p<.05, r =.80 , but the average number of errors for the methods did not differ (F_{ART}(1,15)=.28, r=.13). Neither an interaction effect (F_{ART}(1,15)=.024, r=.04) was found, indicating that the effects were similar in both methods.

Table 3. Descriptive statistics of the several measures in the different conditions and methods.

	Field test		Simulator		Effects		
	Navigation system	Paper map	Built in route guidance system	Paper map	Method	Condition	Interaction
Mean speed	44.98 (4.26)	47.47 (3.02)	48.32 (5.29)	49.04 (3.63)	*	**	NS
SD speed	5.02 (1.88)	4.47 (1.97)	7.24 (3.61)	6.43 (2.17)	**	NS	NS
Route errors	.13 (.34)	1.5 (1.41)	0 (0)	1.44 (1.36)	NS	***	NS

Displayed are means and standard deviations (in brackets).
NS = not significant at α =.05, * p<.05, **p<.01, ***p<.001.

3.2 *Cognitive distraction*

Figures 5a and 5b show the results that were obtained in both studies related to cognitive distraction induced by a phone conversation. The drives were performed on a ring road equipped with several speed limits. Regarding mean speed (Figure 5a), on all speed limit regimes a substantial main effect on method was revealed (50: F(1,13)=21.38, r=.79, p<.001; 70: F(1,13)=44.15, r=.89, p<.001; 80: F(1,15)=10.55, r=.64, p=.005; 100: F(1,14)=27.49, r=.81, p<.001), with simulated mean speed being higher on all speed limit zones. The main effect of condition showed significance for the 80 (F(1,15)=20.41, r=.76, p<.001), and the 100 km/h (F(1,14)=29.59, r=.82, p<.001) speed limit, but not for the 50 (F(1.13)=1.192, r=.29, p=.295) and the 70 km/h (F(1,12)=3.099, r=.45, p=.104) speed limit. None of the interaction effects reached significance (50: F(1,13)=4.589, r=.51, p=.052; 70: F(1,13)=.064, r=.45, p=.805; 80: F(1,15)=1.74, r=.32, p=.207; 100: F(1,14)=3.70, r=.46, p=.075).

SD speed (Figure 5b) showed no main effect for method on the 50 km/h zones (F_{ART}(1,13)=.66, r=.22, p=.432), but the effect was considerable in the 70 (F_{ART}(1,12)=16.59, r=.76, p=.002), 80 (F_{ART}(1,15)=10.48, r=.64, p=.006) and the 100 km/h (F_{ART}(1,14)=14.49, r=.71, p=.002) speed limits, with higher SD speeds in the simulator than in the field test. A main effect on condition was absent in the 50 (F_{ART}(1,13)=.25, r=.14, p=.626), 70 (F_{ART}(1,12)=3.45, r=.49, p=.088) and the 80 km/h (F_{ART}(1,15)=3.34, r=.43, p=.088) speed limits, but did show significance in the 100 km/h zones (F_{ART}(1,14)=9.43, r=.63, p=.008), with higher SD speeds for phoning as compared to baseline. We also found a significant interaction effect on the 100 km/h speed limit (F_{ART}(1,14)=4.6, r=.50, p=.049), indicating absence of relative validity here, as in the simulator a baseline driving showed less speed variation, whereas in field test hardly a difference between conditions surfaced. Interaction effects were non-significant on the 50 (F_{ART}(1,13)=1.98, r=.36, p=.183), 70 (F_{ART}(1,12)=1.47, r=.33, p=.248) and the 80 km/h (F_{ART}(1,15)=.08, r=.07, p=.788) speed limits.

Comparing the numbers of questions in the phone conversation questionnaire that were not answered sufficiently, in the simulator (Mdn=1) and in the field test (Mdn=1), the scores were not substantially different (Z=-.577, p>.05, r=-.14).

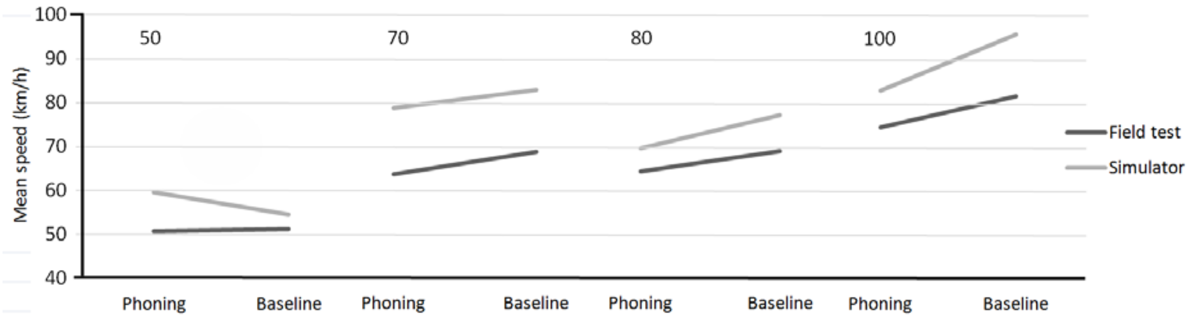

Figure 5a. Mean speed results from the cognitive distraction data, per speed limit. None of the interactions was significant at a=.05.

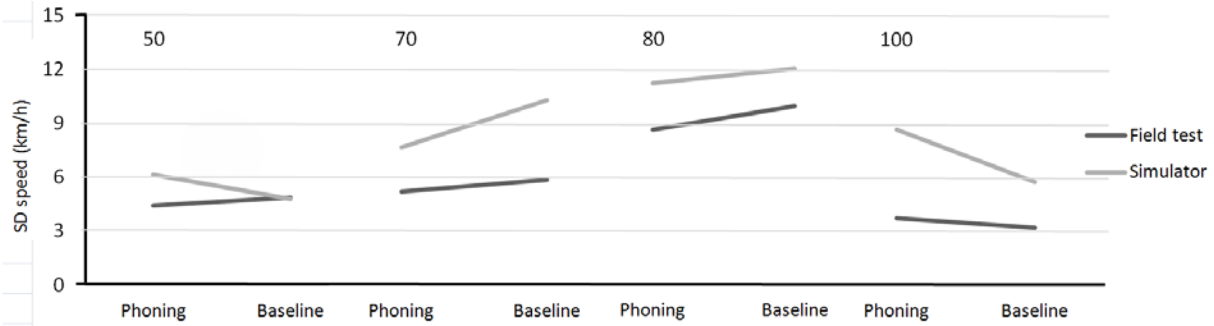

Figure 5b. Speed variation results from the cognitive distraction data, per speed limit. Only the interaction for 100 km/h speed limit was significant at a=.05.

3.3 Baseline driving

Mean and SD speed results were also obtained for driving on an urban 50 km/h speed limit road (B-C, see Table 2), and on a 100 km/h speed limit motorway (C-D in Table 2). The two field test drive results were compared to the simulated baseline drive, as shown in table 4.

Urban mean driving speed showed a substantial main effect (F(2,30)=4.16), post hoc pairwise comparisons revealed that mean speed in the driving simulator was substantially faster than both field test drives (F(1,15)=16.86, r=.73, and F(1,15)=19.44, r=.75). Average SD of speed did not differ significantly (F(2,30)=.252, post hoc pairwise comparisons: F(1,15)=.528, r=.18, F(1,15)=.28, r=.14).

Motorway mean speed was also substantially higher in the driving simulator (F(2,30)=8.80, post hoc comparisons: F(1,15)=5.41, r=.51 and F(1,15)=4.97, r=.50). SD speed measures indicate a higher speed variation in the simulator (F(2,30)=8.80, post hoc comparisons: F(1,15)=14.23, r=.70 and F(1,15)=8.53, r=.60).

Table 4. Descriptive statistics of the measures regarding driving without specific secondary tasks.

Speed limit	Variable	Field test Drive 1	Field test Drive 2	Simulator Baseline	F-Test
50 (urban)	Mean speed	44.32 (4.11)	45.17 (6.20)	51.62 (5.03)	*
	SD Speed	5.59 (2.23)	5.29 (1.34)	4.96 (2.38)	NS
100 (motorway)	Mean speed	101.12 (4.87)	100.85 (4.97)	104.25 (7.15)	***
	SD Speed	3.15 (1.00)	3.18 (1.37)	5.13 (2.47)	***

Displayed are mean scores and standard deviations (in brackets). Note that simulated experimental conditions are not taken into comparison as participants performed destination entries (urban) and texting (motorway) tasks, which are not of interest here.
NS = not significant at α =.05, * $p<.05$, ** $p<.01$, *** $p<.001$.

4. Discussion

We compared field test driving to simulated driving at several levels. Generally, mean speed and variation of speed were higher in the driving simulator than on the real road. Therefore, we found no evidence for absolute validity. However, as results of both studies regarding driver distraction did reveal similar results in the same direction, we found support for relative validity. A detailed discussion of each of the components follows.

4.1 Way finding

Regarding validity of results when investigating way finding using either a paper map or route guidance instructions, results indicate that mean driving speed in the simulator was higher, therefore denying absolute validity. However, the results did vary in the same direction, indicating relative validity. A similar conclusion holds for speed variation. Although the results between methods differed, they did not differ within either method between both conditions, again indicating relative validity. The recorded route errors showed a similar picture, with a main effect for condition, but not for method, and no interaction effect found.

Actually, two methods differed in a number of ways, and care must be taken when drawing conclusions like these. Firstly, while the simulated conditions were counterbalanced across participants, in the field test, the navigation system was always used in the first drive, whereas the paper map was always used in the second drive. This was decided based on evidence that the development of a cognitive map of the driving environment is negatively affected by the turn-by-turn route guidance (Burnett & Lee, 2005; Fenech, Drews, & Bakdash, 2010; Willis, Hölscher, Wilbertz, & Li, 2009). This assertion was not confirmed by the results of our simulator study. Comparing mean speed driven in the map condition in the driving simulator between the two groups (i.e., first drive map, second drive navigation system vs. first navigation system, second map) revealed no significant differences; however, we realise this does not rule out any order effects in the field test.

There were also a number of differences between both maps, for instance regarding street names, scale and familiarity. However, given their relatively simple nature and the presence of useful landmarks on both maps, we do not expect this to have affected our findings regarding relative validity, which is supported by the similar number of errors found. Furthermore, the route guidance function in the simulator, providing two instructions for each turn, was different from a traditional navigation system, because it lacked a moving map and the use of distance to the next turn as a component of the instruction. This may have caused a slight speed increase in the field test, taking into account that simple auditory instructions interfere less with driving performance than more complex instructions, which even led to a speed increase in a study by Dalton, Agarwal, Fraenkel, Baichoo, & Masry (2013).

4.2 Cognitive distraction

During the phone conversation, mean speed in the field test differed substantially from simulated speed for all speed limits, while all interaction effects showed that there was no difference in effect caused by method, again revealing relative, but not absolute validity. With regard to the variability of speed, results were somewhat mixed; within most speed limit zones, a main effect of method occurred, but not on the 50 km/h speed limit stretch. This might be related to these stretches being close to traffic lights, which may have induced somewhat more stop-and-go related variation, specifically in the field test. Moreover, most stretches did not show an interaction effect, except for the 100 km/h speed limit stretch. It is unclear why this occurred, but it might be due to subtle differences in traffic or more difficulties to keep a constant speed in the simulator.

Some cautiousness is warranted here, as cognitive distraction was implemented slightly differently in both methods, through having hands free conversations with a short pause in the field test, but hand held without a pause in the driving simulator. Regarding differences between hand held and hands free driving, a meta-analysis by Caird et al. (2008) suggests there may be a slightly larger speed reduction effect in hand held conversations compared to hands free conversations, which may be reflected in the somewhat lower effects for the field test data (Figure 5a). However, according to Caird et al., the small number of studies that could be incorporated in that meta-analysis, and the fact that one study contributed two effect sizes might have favoured this larger effect for hand held phone conversations while driving.

Furthermore, a different way of counterbalancing was applied in both methods, which also may have had an effect. However, counterbalancing would have probably led to an even stronger effect, in that the field test mean speed results regarding the phone conversation condition would have been rather lower than higher – participants in the phoning condition do not all recognise the route from the first drive, so they compensate more to this lack of familiarity by driving slower. As those effects would be small, they are not likely to induce interaction effects, at least for the 80 and 100 km/h sections. Therefore, despite these counterbalancing issues, we think relative validity is established.

4.3 Baseline driving

When baseline driving was compared between both methods, mixed results for SD speed were obtained. They were similar for both methods in urban driving, but substantially different on the motorway. Furthermore, the results showed differences between mean driving speed, i.e. in the simulator participants drove slightly faster, at both speed limits.

One reason for the absence of absolute validity regarding speed, apart from the 'genuine' differences in driving behaviour and performance itself, may be the fact that measuring speed using GPS is less than perfect, especially when compared to speed recordings very accurately derived from a simulator (cf Godley et al., 2002). Moreover, the differences between mean speed in both methods may disappear if we take into account the fact that the speedometers in the real cars always show a slightly lower speed than the recorded GPS speed, whereas current speed in the simulator was recorded exactly as shown on the simulator speedometer. This may lead to lower recorded field test speed (see the hypothetical example in Table 5), whereas participants in fact kept the same speed. However, this does not clarify the differences in SD speed on the motorway., which suggest we did not fully succeed in mimicking real traffic, that was rather light during most test drives.

Table 5. Example of potential differences between speedometer and reported speed.

	Field test	Driving simulator
Speedometer (hypothetically)	100 km/h	100 km/h
Reported here	95 km/h	100 km/h

4.4 Limitations

Validation of the driving simulator was a secondary objective of both studies, therefore some differences between the two studies could not be avoided. Most are discussed above, and concern differences in experimental setup, tasks, and equipment, and were due to practical constraints. Another source of error may be that in the simulator, some participants reported light simulator sickness, feelings of dizziness or nausea, which may have had an influence on behaviour and performance. However, for older adults, Domeyer, Cassavaugh, & Backs (2013) found that having a (two day) delay between an initial familiarisation drive and a second drive may significantly decrease reported simulator sickness. Given the counterbalanced design, those effects may be ruled out for the simulated effects, but in comparison to field data be the cause of some of the differences.

Likewise, several disturbances such as road works and random error caused by other road users may have had an influence on field test driving, although the direction of these effects can only be guessed and effects are unlikely to be substantial. In addition, the research sample studied may bear some resemblance to the overall population in terms of gender distribution, but their experience in using mobile phones and navigation systems while driving most probably disqualifies them from being representative for all other drivers, although this does not necessarily affect the relationship between effects of distraction in both methods.

4.5 Conclusion

The main significance of the current study is that results concerning driving speed during distracting tasks, as obtained in the current driving simulator, render conclusions in a very similar way and direction as compared to real road (observed) driving, i.e. we found evidence for relative validity with regards to studying effects on speed for distracted driving in this particular simulator. However, in terms of absolute speed, conclusions seem to be much less stable and results should be interpreted cautiously. Still, driver behaviour and performance in traffic psychology research may be validly conducted in at least the simulator applied here.

Acknowledgements

Thanks to dr. Casper Albers for invaluable help regarding the statistical analyses performed here, in particular with the application of the Aligned Rank Transformation (ART), as described in 2.9. This transformation enables researchers to apply nonparametric factorial analyses without the need to abandon familiar ANOVA procedures (Wobbrock et al., 2011). Some of the research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under the project INTERACTION (grant agreement no218560).

References

Aarts, L., & Van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis and Prevention, 38*(2), 215-224.

Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *The American Journal of Occupational Therapy, 64*(2), 336-340. doi: 10.5014/ajot.64.2.336

Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis and Prevention, 40*(3), 1078-1087. doi: 10.1016/j.aap.2007.10.015

Bittner, A., Simsek, O., Levison, W., & Campbell, J. (2002). On-road versus simulator data in driver model development driver performance model experience. *Transportation Research Record, 1803*(02-4103), 38-44. doi: 10.3141/1803-06

Blaauw, G. J. (1980). Driving experience and task demands in simulator and instrumented car: A validation study. (IZF 1980-9).

Blaauw, G. J. (1982). Driving experience and task demands in simulator and instrumented car: A validation study. *Human Factors, 24*(4), 473-486.

Blana, E. (1997). *Driver simulator validation studies: A literature review*: University of Leeds, Institute for Transport Studies.

Brown, T., Dow, B., Marshall, D., & Allen, S. (2007). Validation of stopping and turning behavior for novice drivers in the National Advanced Driving Simulator. Iowa City, Iowa, USA.

Burnett, G. E., & Lee, K. (2005). The effect of vehicle navigation systems on the formation of cognitive maps. In G. Underwood (Ed.), *Traffic and Transport Psychology* (pp. 407-418): Elsevier.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention, 40*, 1282-1293.

Carsten, O. M. J., & Jamson, A. H. (2011). Driving simulators as research tool in traffic psychology. In B. E. Porter (Ed.), *Handbook of Traffic Psychology* (pp. 87-96). London: Academic Press.

Chaloupka, C., & Risser, R. (1995). Don't wait for accidents -- possibilities to assess risk in traffic by applying the `Wiener Fahrprobe'. *Safety Science, 19*(2-3), 137-147.

Dalton, P., Agarwal, P., Fraenkel, N., Baichoo, J., & Masry, A. (2013). Driving with navigational instructions: Investigating user behaviour and performance. *Accident Analysis and Prevention, 50*(0), 298-303. doi: 10.1016/j.aap.2012.05.002

Davidse, R. J., Hagenzieker, M. P., Van Wolffelaar, P. C., & Brouwer, W. H. (2009). Effects of in-car support on mental workload and driving performance of older drivers. *Human Factors, 51*(4), 463-476. doi: 10.1177/0018720809344977

De Waard, D., Van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999a). Driver Behavior in an Emergency Situation in the Automated Highway System. *Transportation Human Factors, 1*(1), 67-82.

De Waard, D., Van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999b). Reply to comments on "Driver behavior in an emergency situation in the automated highway system". *Transportation Human Factors, 1*(1), 87-89.

De Winter, J., De Groot, S., Mulder, M., Wieringa, P., Dankelman, J., & Mulder, J. (2009). Relationships between driving simulator performance and driving test results. *Ergonomics, 52*(2), 137-153.

Domeyer, J. E., Cassavaugh, N. D., & Backs, R. W. (2013). The use of adaptation to reduce simulator sickness in driving assessment and research. *Accident Analysis & Prevention, 53*(0), 127-132. doi: http://dx.doi.org/10.1016/j.aap.2012.12.039

Elvik, R., Christensen, P., & Amundsen, A. (2004). Speed and road accidents. An evaluation of the Power Model. . Oslo: Institute of Transport Economics TOI.

Evans, L. (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold.

Farber, E. I. (1999). Comments on "Driver behavior in an emergency situation in the automated highway system". *Transportation Human Factors, 1*(1), 83-85.

Fenech, E. P., Drews, F. A., & Bakdash, J. Z. (2010). *The effects of acoustic turn-by-turn navigation on wayfinding*, San Francisco, CA.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention, 37*(3), 461-472.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis and Prevention, 34*, 589-600.

Groeger, J. A., Carsten, O. M. J., Blana, E., & Jamson, H. (1999). Speed and distance estimation under simulated conditions. In A. G. Gale (Ed.), *Vision in Vehicles - VII* (pp. 291-299). Amsterdam: Elsevier Scientific (North Holland).

Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Åkerstedt, T. (2013). Sleepy driving on the real road and in the simulator—A comparison. *Accident Analysis and Prevention, 50*, 44-50. doi: 10.1016/j.aap.2012.09.033

Harms, L. (1996). Driving performance on a real road and in a drving simulator: results of a validation study. In A. G. Gale (Ed.), *Vision in Vehicles - V*. Amsterdam: Elsevier Science B. V. (North Holland).

Hoogendoorn, R. G., Van Arem, B., Hoogendoorn, S., Brookhuis, K. A., & Happee, R. (2012). Towards a stochastic model of driving behavior under adverse conditions. *Advances in Human Aspects of Road and Rail Transportation* (pp. 439-448): CRC Press.

Hoskins, A. H., & El-Gindy, M. (2006). Technical report: Literature survey on driving simulator validation studies. *International Journal for Heavy vehicle Systems, 13*(3), 241-252.

Jamson, S. L., & Jamson, A. H. (2010). The validity of a low-cost simulator for the assessment of the effects of in-vehicle information systems. *Safety Science, 48*(10), 1477.

Joksch, H. C. (1993). Velocity change and fatality risk in a crash—A rule of thumb. *Accident Analysis and Prevention, 25*(1), 103-104. doi: http://dx.doi.org/10.1016/0001-4575(93)90102-3

Jones, S. R. G. (1992). Was there a Hawthorne effect? *The American Journal of Sociology, 98*(3), 451-468.

Kaptein, N. A., Theeuwes, J., & Van der Horst, R. (1996). Driving simulator validity: some considerations. *Transportation Research Record*(1550), 30-36.

Klee, H., Bauer, C., Radwan, E., & Al-Deek, H. (1999). Preliminary validation of driving simulator based on forward speed. *Transportation Research Record*(1689), 33-39. doi: 10.3141/1689-05

Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: on-road versus simulated driving. *Accident Analysis and Prevention, 35*(5), 797-803. doi: 10.1016/s0001-4575(02)00083-0

Lee, H. C., Falkmer, T., Rosenwax, L., Cordell, R., Granger, A., Vieira, B., & Lee, A. (2007). Validity of driving simulator in assessing drivers with Parkinson's disease. *Advances in Transportation Studies*.

Lee, H. C., & Lee, A. H. (2005). Identifying older drivers at risk of traffic violations by using a driving simulator: A 3-year longitudinal study. *The American Journal of Occupational Therapy, 59*(1), 97-100.

Lee, J. D., McGehee, D., Brown, J., Richard, C., Ahmad, O., Ward, N., . . . Lee, J. (2011). Matching simulator characteristics to highway design problems. *Transportation Research Record: Journal of the Transportation Research Board*(2248), 53-60. doi: 10.3141/2248-07

May, A. J., & Ross, T. (2006). Presence and quality of navigational landmarks: Effect on driver performance and implications for design. *Human Factors, 48*(2), 346-361.

Mayhew, D. R., Simpson, H. M., Wood, K. M., Lonero, L., Clinton, K. M., & Johnson, A. G. (2011). On-road and simulated driving: Concurrent and discriminant validation. *Journal of Safety Research, 42*(4), 267-275. doi: 10.1016/j.jsr.2011.06.004

McAvoy, D. S., Schattler, K. L., & Datta, T. K. (2007). *Driving simulator validation for nighttime construction work zone devices.* Paper presented at the 86th Annual Meeting of the Transportation Research Board, Washington, DC.

McRuer, D. T., & Klein, R. H. (1976). Comparison of driver dynamics with actual and simulated visual displays. *Transportation Research Record*(611), 46-48.

Mullen, N., Charlton, J., Devlin, A., & Bédard, M. (2011). Simulator validity: Behaviors observed on the simulator and on the road. In D. L. Fisher, M. Rizzo, J. K. Caird & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology*: CRC Press.

Pacejka, H. B., & Bakker, E. (1992). The magic formula tyre model. *Vehicle System Dynamics, 21*(sup001), 1-18. doi: 10.1080/00423119208969994

Pereira, M., Hamama, H., Bruyas, M.-P., & Simoes, A. (2008). *Effect of additional tasks in driving performance: Comparison among three groups of drivers.* Paper presented at the European Conference on Human Centred Design for Intelligent Transport Systems, Lyon, France.

Rakauskas, M. E., Ward, N., Bernat, E., Cadwallade, M., & De Waard, D. (2005). Driving performance during cell phone conversations and common in-vehicle tasks while sober and drunk (pp. 113). Minneapolis: Minnesota Department of Transportation.

Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics, 42*(8), 1015-1037.

Reimer, B., D'Ambrosio, L. A., Coughlin, J. F., Kafrissen, M. E., & Biederman, J. (2006). Using self-reported data to assess the validity of driving simulation data. *Behavior Research Methods, 38*(2), 314-324.

Santos, J., Merat, N., Mouta, S., Brookhuis, K., & De Waard, D. (2005). The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F, 8*, 135-146.

Shechtman, O. (2010). Validation of driving simulators. *Advances in Transportation Studies*(SPEC. ISSUE), 53-62. doi: 10.4399/97888548416596

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: A validation study. *Traffic Injury Prevention, 10*(4), 379 - 385.

Shinar, D., & Ronen, A. (2007). Validation of Speed Perception and Production in a Single Screen Simulator. *Advances in Transportation Studies*.

Stelling, A., & Hagenzieker, M. P. (2012). *Afleiding in het verkeer [Distraction in traffic]*. (R-2012-4). Leidschendam: SWOV Institute for Road Safety Research.

Stoner, H. A., Fisher, D. L., & Mollenhauer, M. A. (2011). Simulator and scenario factors influencing simulator sickness. In D. L. Fisher, J. K. Caird, M. Rizzo & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology: An overview* (pp. 14.11 - 14.20). Boca Raton: CRC Press, Taylor & Francis Group.

Törnros, J. (1998). Driving behaviour in a real and a simulated road tunnel—a validation study. *Accident Analysis and Prevention, 30*(4), 497-503. doi: 10.1016/s0001-4575(97)00099-7

Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics, 53*(3), 404-420. doi: 10.1080/00140130903464358

Waugh, J. D., Glumm, M. M., Kilduff, P. W., Tauson, R. A., Smyth, C. C., & Pillalamarri, R. S. (2000, July 1, 2000). *Cognitive workload while driving and talking on a cellular phone or to a passenger.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2000. Prentice Hall, New Jersey.

Willis, K. S., Hölscher, C., Wilbertz, G., & Li, C. (2009). A comparison of spatial knowledge acquisition with maps and mobile maps. *Computers, Environment and Urban Systems, 33*(2), 100-110. doi: http://dx.doi.org/10.1016/j.compenvurbsys.2009.01.004

Witte, T. H., & Wilson, A. M. (2004). Accuracy of non-differential GPS for the determination of speed over ground. *Journal of Biomechanics, 37*(12), 1891-1898. doi: http://dx.doi.org/10.1016/j.jbiomech.2004.02.031

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). *The aligned rank transform for nonparametric factorial analyses using only anova procedures*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada.

Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis and Prevention, 40*(1), 274-288. doi: 10.1016/j.aap.2007.06.007

Young, K. L., & Regan, M. A. (2007). Driver distraction: A review of the literature. In I. J. Faulks, M. A. Regan, M. Stevenson, J. Brown, A. Porter & J. D. Irwin (Eds.), *Distracted driving* (pp. 379-405). Sydney, NSW: Australasian College of Road Safety.

Young, K. L., Regan, M. A., & Lee, J. D. (2009). Measuring the effects of driver distraction: Direct driving performance methods and measures. In M. A. Regan, J. D. Lee & K. L. Young (Eds.), *Driver Distraction - Theory, effects & mitigation* (pp. 85-105). Boca Raton: Taylor & Francis Group.