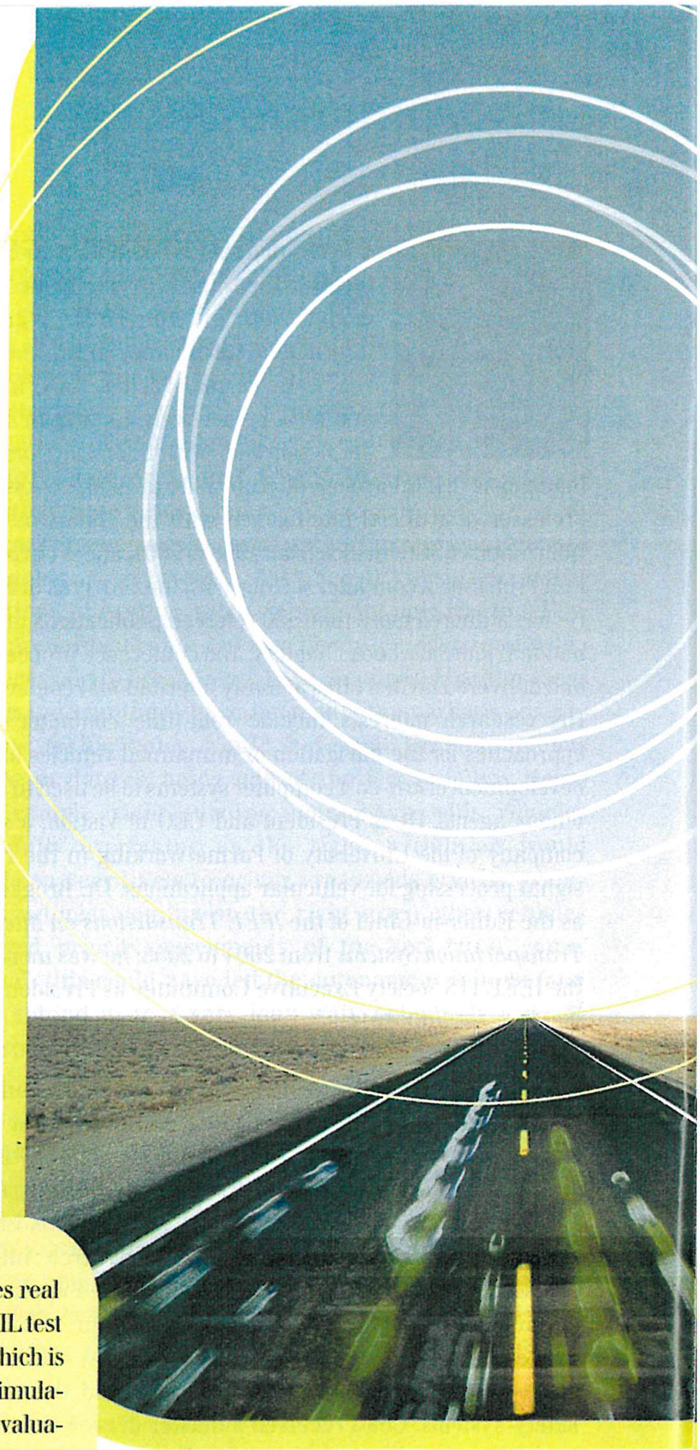
Driving Behavior and Simulator Sickness While Driving the Vehicle in the Loop: Validation of Longitudinal Driving Behavior



***Abstract*—**The Vehicle in the Loop (VIL) is a simulator, which combines real driving experience with the replicability and safety of simulators. In the VIL test setup a real test vehicle is combined with a virtual testing environment which is displayed to the user via a head-mounted display (HMD). In theory, this simulation concept renders the VIL uniquely suited to the development and evaluation of numerous automotive applications, including driver assistance systems. Aiming to assess the extent to which the VIL elicits realistic driving responses, a validation study was performed. In this first validation study the focus was on longitudinal driving behavior. 44 participants performed five common traffic maneuvers in reality and the VIL setup. Simulator sickness was assessed with the simulator sickness questionnaire (SSQ). Descriptive and inferential analyses of the data showed that the VIL achieves relative validity concerning brake pressure and reaction times and absolute validity concerning the steering angle. However, subjects showed longer reaction times and accelerated more smoothly while driving the VIL. One possible explanation for these results could be the presence of simulator sickness. Overall, the study indicates that the VIL represents a suitable testing method for the evaluation and development of driver assistance systems. The study also provided clear indications for the future development of the VIL.

Ines Karl, Guy Berg, Fabian Rüger, and Berthold Färber
Universitaet der Bundeswehr Muenchen, Human Factors Institute, 85577 Neubiberg, Germany

Digital Object Identifier 10.1109/MITS.2012.2217995
Date of publication: 23 January 2013

1939-1390/13/$31.00©2013IEEE

© EYEWIRE

I. Introduction

More than 30,000 people die every year on European roads [40]. *Advanced driver assistance systems* (ADAS) have the potential to reduce the number of traffic accidents and road fatalities. Especially collision avoidance and collision mitigation systems can help the driver to manage safety-critical situations. These systems use various sensor techniques (e.g., video, radar, and lidar sensors) in order to detect critical situations and to assist the driver in avoiding or mitigating a collision with another vehicle. Due to their invariably complex interaction with the environment, ADAS require high standards of system functionality and system validation. Additionally (fail-)safety and error-tolerance are essential in order to guarantee a perfect human-machine-interaction [16], [14]. ADAS that are not fully compatible with environmental or driver demands can have negative consequences for driving comfort, traffic safety, and customer acceptance.

In order to test the functionality of these systems, the realistic and reproducible generation of critical traffic situations represents a significant challenge in the development of ADAS. The investigation of drivers' reactions to system limits and faults also requires appropriate and reliable testing methods.

As with any technical system, ADAS can only be tested within the limits of reality and reproducibility that are stipulated by the testing environment. Hence, the choice of the appropriate testing environment is crucial in the evaluation and subsequent development of ADAS. Generally, three different testing approaches can be distinguished: driving simulators (1), quasi-experimental field studies (2) and field operational tests (3). The suitability of currently prevalent testing methods is limited by a number of factors. On the one hand, while driving behavior in naturalistic driving studies is observed unobtrusively and therefore elicits the highest degree of realistic behavior on part of the test subjects, they are very time and cost intensive. Additionally, substantial effort has to be put into data analysis and special algorithms are needed in order to facilitate data processing (for an example see [1]). Another disadvantage is the lack of control and replicability of specific traffic situations. In contrast, quasi-experimental field studies still offer a relative high degree of face validity combined with at least some control over the traffic situations of interest. However, the testing of some ADAS, in particular that of safety systems such as collision avoidance systems, can be too dangerous when other traffic participants are involved. Furthermore, the repeated replication of specific traffic situations for a larger number of test subjects is very difficult.

On the other hand, driving simulators are a common research tool to investigate advantages and disadvantages of new in-car technologies, including ADAS. The most essential benefits of driving simulators are the safe and reliable replication of traffic situations and the control of environmental circumstances [2], [7]. Chang and Chou [12], for example, used a bus-driving simulator in order to investigate rear-end collision warning systems. However, one problem with driving simulators concerns their validity. To what extent research results obtained in a driving simulator can be transferred to reality highly depends on the aim of the driving simulator study. Whereas new human-machine interfaces can be easily evaluated in driving simulators with even low physical correspondence between the simulated and the real word, the testing of ADAS requires a higher degree of fidelity. However, even in high-fidelity driving simulators with good motion performance, the precise replication of physical driving sensations and vehicle dynamics is difficult and can only be achieved by a few driving simulators such as the simulator of the Daimler AG in Sindelfingen (Germany). The operation of such simulators, however, is very expensive. Additionally, studies in driving simulators are often afflicted by simulator sickness due to either methodological issues concerning the studies' design or to the simulators themselves. Simulator sickness, however, often leads to driving behavior that is dissimilar to driving in the real world [7].

In the *vehicle in the loop* (VIL) test setup a real test vehicle is combined with a simulated testing environment which is displayed to the user via a *head-mounted display* (HMD) [9]. In this simulated environment, objects that are relevant to particular traffic situations, such as pedestrians, cyclists or leading cars can be portrayed. However, at the current stage of development the cockpit and the instrument panel are not displayed in the virtual world. The VIL is not moving on a public road, but on a secured test track. The virtual environment matches exactly the boundaries of the test track, thus ensuring the safety of the driver. One huge advantage of the VIL is that ADAS can be directly tested in a real vehicle with real vehicle dynamics. Additionally, as simulators, the VIL test setup enables the safe and controlled replication of specific traffic situations. Hence, the VIL combines the main advantages of both test methods.

The original VIL was developed by Bock [9] and operated with an augmented reality display. In a first validation study, Bock could demonstrate behavioral validity of his

VIL setup [9]. However, positioning and visualization of the simulated traffic caused various difficulties. Luminance and contrast of the virtual vehicles, for example, were low during direct sunlight. Furthermore, the exact positioning of the vehicles on the street was not always possible. Therefore, the impression emerged at times that the virtual vehicles "flew over" or "sank into" the street [8]. A similar system was used by Moussa, Radwan, and Hussain [35] to investigate left-turn maneuvers at two-way stop-controlled intersections. They installed a fixed camera in the window screen which recorded a live stream of the road ahead. The road was segmented in real time and virtual vehicles were placed on the road in front of the driver. The modified stream was displayed to the driver via an HMD. However, as the video camera recording the scene was held in a fixed position, the field of view of the driver was limited which caused problems while driving on curved segments. Furthermore, virtual objects were misaligned when the road segmentation was flawed.

To address the problems of Bock [8] and Moussa et al. [35], the Carmeq GmbH, as a patent licensee, and the Universitaet der Bundeswehr Muenchen (UniBw) further developed the VIL. Instead of an augmented reality, a virtual reality (VR) was used to simulate the road and traffic environment (see for instance [4] or [44]). Thus, the visual channel is completely disconnected from reality (see Fig. 1).

The VIL setup with a virtual reality offers many advantages regarding the testing and validation of ADAS. However, results from VIL studies can only generalize to reality, if validity can be assumed. Therefore, this paper presents the results of a first validation study based on general driving behavior, using the VIL with a virtual reality.

First, the technical setup of the VIL is explained before the topic of driving simulator validity is discussed in detail. Subsequently, the method and results of our VIL validation study are presented.

II. Theory

A. The Vehicle in the Loop

The VIL is based on the simulation software "Virtual Test Drive" (VTD) which offers the required core functionality [38]. Included are visualization, traffic simulation and sensor model components. The sensor models, which simulate sensors used in the automotive domain, provide an interface to the virtual environment, so that the developer can connect a particular ADAS with the VIL in order to test and evaluate its functionality. Furthermore, a driving interface with different input modalities is provided. Thus, with the VTD the user can drive freely through the virtual world and experience predefined scenarios, which are generated by the traffic simulation module. With the sensor model module, object data from the current environment situation is provided and can be used to control ADAS actuators (see also Fig. 2).

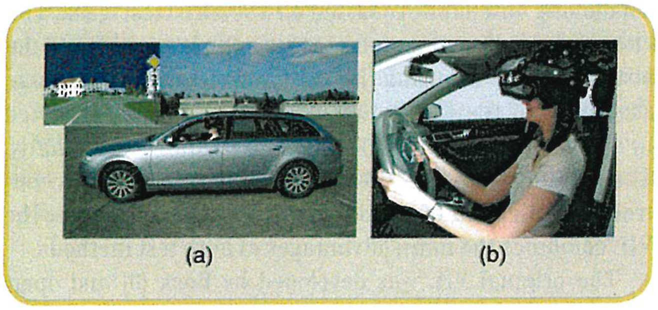


FIG 1 (a) Vehicle in the loop—(b) driver wearing an HMD on which images from the VR environment are displayed (upper left).

When VTD is used in the VIL mode, visualization, positioning and the layout of the virtual world are adopted. The layout of the virtual world is stipulated by the layout of the real test field on which the real vehicle can drive. Thus, the virtual world is superimposed onto the real test track in such a way that for every position in the virtual world an equivalent position on the real test track exists. In order to ascertain the current position of the vehicle during simulation, an inertial measurement unit (IMU) is used. The used IMU is an iTraceRT-F200 from iMAE Navigation with dGPS, which offers a long-time accuracy of two centimeters. As the positions of real and virtual world are equivalent, the vehicle is set to the appropriate position in the virtual world. With the use of a head mounted display for the visualization, the natural field of view of the driver is maintained. However, an additional head tracker to measure the viewing angle is needed. In order to display the correct viewing angle in the HMD, the global gaze direction has to be calculated based on the vehicle orientation and position, as well as the driver head pose (see Fig. 3).

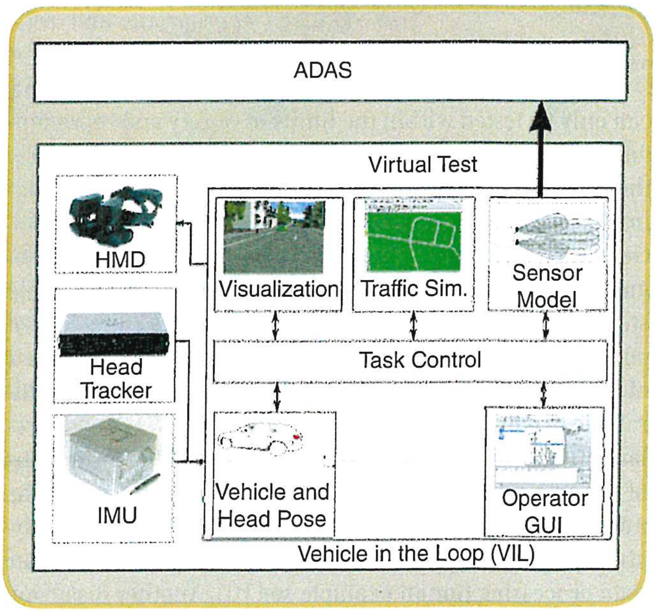


FIG 2 Technical setup of the VIL.

Based on the experience that was gained from the hardware components used in the Bock and Carmeq setup [9], [44], hardware changes were made. In the original design, the head tracker was tethered and restricted in the viewing angle. Therefore, the redesigned system opted for the Personal Space Tracker, a wireless optical motion tracker, which allows for a 360° tracking in the tracker's field of view.

Furthermore, in the original setup the HMD was chosen based on its see-through capability. That HMD featured a small field of view, however, which posed a considerable disadvantage in cross-traffic scenarios. With the introduced VIL virtual reality mode the see-through option was not mandatory anymore. Thus, the stereoscopic NVIS SX111 HMD with a total of 111° FOV was selected for the present setup. However, with the soft- and hardware used in our VIL setup, a time lag of about 150 ms between head movements and the update of the visual scene occurred.

B. Driving Simulator Validity

One of the most essential questions about driving simulators is: To what extent do the results measured in a simulator represent those obtained in the real world? [11]. The research literature of the last 50 years referred to this problem as *simulation validity.* To this date, however, driving simulator validation studies remain rare [7], [11].

Driving simulator validity is a multi-dimensional problem. According to Mudd [36] and Blaauw [6], it can involve physical (*physical validity*) and behavioral (*behavioral validity*) dimensions. Furthermore, Blana [7] points out the importance of subjective experiences and sensations of the people driving in a simulator, which she called *face validity.*

Physical Validity

Blaauw [6] defines physical validity as the physical correspondence between reality and the driving simulator, including layout and dynamic characteristics (p. 473). Physical validity or *fidelity* is considered high, when the simulator reproduces the vehicle handling and driving experiences of real car driving, [17], [19]. Therefore, significant cues have to be represented accurately by the driving simulator to the driver. These include visual, motion, audio, and tactile cues [19]. Visual scenarios, e.g., are usually created by an image generator and projected onto curved or flat screens [28]. Alternatively, HMDs may be used for the presentation of visual cues. Considerable emphasis is also placed on the simulation of vehicle motions, because driving tasks like speed and distance estimation are highly dependent on the perception of different vehicle forces, e.g., acceleration and deceleration [19]. Motion cues are often provided by a motion platform, which allows for the simulation of longitudinal, lateral, and vertical movements [28]. However, even in high-fidelity driving simulators, the perception of movement differs from real driving experiences. To define the fidelity of a specific driving simulator, it is essential to know which experimental artifacts and model errors can emerge during the cue generation process and how these errors might influence the behavior of simulator drivers.

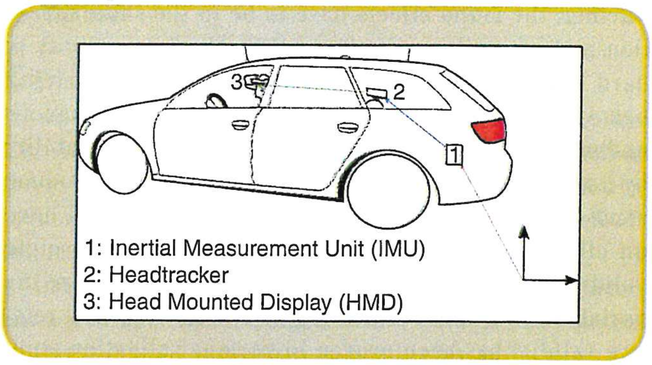


FIG 3 Point of view determination in the global coordinate system.

Another criterion of a successful driving simulator, and thus physical validation, is, according to Harms [20], the absence of simulator sickness.

Behavioral Validity

Although physical validity is important for the test of safety critical ADAS, some have argued that early driving simulator studies focused on it too much [17]. On the other side, behavioral correspondence was rarely mentioned [7], due to the widely-held assumption that behavioral validity will necessarily result from high simulator fidelity [17]. However, according to Blaauw [6], physical and behavioral validity do not have to be related. Hence, when employing simulators as a research tool, one might argue that not only physical validity but also behavioral correspondence needs to be taken into consideration.

Blaauw [6] defines behavioral validity as "the correspondence between the behavior of the human operator in the simulator and in the real, operational system" (p. 473). Thus, behavioral validity is one essential requirement for the generalization of simulator results to real world driving and vital when decisions about the design of future ADAS and other vehicle technologies are made based on simulator studies.

One method to measure behavioral validity is to compare the driving performance and driving behavior in a simulator and a real car during identical tasks and circumstances. In general, two types of behavioral validity can be distinguished. *Absolute validity* is reached when the measured variables have the same numerical values [6]. *Relative validity*, however, requires the comparison between at least two experimental conditions, both in reality and in simulation. Relative validity is achieved when performance differences between experimental conditions are found in reality and in simulation.

Further, the found effects have to be in the same direction and of similar magnitude [17]. Absolute validity is hard to achieve for a number of reasons. First, driving behavior of a single driver is influenced by circumstances and motivational aspects and thus, has a high variability by itself [11]. Second, driving simulators always produce some kind of artificial environment, which could have an effect on driving behavior [37]. Therefore, absolute validity could only be demonstrated for specific driving variables in a few studies (e.g., [6]). In contrast, relative validity has been proven in various validation studies for different driving performance variables, e.g., speed [17], [46], lateral position [46], [6], or mean braking onset [25]. Because of the high variability of driving behavior, Törnros [46] points out that relative validity is sufficient for the application of driving simulators as research tools (p. 497).

However, due to the wide variety of different driving simulators with regard to simulator equipment, software and environment, behavioral validity cannot be automatically transferred from one simulator to another [37]. Furthermore, Kaptein, Theeuwes, and van der Horst [26] emphasize that behavioral validity of a specific driving simulator is highly task dependent. Hence, it is necessary to validate driving simulators before using them as research tools to investigate specific research questions.

Face Validity

In addition to physical and behavioral validity, Blana [7] points out the importance of face validity. According to her, face validity cannot be objectively measured, but it describes how realistic the simulator looks to the subjects. Consequently, subjective judgments are typically used for the determination of face validity. Face validity is improved by a realistic simulator design concerning the provided visual, audio, tactile and motion cues, whereby their aesthetic rather than their physical characteristics are of principle interest. Therefore, while a simulator with physical validity is likely to show face validity, the opposite is not necessarily the case. Hence, although face validity is useful to acquire, it should never substitute physical and behavioral validity.

C. Simulator Sickness

Driving simulators offer a safe method to investigate driving behavior. However, they often encourage simulator sickness, which can negatively influence their behavioral validity [5]. As Silverman and Slaughter [41] or Cobb et al. [13] showed, simulator sickness can lead to a modification of driving behavior in order to reduce simulator sickness symptoms, e.g., reduction of head movements. Additionally, simulator sickness is often responsible for high dropout rates in simulation studies [13].

Simulator sickness is a special type of motion sickness, but less severe and of lower occurrence within the population [30], [29]. Like motion sickness, simulator sickness consists of a variety of different symptoms, for instance nausea, dizziness, vertigo, sweating or eyestrain [30], whereas oculomotor symptoms are the dominating ones [31].

Several theories exist which explain how simulator sickness occurs. The most widely accepted theory is the sensory conflict theory from Reason and Brand [39]. According to this theory, simulator sickness symptoms arise when there is a conflict between the stimuli received from different sensory systems, such as the visual and the vestibular system. For example, in a fixed-based driving simulator the mismatch between visual cues that signal motion and vestibular cues which notify standstill, can lead to simulator sickness. This experience of self-movement without actual physical movement is referred to as vection [43]. However, simulator sickness is not only observed with fixed-based driving simulators, but also with motion-based and head-mounted simulators [45]. One reason is that according to Hettinger and Riccio [23] simulator sickness is often visually induced. For example, high rates of optic flow can lead to more and severe simulator sickness symptoms [23]. Optic flow is the perceived motion of elements in the visual scene that appears from the relative motion of the viewer to the environment [18]. In reality, optic flow provides important information for speed and motion perception. However, optic flow can contribute to simulator sickness if the perceived motion does not correspond to the vestibular stimuli [45]. Hence, the visual scene should include only as many objects that encourage optic flow, like trees, houses and so forth, as are needed in order to provide the perception of motion on the one hand and to reduce simulator sickness on the other hand [45].

Besides optic flow, other technical system factors contribute to the occurrence of simulator sickness [32], [3]. For example, Cobb et al. [13] mentioned that in head-mounted simulators simulator sickness often arises as a result of technical problems with the HMD. The weight of the HMD itself can lead to user discomfort and simulator sickness symptoms. Furthermore, due to the required head tracking process, transmission delays occur which influence the update rate of the visual scene [13], [45]. These latencies can encourage discrepancies between visual and vestibular stimuli and therefore contribute to simulator sickness. In order to avoid simulator sickness, latencies should not exceed a value of 50 ms [27].

Additionally, the duration and exposure to the driving simulator also influence the occurrence of simulator sickness. Kennedy, Stanney, and Dunlap [32] could show that longer, coherent time periods (e.g., 60 minutes vs. 20 minutes) spent in the simulator led to more simulator sickness symptoms. However, repeated exposures (e.g., 2 times a month compared to 6 times a month) promoted adaptation and therefore reduced sickness outcomes [32]. Kennedy, et al. [30] suggest that a time interval of two to five days is optimal in order to adapt to simulator sickness. Moreover, characteristics of the driving task influence the occurrence of simulator sickness. As Mourant, et al. demonstrated, driving on curves and in city environments led to more simulator sickness than driving on straight roads [34].

Further, user characteristics, like gender or age, can also contribute to simulator sickness. For example, Brooks et al. [10] pointed out that older participants suffer from more simulator sickness symptoms.

In most studies, simulator sickness is measured with the Simulator Sickness Questionnaire (SSQ) from Kennedy et al. [30]. This questionnaire consists of 16 different symptoms, which contribute to three different subscales: nausea (N), which includes gastrointestinal symptomatology, oculomotor disturbances (O), which contain visual problems, and disorientation (D), which involves vestibular symptoms like dizziness or vertigo. Subjects have to rate the severity of each symptom on a 4-point scale from 0 (none) to 3 (severe). In addition to the three symptom clusters, a total sickness score is computed, which can be considered as the most reliable indicator of simulator sickness severity [30]. The highest possible total score is 235.

Simulator sickness is likely to occur in every kind of driving simulator. As simulator sickness can alter driving behavior, it is essential to account for simulator sickness in driving simulation and especially validation research.

III. Research Questions

For the safe, reliable, realistic and reproducible testing of ADAS (especially for active safety systems), new test methods are needed which overcome the disadvantages of driving simulators and real test drives. As explained above, the VIL offers many advantages compared to real test drives and driving simulators. However, results obtained in VIL studies can only be transferred to reality, if behavioral validity of the VIL test setup is ascertained. Thus, four main research questions have been addressed. First, it was one aim of this research study to investigate if the VIL displays absolute behavioral validity for longitudinal driving maneuvers. The second research question focused on the systematic examination of relative validity. Thirdly, face validity should be explored. In this first validation study lateral driving behavior was deliberately omitted because of existing problems with the simulation of turning maneuvers.

As the VIL might, just like driving simulators, have an issue with simulator sickness, the fourth aim of the current research study was to investigate the occurrence of simulator sickness while driving the VIL.

IV. Method

A. Driving Maneuvers

In order to investigate behavioral validity of the VIL setup, five different driving maneuvers where chosen. Each of these frequently occurs in every day driving and can be implemented both in the VIL and in reality on a closed-off test track.

In the first maneuver participants were instructed to follow a leading vehicle with a constant distance of 15 meters while the leading vehicle drove steadily 50 km/h.

In the second maneuver, the participants also had to follow the leading vehicle with a constant distance of 15 meters. First, the car in front drove 40 km/h. Then, accelerated with 1 m/s^2 to 60 km/h, continued with constant speed, and finally decelerated with 1 m/s^2 to 40 km/h again at the end. Participants were told to adapt their driving behavior to that of the leading vehicle in order to keep a constant distance of 15 meters.

The task in the third driving behavior constituted overtaking of a leading vehicle that drove at a constant speed of 50 km/h. Again, participants were instructed to follow the leading vehicle with a constant distance of 15 meters. At a certain point at the test track, the examiner asked the subjects to overtake the leading vehicle. No speed limits were given.

The fourth task was to stop behind a vehicle that was already waiting at a junction. The instruction in the fourth maneuver was to drive with a constant speed of 50 km/h. No instructions were given concerning the vehicle already waiting at a junction. Subjects were only told to adapt their driving behavior to the traffic situation.

The last task entailed an unanticipated braking maneuver. The subjects had to follow a leading vehicle with a constant distance of 15 meters. The speed of the leading vehicle was 55 km/h. Unexpected for the subjects, the leading vehicle decelerated with 3.50 m/s^2 to 30 km/h. The instruction in the fifth driving maneuver was the same as in tasks one and two.

The corresponding time headways for tasks 1, 2, 3, and 5 are summarized in Table 1. For all driving maneuvers

Table 1. Driving maneuvers and time headways.

Driving Maneuver	Time Headway
(1) Following a leading vehicle with constant speed of 50 km/h	1.1 s
(2) Following a leading vehicle with varying speed (40–60–40 km/h)	1.35 s–0.90 s–1.35 s
(3) Overtaking a leading vehicle (50 km/h)	1.1 s
(4) Anticipated braking maneuver	No leading car
(5) Unanticipated braking maneuver (55 km/h)	0.98 s

the section of measurement was 500 meters. In order to control the speed profile of the leading vehicle in reality, the driver of this car received an intensive training before the study. As additional help, optical hints at the roadside marked the starting and endpoint of the acceleration and deceleration phases.

B. Experimental Design and Measurement Variables

As mentioned in section II (C), high rates of optic flow can intensify simulator sickness symptoms. In order to investigate whether a visual scene with lots of elements that encourage optic flow has an influence on simulator sickness as well as on driving behavior, two different visual environments were designed. The first environment was more complex and consisted of a number of static objects, like trees or buildings, in order to provoke a high degree of optic flow (in the following stated as "complex virtual environment"). The second virtual world, however ("simple virtual environment"), matched the real environment of the test track and contained only few objects that could encourage optic flow and was therefore very simple. Both virtual worlds as well as the real test track are depicted in Fig. 4.

A 5 (driving maneuvers) × 3 (driving environment) within-subjects experimental design was chosen in order to examine the research questions previously outlined. The effects of three different driving environments (reality/VIL complex virtual environment/VIL simple virtual environment) in five different driving situations (following a leading vehicle with constant speed/following a leading vehicle with varying speed/overtaking a leading vehicle/anticipated braking maneuver/unanticipated braking maneuver) on driving behavior, perceived realism and simulator sickness were investigated. Latin square method was used in order to control for serial effects.

Various driving parameters, which were most suited to assess the participants' braking, steering and distance behavior, constituted the objective measures of driving behavior. Absolute validity was determined by the comparison of the averaged values of these parameters between driving in a real car and driving the VIL. Relative validity was assessed by the comparison of subjects' braking behavior in the two experimental conditions "anticipated braking maneuver" and "unanticipated braking maneuver" as well as by the comparison of their reaction times in the two driving maneuvers "following a leading vehicle with varying speed" and "unanticipated driving maneuver" between reality and the VIL conditions. Table 2 summarizes the objective driving parameters for each of the five driving maneuvers.

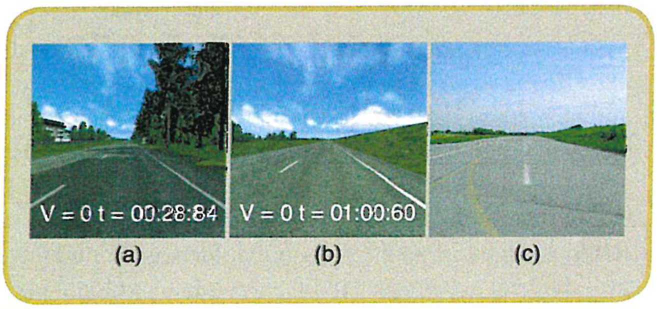


FIG 4 (a) Complex and (b) simple virtual world, and (c) reality.

Simulator sickness was measured by the simulator sickness questionnaire from Kennedy et al. [30]. Scores for the three subscales of nausea, oculomotor disturbances, and disorientation were computed as well as the total sickness score.

Additionally, perceived realism was assessed with a single-item 5-point Likert scale.

C. Test Setup and Data Recording

For the validation study, an off-the-shelf Audi A6 Avant was used. Only a CAN interface to read controller messages and record objective vehicle data was additionally implemented. The VIL hardware and a second PC, which was used to record the reality scenarios, were installed inside the trunk.

During the simulator scenarios, data were recorded with the Virtual Test Drive software, at a frequency of 60 Hz. In addition to information pertaining to the vehicle's state, (e.g., steering angle, velocity, etc.) the positions of all vehicles in the simulation were stored simultaneously. The distance from the VIL to the leading vehicle was calculated offline by use of the Euclidean distance measure.

Table 2. Driving maneuvers and objective measures.

Driving Maneuver	Objective Measures
(1) Following a leading vehicle with constant speed	• Standard deviation of the average distance to the leading vehicle
(2) Following a leading vehicle with varying speed	• Maximum position of accelerator pedal • Maximum longitudinal acceleration • Time between acceleration of leading vehicle and increase of accelerator pedal of at least 10% • Maximum brake pressure • Maximum longitudinal deceleration • Time between deceleration of leading vehicle and zero-position accelerator pedal
(3) Overtaking a leading vehicle	• Maximum steering angle • Maximum accelerator position • Maximum longitudinal acceleration
(4) Anticipated braking maneuver	• Maximum brake pressure • Maximum longitudinal deceleration • Distance to leading vehicle at point of start braking • Distance to leading vehicle at point of standstill
(5) Unanticipated braking maneuver	• Maximum brake pressure • Maximum longitudinal deceleration • Time between deceleration of leading vehicle and zero-position accelerator pedal

During the real-life scenarios, all data were recorded with the Vector CANape software. Vector CANape is a software tool for the development, calibration and diagnostics of electronic control units. The distance to the leading vehicle was measured with the built-in radar of the adaptive cruise control.

D. Procedure

Prior to the experimental trials, a short video of the VIL was shown, the main operational principles of the VIL were explained and signed consent was obtained. Next, participants had time to familiarize themselves with the handling of the experimental vehicle. This training lasted for approximately 10 minutes and included full braking applications and driving with different speeds.

Subsequently, participants had the opportunity to become familiar with driving the VIL in another test phase, which also took around 10 minutes to complete. Before each of the three experimental trials, participants had to follow a leading vehicle and practiced to keep a constant distance of 15 meters behind that leading car. The examiner monitored the actual distance of the participants and helped them to hold an exact distance of 15 meters by giving them feedback about their actual headway. Following the training phases, the experimental trials were conducted, which entailed three phases: driving in reality, driving VIL with a complex environment, and driving VIL with a simple environment. Each of these three experimental phases lasted for about 30 minutes. The entire procedure took on average 2.5 hours per participant, over a period of two days.

As mentioned above, the order of experimental conditions systematically varied in order to control for serial effects. Consequently, the sequence of conditions changed for each participant. An exemplary procedure is detailed in the following. Upon completion of the training phase, a participant would e.g., first complete the five investigated driving maneuvers in reality. After a short break, the participant would execute all five driving maneuvers in the VIL with a complex virtual reality and afterwards complete the realism questionnaire as well as the SSQ. In order to avoid overstrain, the second VIL phase never followed the first one directly, but took place on another day. Prior to this second VIL phase, the participant would be given another opportunity to familiarize themselves with the vehicle. S/he would then execute the same driving maneuvers (in a different order) in the simple virtual world and fill out another SSQ and realism questionnaire. Finally, the participant was debriefed and compensated for his/her efforts with 30 EUR.

E. Subjects

Because of the HMD, people with glasses could not take part in this study. Another exclusion criterion was driving experience of less than 10,000 km per year. A total of 44 subjects (15 female and 29 male) participated in this validation study. However, two people could not finish the study because of severe simulator sickness symptoms and were excluded from further analysis. Subjects were students or staff members of the Universitaet der Bundeswehr Muenchen and were recruited around campus. The mean age of the sample was 29 years (SD = 10) with a range between 20 and 56 years. The participants had an average driving experience of 23,797 km per year (SD = 14,033). None of the subjects has taken part in a driving simulator study before.

V. Results

A. Simulator Sickness

None of the participants reported any sickness or illness symptoms prior to the test phase. However, two participants had to break off during their first VIL test drive due to severe simulator sickness symptoms and their data were excluded from further analysis. 2% of the subjects reported no symptoms at all in the SSQ when driving VIL in the simple environment, whereas 12% complained about severe symptoms.

Table 3. Means and standard deviations for simulator sickness symptoms.

	VIL Simple Virtual World		VIL Complex Virtual World	
Symptoms SSQ	Mean	SD	Mean	SD
General discomfort	0.60	0.63	0.79	0.75
Fatigue	0.24	0.48	0.24	0.48
Headache	0.33	0.57	0.26	0.50
Eyestrain	0.98	0.90	0.81	0.63
Difficulty focusing	1.02	0.84	0.98	0.81
Increased salivation	0.05	0.22	0.05	0.22
Sweating	0.40	0.63	0.40	0.63
Nausea	0.48	0.71	0.45	0.71
Difficulty concentrating	0.55	0.74	0.48	0.63
Fullness of head	0.48	0.74	0.33	0.61
Blurred vision	0.48	0.67	0.57	0.67
Dizzy (eyes open)	0.29	0.55	0.26	0.63
Dizzy (eyes closed)	0.05	0.22	0.12	0.40
Vertigo	0.26	0.50	0.29	0.60
Stomach awareness	0.31	0.68	0.36	0.66
Burping	0.02	0.15	0.07	0.34

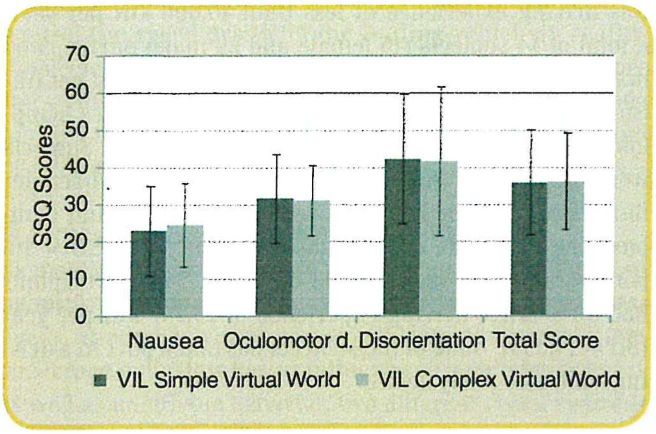


FIG 5 Simulator sickness scores (bars = SD).

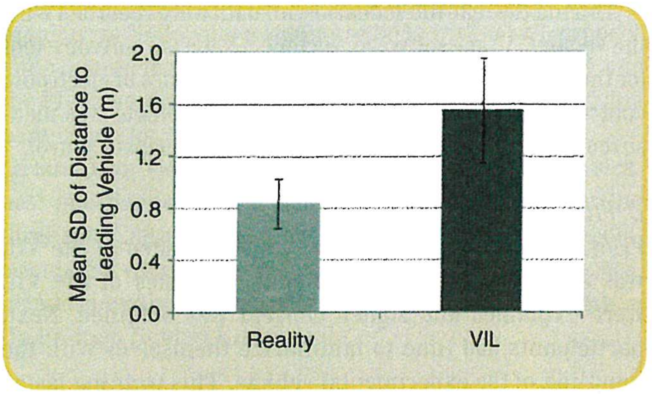


FIG 6 Following a leading vehicle with constant speed—comparison between reality and VIL (bars = SD).

5% showed no simulator sickness symptoms after driving VIL in the complex environment, whilst 9% experienced intense sickness symptoms. Table 3 provides an overview of the average scores for each individual symptom. In order to assess the extent of simulator sickness in more detail, the nausea, oculomotor disturbances, and disorientation scores as well as the total sickness scores were computed according to Kennedy et al. [30]. Wilcoxon signed-rank tests were conducted to compare the respective scale and total sickness scores between driving VIL in a complex and simple virtual environment. As depicted in Fig. 5, disorientation symptoms were more severe than nausea ones or oculumotor disturbances in both experimental conditions. At an α-level of 5%, no statistically significant differences were found between the two VIL conditions with regard to simulator sickness.

B. Absolute Validity

In order to compare the influence of the two VIL conditions on driving behavior, dependent t-tests were conducted. In none of the five driving maneuvers, the design of the virtual world (complex vs. simple) had a significant effect on the measured driving parameters at an α-level of 5%. Therefore, for further statistical analysis the data of the two VIL conditions could be subsumed and compared with those of the reality condition.

Since absolute validity can be assumed when the numerical values are equal in both experimental conditions, it was tested whether there are differences between driving in reality and driving in the VIL. For this purpose, dependent t-tests and Wilcoxon signed-rank tests were used depending on whether the assumptions for parametric tests had been met or violated. Since the aim of this study is to show that driving in the VIL does not differ from driving in reality, it is essential for the statistical analysis to keep the probability of a Type II error (i.e., the likelihood of failing to find a difference that is really there) at a minimum level of 5%. Due to the fact that there is a trade-off between Type I and Type II errors, an α-level of 20% was chosen, see e.g., [15]. An a priori statistical power analysis revealed that, given a sample size of at least 40 subjects and an assumed medium effect size, an α-level of 20% would be sufficient to reach a statistical power of 95%. In order to characterize the magnitude of observed effects, Pearson's correlation coefficient r was calculated. In the following, the five driving maneuvers are analyzed in detail.

Following a Leading Vehicle with Constant Speed

In this scenario, participants were asked to keep a constant distance of 15 meters to the car driving in front of them. In reality the mean following distance was 18 meters (SD = 1.72) compared to 17 meters (SD = 2.57) in the VIL condition. As depicted in Fig. 6, the distance to the leading vehicle varied significantly more when driving in the VIL compared to driving in reality ($z = -4.55$, $p < 0.001$, $r = -0.50$).

Following a Leading Vehicle with Varying Speed

Here, the participants were also instructed to follow a leading vehicle with a constant distance of 15 meters. However, this time, the leading car varied its driving speed consequently; subjects had to continuously adapt their own driving speed in order to keep the distance at the constant level of 15 meters. The driving maneuver was divided into an acceleration and a deceleration phase.

As shown in Fig. 7, participants (N = 41) pressed the accelerator pedal more forcefully in reality than they did during driving the VIL ($z = -3.16$, $p < 0.001$, $r = -0.35$). Consequently, the mean maximum longitudinal acceleration was also higher in reality ($z = -4.10$, $p < 0.001$, $r = -0.45$). Further, with an average reaction time of 1.09 s (SD = 0.52 s) participants (N = 39) reacted more quickly to the changing speed of the leading vehicle while driving in reality compared to driving the VIL (M = 2.38 s, SD = 0.85 s; $z = -5.43$, $p < 0.001$, $r = -0.61$).

In the deceleration phase, however, subjects (N = 41) braked more forcefully while driving the VIL compared to driving in reality ($z = -2.31$, $p = 0.02$, $r = -0.26$). Therefore, the mean maximum deceleration was also higher ($z = -3.56$, $p < 0.001$, $r = -0.45$). On the other hand, participants (N = 36) reacted more quickly to the braking leading vehicle in

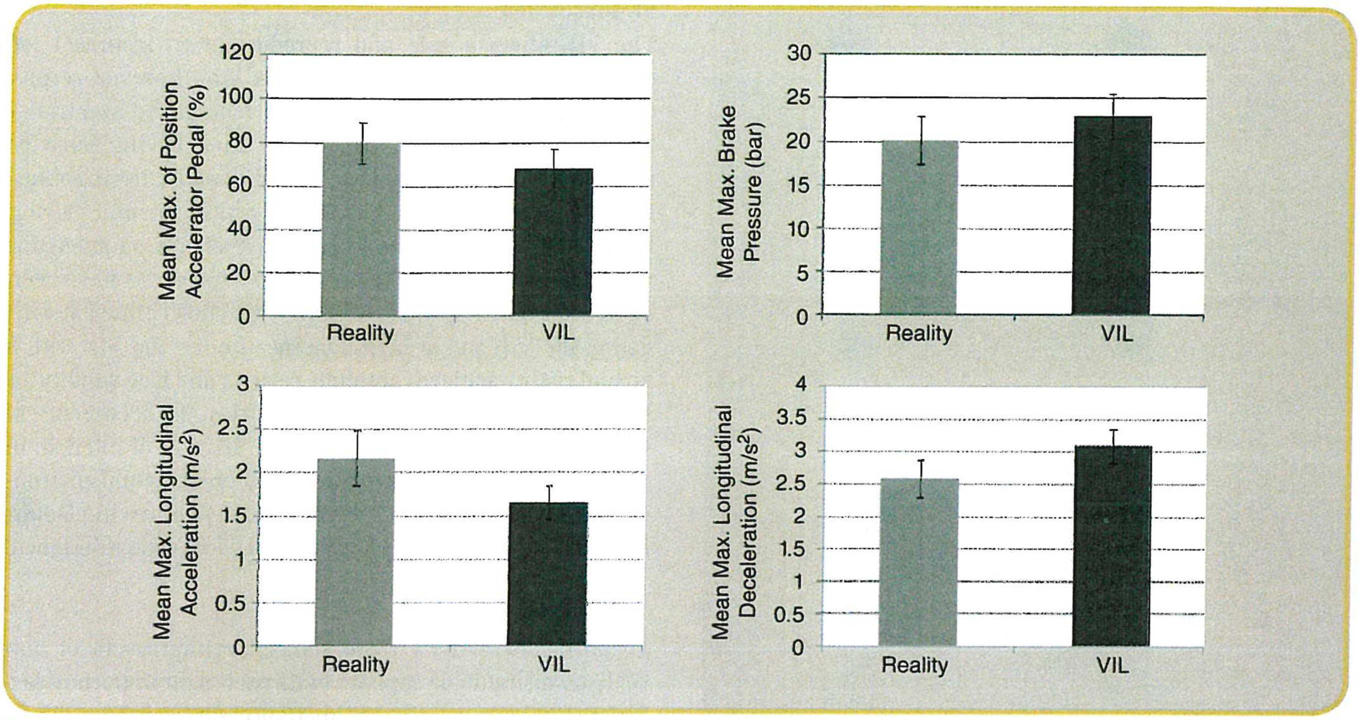


FIG 7 Following a leading vehicle with varying speed—comparison between reality and VIL (bars = SD).

reality (M = 0.93 s, SD = 0.76 s), than in the simulation condition (M = 1.79 s, SD = 0.68 s; z = −3.81, $p < 0.001$, r = −0.45).

Overtaking a Leading Vehicle

During this maneuver, participants were asked to overtake a leading vehicle. As displayed in Fig. 8, subjects pressed the accelerator pedal more forcefully in reality compared to driving the VIL (z = −4.37, $p < 0.001$, r = −0.48). The mean maximum acceleration was also higher in reality (z = −5.61, $p < 0.001$, r = −0.61). No significant differences, however, were found between the two driving modes regarding the mean maximum steering angle (z = −0.27, p = 0.79).

Anticipated Braking Maneuver

In this maneuver, subjects had to stop behind a vehicle that was already standing at a junction. On average, participants started braking 79.68 meters in front of the standing vehicle while driving the VIL compared to 49.90 meters in reality (t(41) = −14.41, $p < 0.001$, r = 0.91). Additionally, subjects stopped more closely to the standing vehicle in reality than in the VIL conditions (z = −1.87, p = 0.06, r = −0.20). No significant differences between the two driving modes were found concerning the average brake pressure (z = −0.38, p = 0.70) as well as the mean maximum longitudinal deceleration (z = −0.18, p = 0.86). The individual values are summarized in Fig. 9.

Unanticipated Braking Maneuver

As depicted in Fig. 10, subjects took more time to react to the unanticipated braking of the leading vehicle while driving the VIL compared to driving in reality (t(37) = −12.35, $p < 0.001$, r = 0.90). In addition, participants braked with higher force in reality than in the simulation condition (z = −2.56, p = 0.01, r = −0.28). Therefore, the mean maximum longitudinal deceleration was also significantly higher in reality (z = −2.88, $p < 0.001$, r = −0.32).

C. Relative Validity

In order to assess relative validity, the mean maximum brake pressure in the anticipated braking maneuver was compared with that in the unanticipated braking maneuver, both in reality and in simulation. For a better overview, the mean maximum brake pressure for the two experimental conditions is depicted in Fig. 11. Relative validity can be assumed when brake pressure differences are of the same direction and order in both experimental conditions.

As can be seen in Fig. 11, in both experimental conditions participants braked more forcefully in the unanticipated braking maneuver. In reality, the mean maximum brake pressure was 3.5 times higher in the unanticipated braking maneuver than in the anticipated maneuver, whereas in the VIL condition the mean maximum brake pressure tripled.

Additionally, relative validity was assessed by the comparison of the average time that participants took to react to the braking leading car with the release of the accelerator pedal in the unanticipated braking maneuver and the maneuver in which participants followed the vehicle with varying speed. Again, relative validity can be assumed when reaction time differences are of the same direction and approximate magnitude in both experimental

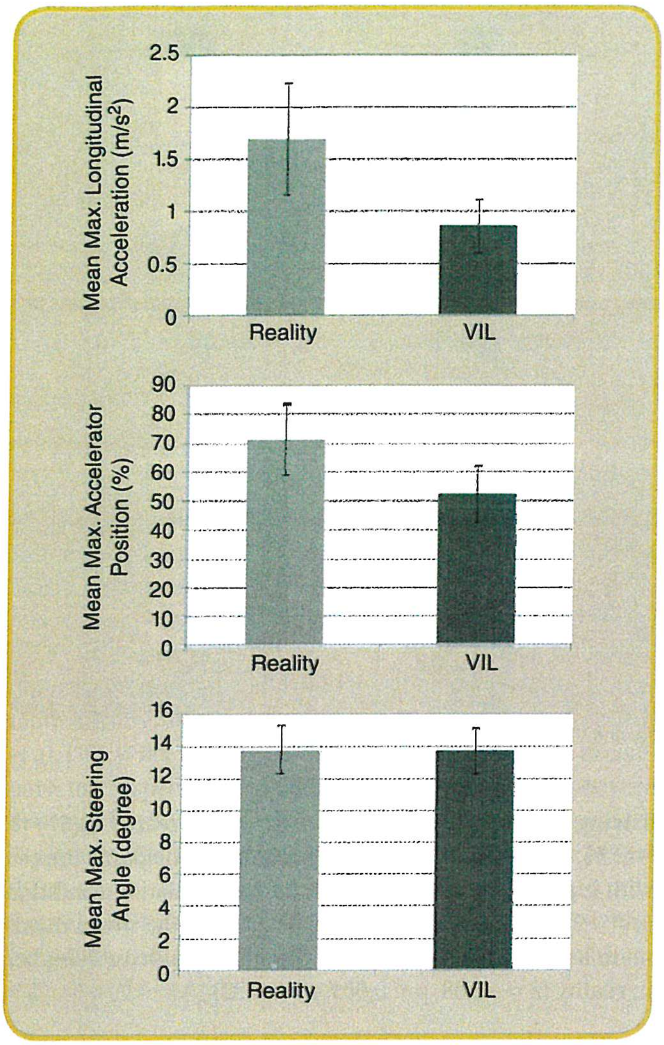


FIG 8 Overtaking a leading vehicle—comparison between reality and VIL (bars = SD).

conditions (reality and VIL). Mean reaction times are depicted in Fig. 12. It can be seen that in both maneuvers subjects reacted more quickly in reality compared to the simulation. However, in both, the VIL and the reality conditions, drivers released the accelerator pedal more quickly in the unanticipated braking maneuver than in the maneuver in which participants followed the vehicle with varying speed. In reality, subjects reacted 2.5 times faster in the unanticipated braking maneuver than in the maneuver with varying speed, compared to 2.2 times in the VIL condition.

D. Perceived Realism

After each VIL condition, participants were asked to indicate how realistic the VIL environment looks to them on a single-item Likert scale ranging from 1 (very unrealistic) to 5 (very realistic). Perceived realism ratings for both VIL conditions are displayed in Fig. 13. Nearly 80% of the participants conceived the VIL as realistic in both conditions. Only 10% of the subjects stated that the VIL appears unrealistic to them.

VI. Discussion

The VIL offers a safe and reproducible environment for the development and evaluation of ADAS. However, results obtained in VIL studies can only be generalized and transferred to reality, if the VIL elicits the same driving behavior as it is observed in the real world. Additionally, driving simulators often elicit simulator sickness, which can alter driving behavior and as a result have negative effects on simulator validity. Therefore, four main research questions have been addressed. The first three research questions aimed at validating the VIL and at investigating whether the VIL with a virtual reality achieves absolute, relative and face validity for longitudinal driving maneuvers. Fourthly, the occurrence of simulator sickness while driving the VIL was of interest. In order to investigate the influence of different virtual environment configurations (complex vs. simple) on driving behavior and simulator sickness two virtual worlds were designed.

A. VIL Validity

According to Blana [7] the subjective impression of how realistic a simulator appears to its user is an important factor of simulator validity as this impression can have effects on driving motivation and driving behavior. For example, an unrealistic driving simulator may motivate the subjects to behave differently as they would in reality. Based on the results presented in section V (D) the VIL achieves good face validity. 80% of the subjects perceived the VIL as (very) realistic. Especially the realistic representation of driving dynamics as well as vehicle noise were positively noted. No significant differences were found between the complex and simple virtual world concerning subjective appraisals. However, as reported by Berg, Karl, and Färber [4], subjective immersion is higher for the complex world compared to the simple one. In order to generate a realistic driving environment, the complex virtual world may therefore be preferred. On the other hand, the operation of the complex virtual world requires more computational power, which may have negative consequences concerning other system factors like time lags. This dilemma cannot be resolved in general, but has to be considered with respect to the study aims.

Besides face validity, behavioral validity plays a significant role when simulator results are applied to reality. In order to assess absolute behavioral validity, five different driving maneuvers were compared between two VIL and one real test drive. As no statistical significant differences were found between driving the VIL with a complex and a simple virtual world, no further distinctions are made between the two VIL test drives. Absolute validity was only achieved for three individual driving parameters. Specifically, no differences were found concerning the maximum steering angle when overtaking a leading vehicle. Additionally, the maximum brake pressure as well as the maximum longitudinal deceleration did not differ significantly in the anticipated braking maneuver where subjects were

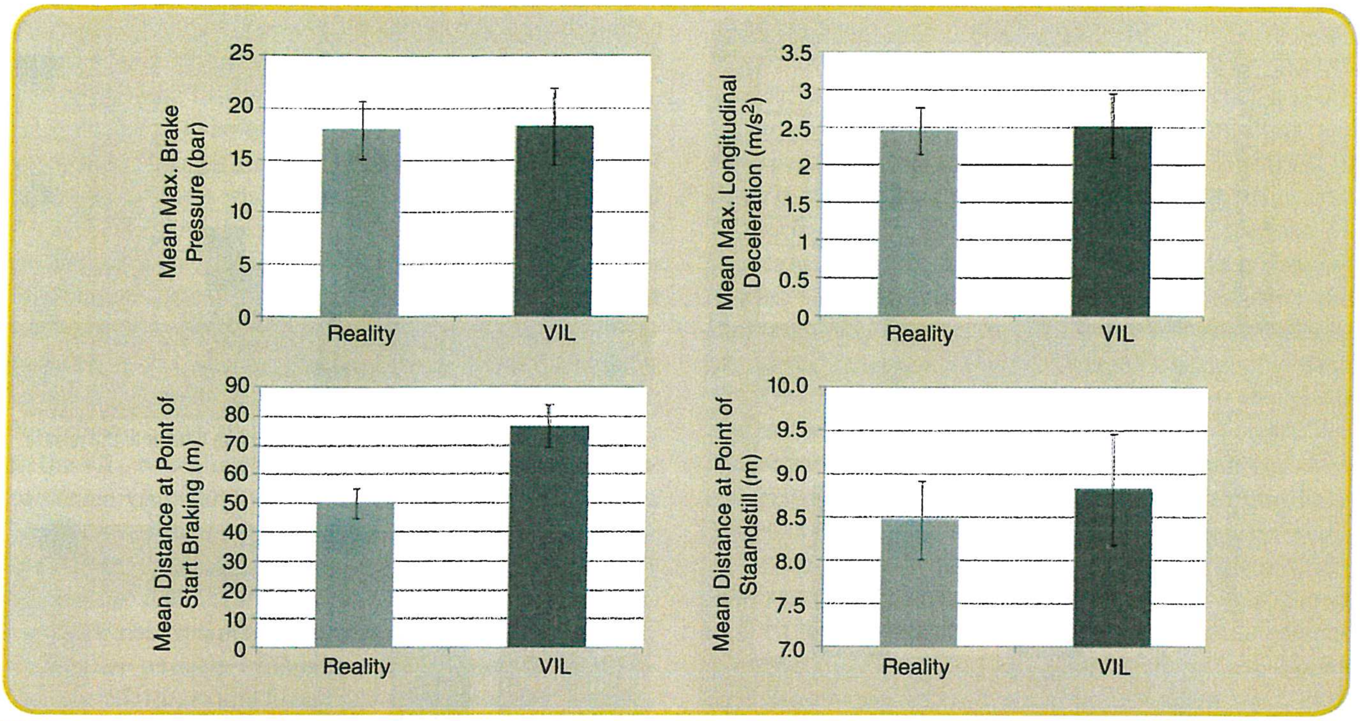


FIG 9 Anticipated braking maneuver—comparison between reality and VIL (bars = SD).

required to stop behind a standing vehicle. However, for all other driving parameters significant differences emerged. For example, it was more difficult for the subjects to keep a constant distance of 15 meters to the leading vehicle in the VIL condition. In addition, participants took less time to react to speed changes of the leading vehicle in reality than while driving the VIL. As many subjects mentioned, the perception of distance as well as the perception of distant objects was more difficult in the simulation than in reality, which might have contributed to the observed higher reaction times. As a consequence, participants had to brake with higher forces during the maneuver "following a leading vehicle with varying speed" in order to keep a constant distance of 15 meters. Moreover, subjects initiated the deceleration process earlier, i.e., at greater distance behind the leading vehicle in the anticipated driving maneuver.

However, in general the tendency to react more slowly and to accelerate more smoothly was found while driving the VIL. Various causes are conceivable for this trend. On the one hand, subjects might have perceived higher acceleration and deceleration forces as unpleasant because of the inertia of the HMD. On the other hand, the criticality of the driving situation could have been experienced as less severe in the VIL as in reality and therefore led to different braking patterns. Additionally, subjects mentioned that the perception of distances was more difficult in the virtual world than in reality. This could also be responsible for differences concerning the longitudinal driving behavior. The absence of the cockpit and the instrument panel in the virtual world could have exacerbated the perception of distances in the VIL. In the VIL, the driver's perspective of the road ahead is quite different as in a real car. This may have led to a slightly different driving behavior. For future VIL studies, it is required to integrate the cockpit in the simulation. This is also necessary if in-vehicle displays should be tested with the VIL test setup.

Another possible explanation for the lack of absolute validity could be the occurrence of simulator sickness, which is discussed in detail in the following section.

Although absolute validity was only found for three driving parameters, these results are not unexpected. Blaauw [6] for example could achieve absolute validity for his driving simulator only for longitudinal, but not for lateral vehicle control. However, Törnros [46] points out that relative validity is sufficient for the application of driving simulators as research tools (p. 497).

In order to demonstrate relative validity the mean maximum brake pressure as well as the mean reaction times of two experimental conditions were compared, both in reality and in the simulation. In both cases relative validity could be well established. Although there were significant differences concerning the absolute values, the found effects were in the same direction and of similar magnitude. Thus, it could be shown that subjects were able to realize the various situation demands and to react situation-specifically. For example, both in reality and in the simulation, subjects reacted more quickly and braked more forcefully in the unanticipated braking maneuver.

In conclusion, the VIL elicits situation-specific driving behavior that is comparable to driving in the real world. Hence, it seems that the VIL offers high potential for the

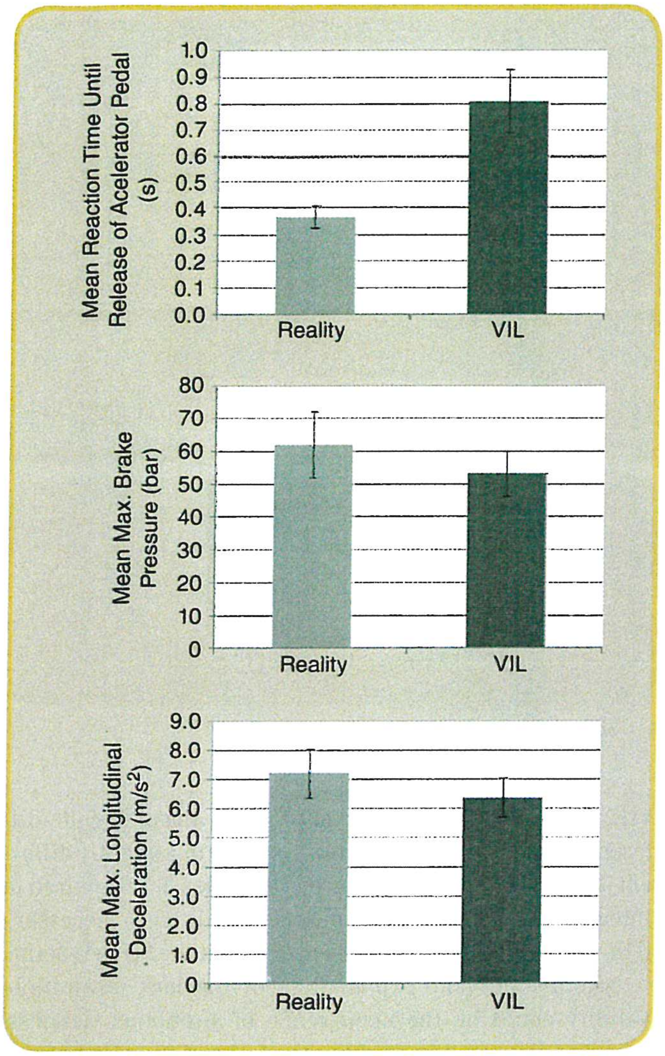


FIG 10 Unanticipated braking maneuver—comparison between reality and VIL (bars = SD).

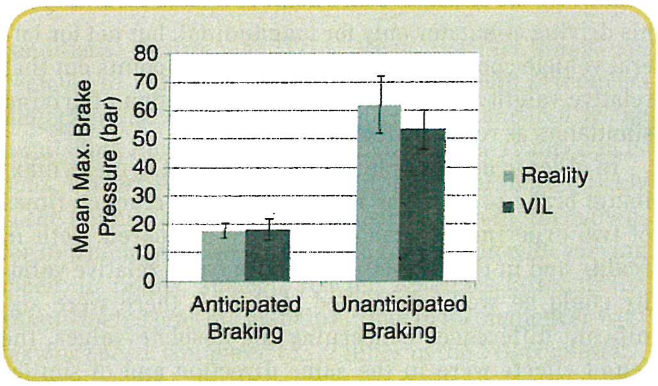


FIG 11 Effects of maneuver type on mean max. brake pressure (bars = SD).

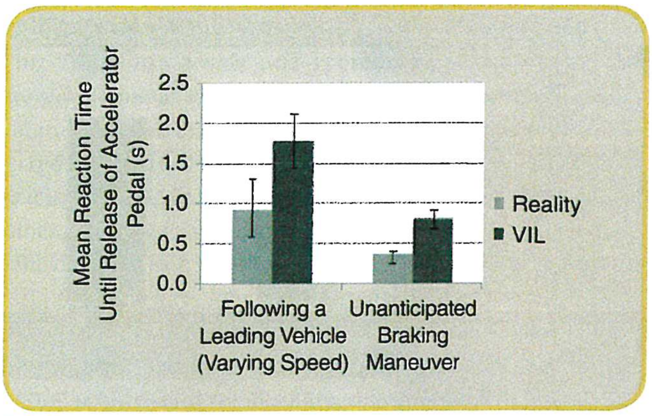


FIG 12 Effects of maneuver type on mean reaction time until the release of the accelerator pedal (bars = SD).

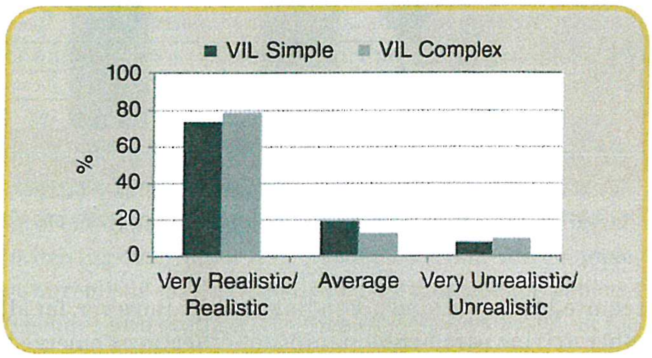


FIG 13 Perceived realism for each VIL condition.

development and evaluation of ADAS. In particular driving behavior during system failures, e.g., the malfunction of the automatic braking system or the adaptive cruise control can thus be investigated in a safe environment without endangering other traffic participants. However, due to the lack of absolute validity, the VIL is not yet suited as a research tool for the fine tuning of comfort systems. Conceivably, the introduction of a correction factor would mitigate the differences in the absolute numerical values between reality and simulation, which warrants further research. Additionally, the present study focused on longitudinal driving maneuvers. Thus, before investigating ADAS with the VIL test setup that supports the driver in his lateral driving behavior, an additional validation study is vital. It is assumed that simulator sickness is even more problematic when turning maneuvers are driven in simulation. That is one reason why the obtained results cannot simply be transferred to lateral driving maneuvers.

B. Simulator Sickness

As mentioned above, validation results have to be treated with caution if simulator sickness occurs. Hence, simulator sickness was measured by the Simulator Sickness Questionnaire [30]. Although the literature suggests that simulation sickness might be exacerbated with increasingly detailed virtual landscapes, severity as well as symptom constellation did not differ between the complex and the simple virtual worlds. Consequently, no further distinction between the two virtual environments was made. Overall, 4% of the subjects could not finish the experiment because of severe simulator sickness symptoms. Only 2% experienced no sickness symptoms at all, whereas the remaining 94% reported at least a minor increase in simulator sickness symptoms measured with the SSQ. According to Kennedy et al. [31] total sickness scores higher than 20 indicate perceivable discomfort. Sickness scores higher than 100, however, point to severe illness. In both VIL conditions the mean total sickness score is 36, wherefore the magnitude of simulator sickness while driving the VIL can be considered unpleasant. These results are comparable with other studies assessing virtual reality-induced simulator sickness (see for example [13], [21]). However, based on the assumption that vection is a major contributing factor of simulator sickness [22], it was assumed that the VIL induces less simulator sickness symptoms compared to other driving simulators. Nevertheless, the accurate representation of motion cues in the VIL, which ought to reduce the conflict between the presented visual and vestibular cues, could not reduce the occurrence of simulator sickness. One possible explanation could be that there are still residual vection conflicts induced by measurement errors. However, these small inconsistencies could not explain the high amount of simulator sickness.

Another possible explanation for the study's results is given by Hettinger and Riccio [23]. They point out that vection is only one of two major causes of simulator sickness. Simulator sickness also often occurs when there are large time lags between head movements and the update of the visual scene. As mentioned above, these latencies should not exceed 50ms [27]. With the soft- and hardware used in the presented VIL setup, a time lag of nearly 150ms occurred. These latencies might have led to slight discrepancies between the visual and vestibular inputs and thereby encouraged simulator sickness symptoms. This is in line with the finding that the amount of simulator sickness in Navy simulators was exacerbated by increasing transmission delays [43]. Additionally, Cobb et al. [13] showed that greater time lags resulted in higher disorientation scores at the SSQ. As pointed out in section V, disorientation symptoms were most severe in the present study, followed by oculomotor and nausea symptoms. In contrast, Moss et al. [33] found that simulator sickness during HMD exposure did not increase with longer time delays. However, they could show that an increase of exposure duration encouraged simulator sickness. Therefore, another possible explanation for the high amount of simulator sickness in the present study could be that the exposure times of about 30 minutes have been too long.

Kennedy et al. [31] showed that the severity and the symptom constellation of simulator sickness can vary between different simulator types and thereby provide valuable cues to identify the causes of simulator sickness. Stanney and Kennedy [37], among others, pointed out that simulator induced sickness usually has an O (oculomotor disturbances) > N (nausea) > D (disorientation) profile whereas exposure to virtual environments led to a D > N > O profile. The VIL, however, elicited a slightly different, D > O > N, pattern. That is, consistent with other VR exposures, disorientation symptoms, such as dizziness or vertigo, were the dominating ones, whereas symptoms of nausea were least frequently reported. According to the results of Cobb et al. [13], one major cause for the onset of simulator sickness in the present study may have been the large transmission time lags. In contrast, in this VIL study nausea symptoms were less severe compared to disorientation and oculomotor symptoms. Yet, the absolute nausea scores are comparable with other simulator sickness studies, see for example [13] or [21]. On the other hand, the disorientation as well as the oculomotor disturbance scores are much higher compared to the results obtained by Cobb et al. [13] or Häkkinen et al. [21]. This indicates that HMD design and technical factors are a major cause of simulator sickness while driving the VIL with a virtual reality.

Indeed, besides the latencies further factors may have contributed to the observed simulator sickness. For example, weight and inertia of HMDs have been linked to simulator sickness. In fact, HMDs as light as 600g were found to elicit simulator sickness [45]. The HMD of the present study weighs 1400g. This led, to pronounced, involuntary head movements when driving on an uneven road. Furthermore, especially during hard braking maneuvers the HMD swung back and forth, which the subjects tended to experience as unpleasant and might have encouraged simulator sickness. In addition, Häkkinen et al. [21] showed that disorientation as well as total sickness scores were significantly higher when using stereoscopic compared to non-stereoscopic displays. Both scores are comparable with the disorientation and total score of the present study. Hence, it is conceivable that the use of an HMD with stereoscopic depth may also be a major contributing factor to simulator sickness.

In summary, driving the VIL led to at least minor simulator sickness symptoms. Several HMD software and hardware design factors may have contributed to the dominance of disorientation symptoms in the present study. Clearly, further research is necessary to investigate the influence of these factors in more detail.

As mentioned above, simulator sickness can have negative consequences for driving simulator validation. Subjects often modify their driving behavior in order to reduce sickness symptoms as demonstrated by Cobb et al. [13]. In the present study behavior adaptation in order to reduce simulator sickness symptoms could be a possible explanation for some effects found in the present study. For instance, subjects could have been braking more gently in the VIL conditions in order to avoid the HMD rocking back and forth which was experienced as unpleasant. Hence, it is possible that simulator sickness is at least in some cases responsible for the lack of absolute validity in the present study. Other consequences, such as a loss of motivation or the lack of concentration, are also possible [45]. Although other explanations are conceivable, validation results have to be treated carefully as long as simulator sickness emerges. Therefore, further development of the VIL is required in order to reduce simulator sickness and thus facilitate the interpretation of validation research results.

VII. Outlook

The present study provided clear indications for future improvements of the VIL, so that it may be utilized as a safe and reliable simulation tool for the evaluation and development of ADAS. As a measure of improvement, simulation sickness has to be decreased. To counter simulation sickness, latency in head tracking needs to be minimized. Furthermore, the majority of the subjects complained about the slow image update during head movements. With the current hardware setup a time lag of approximately 30 ms is induced. Therefore, one of the major goals for further VIL improvements is to reduce the time lag in head tracking. To minimize the adverse influence of the used HMD on the development of simulation sickness, a lighter device (with the disadvantage of a smaller field of view) and the possibility to run the HMD in mono is also recommended, with which, a direct comparison of the different HMD influences on simulation sickness can be investigated in future studies. Further measures concerning the study design proved to be useful in the mitigation of simulator sickness. Kennedy, Stanney, and Dunlap [32] as well as Hoffman and Buld [24] showed that repeated exposure to simulators reduces the severity of sickness symptoms. In accordance with this study, participants in the present study also reported to experience the second VIL test drive as less unpleasant although no significant differences in the SSQ were found between the first and the second VIL test drive. Therefore, we recommend to introduce at least two training sessions in order to get familiar with the VIL. Moreover, it would seem useful to develop a specialized VIL training program in order to accelerate subjects' adaptation to the VIL and to reduce simulator sickness. In addition, time spent in the VIL should be limited to a maximum of 20 minutes.

Perceived realism may be further improved by displaying the ego car to the driver. Presently, the driver does not receive any visual feedback of the vehicle's interior. In fact, due to the direct view to the environment, the driver has the illusion of flying over the street. Therefore, further development efforts should strive to superimpose a vehicle in order to enhance the perceived realism of the driving experience.

Additionally, the current validation study focused on longitudinal driving behavior. In order to use the VIL test setup for lateral driving behaviors, another validation study is needed.

In conclusion, the VIL is a valuable research tool for the development and evaluation of ADAS, especially of collision avoidance systems like the automatic braking guard or the avoidance assistant. The present study demonstrated that, due to its proven relative validity, the behavior of drivers at system boundaries can be examined reliably and safely with the VIL. For the application of the VIL to the testing of comfort systems, however, absolute validity is required. While the VIL, at its present state of development does not yet show this absolute validity, the present study provided valuable insight into further developments that are required to achieve this goal.

About the Authors

Ines Karl studied psychology at the University of Regensburg, where she graduated in 2008. Between 2008 and 2009 she was working for MTO Psychological Research and Consulting GmbH in Tuebingen as a research consultant in the HMI team. Since the beginning of 2010 she has been working as a research assistant at the Human Factors Institute, Faculty for Aerospace Technology at the Universitaet der Bundeswehr Muenchen. Her main research interests lie in the field of advanced driver assistance systems, driver intention recognition and driving simulator validation.

Guy Berg studied information technology at Technische Universitaet Muenchen, where he graduated in 2008. Since 2008 he has been working as a research assistant at the Human Factors Institute, Faculty for Aerospace Technology at the Universitaet der Bundeswehr Muenchen. His main research interests lie in the field of advanced driver assistance systems and development of driving simulators.

Fabian Rüger studied aerospace engineering at the Universitaet der Bundeswehr Muenchen, where he graduated in 2011. Since 2011 he has been working as a research assistant at the Human Factors Institute, Faculty for Aerospace Technology at the Universitaet der Bundeswehr Muenchen. His main research interests lie in the field of advanced driver assistance systems, eye tracking systems and driving simulators.

Berthold Färber studied psychology at the University of Regensburg, where he received his doctoral degree in 1980. Since 1989 he is professor of human factors studies at the Universitaet der Bundeswehr, Munich. His research topics are traffic safety, human factors for advanced driver assistance systems and human-robot interaction.

References

[1] C. Ahlstrom, T. Victor, C. Wege, and E. Steinmetz, "Processing of eye/head-tracking data in large-scale naturalistic driving data sets," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 553-564, 2012.

[2] R. W. Allen, T. J. Rosenthal, and M. L. Cook, "A short history of driving simulation," in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, and J. D.Lee, Eds. Boca Raton, FL: CRC Press, 2011, pp. 1-16.

[3] G. J. Andersen, "Sensory and perceptual factors in the design of driving simulation displays," in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, J. D. Lee, Eds. Boca Raton, FL: CRC Press, pp. 1–11, 2011.

[4] G. Berg, I. Karl, and B. Färber, "Vehicle in the loop validierung der virtuallen validität," in *Der Fahrer im 21. Jahrhundert. Fahrer, Fahrerunterstützung und Bedienbarkeit*, Düsseldorf, Germany: VDI-Verlag GmbH, 2011, pp. 143–154.

[5] A. C. J. Bittner, B. Gore, and B. Hooey, "Meaningful assessments of simulator performance and sickness: Can't have one without the other," in *Proc. 41st Human Factors Ergonomics Society Annu. Meeting*, 1997, pp. 1089–1093.

[6] G. J. Blaauw, "Driving experience and task demands in simulator and instrumented car: A validation study," *Human Factors*, vol. 24, no. 4, pp. 473–486, 1982.

[7] E. Blana, "Driving simulator validation studies: A literature review," *Institute of Transport Studies*, Univ. Leeds, Leeds, U.K. 1996, p. 6.

[8] T. Bock, "Vehicle in the loop–test-und simulationsumgebung für fahrerassistenzsysteme," Ph.D. dissertation, Germany, Munich: TU München, 2008.

[9] T. Bock, M. Maurer, and G. Färber, "Validation of the vehicle in the loop–a milestone for the simulation of driver assistance systems," in *Proc. IEEE Intelligent Vehicles Symp.*, 2007, pp. 612–617.

[10] J. O. Brooks, R. R. Goodenough, M. C. Crisler, N. D. Klein, R. L. Alley, B. L. Koon, W. C. Logan Jr., J. H. Ogle, R. A. Tyrrell, and R. F. Wills, "Simulator sickness during driving simulation studies," *Accident Anal. Prevent.*, vol. 42, no. 3, pp. 788–796, 2010.

[11] J. K. Caird and W. J. Horrey, "Twelve practical and useful questions about driving simulation," in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J.K. Caird, and J. D. Lee, Eds. Boca Raton, FL: CRC Press, 2011, pp. 1–18.

[12] C.-Y. Chang and Y.-R. Chou, "Development of fuzzy-based bus rear-end collision warning thresholds using a driving simulator," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 360–365, 2009.

[13] S. V. G. Cobb, S. Nichols, A. Ramsey, and J. R. Wilson, "Virtual reality-induced symptoms and effects," *Presence*, vol. 8, no. 2, pp. 169–186, 1999.

[14] S. Cotter, J. Hopkin, A. Stevens, A. Burrows, P. Kompfner, and M. Flament, "The institutional context for advanced driver assistance systems: A code of practice for development," in *Proc. 13th World Congr. Intelligent Transport Systems*, 2006, p. 8.

[15] A. Field, *Discovering Statistics Using SPSS*. London, U.K.: SAGE Publications Ltd, 2009, pp. 1–856.

[16] O. Gietelink, J. Ploeg, B. De Schutter, and M. Verhaegen, "Development of advanced driver assistance systems with vehicle hardware-in-the loop simulations," *Veh. Syst. Dynamics*, vol. 44, no. 7, pp. 569–590, 2006.

[17] S. T. Godley, T. J. Triggs, and B. N. Fildes, "Driving simulator validation for speed research," *Accident Anal. Prevent.*, vol. 34, no. 5, pp. 589–600, 2002.

[18] E. B. Goldstein, *Sensation and Perception*. Belmont, CA: Wadsworth Publishing, 1989, pp. 1–459.

[19] J. Greenberg and M. Blommer, "Physical fidelity of driving simulators," in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, J. D. Lee, Eds. Boca Raton, FL: CRC Press, 2011, pp. 1–24.

[20] L. Harms, "Driving performance on a real road and in a driving simulator: Results of a validation study," in *Vision in Vehicles–V*, A. G. Gale, I. D. Brown, C. M. Haslegrave, and S. P. Taylor, Eds. Amsterdam, New York: Elsevier, 1996, pp. 19–26.

[21] J. Häkkinen, M. Pölönen, J. Takatalo, and G. Nyman, "Simulator sickness in virtual display gaming: A comparison of stereoscopic and non-stereoscopic situations," in *Proc. Mobile HCI*, 2006, pp. 227–229.

[22] L. J. Hettinger, K. S. Berbaum, R. S. Kennedy, W. P. Dunlap, and M. D. Nolan, "Vection and simulator sickness," *Military Psychol.*, vol. 2, no. 3, pp. 171–181, 1990.

[23] L. J. Hettinger and G. E. Riccio, "Visually-induced motion sickness in virtual environments," *Presence*, vol. 1, no. 3, pp. 306–310, 1992.

[24] S. Hoffmann and S. Buld, "Driving in a simulator design and evaluation of a training programme," in *VDI Berichte Nr. 1960. Integrierte Sicherheit und Fahrerassistenzsysteme*. Düsseldorf, Germany: VDI-Verlag, 2006, pp. 113–132.

[25] J. D. Hoffman, J. D. Lee, T. L. Brown, and D. V. McGehee, "Comparison of driver braking responses in a high-fidelity simulator and on a test track," *Transp. Res. Record.*, vol. 1803, pp. 59–65, Mar. 2002.

[26] N. A. Kaptein, J. Theeuwes, and R. van der Horst, "Driving simulator validity: Some considerations," *Transp. Res. Record*, vol. 1550, pp. 30–36, Jan. 1996.

[27] A. Kemeny, " Simulation and perception of movement," in *Proc. Driving Simulation Conf.*, 2000, pp. 13–22.

[28] A. Kemeny and F. Panerai, "Evaluating perception in driving simulation experiments," *Trends Cogn. Sci.*, vol. 7, no. 1, pp. 31–37, 2003.

[29] R. S. Kennedy, L. J. Hettinger, and M. G. Lilienthal, "Simulator sickness," in *Motion and Space Sickness*, G. H. Crampton, Ed. Boca Raton, FL: CRC Press, 1988, pp. 317–341.

[30] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness Questionnaire: An enhanced Method for quantifying simulator sickness," *The Int. J. Aviation Psychol.*, vol. 3, no. 3, pp. 203–220, 1993.

[31] R. S. Kennedy, N. E. Lane, M. G. Lilienthal, K. S. Berbaum, and L. J. Hettinger, "Profile analysis of simulator sickness symptoms: Application to virtual environment systems, *Presence*, vol. 1, no. 3, pp. 295–301, 1992.

[32] R. S. Kennedy, K. M. Stanney, and W. P. Dunlap, "Duration and exposure to virtual environments: Sickness curve during and across sessions," *Presence*, vol. 9, no. 5, pp. 463–472, 2000.

[33] J. D. Moss, J. Austin, J. Salley, J. Coats, K. Williams, and E. R. Muth, "The effects of display delay on simulator sickness," *Displays*, vol. 32, no. 4, pp. 159–168, 2011.

[34] R. R. Mourant, P. Rengarajan, D. Cox, Y. Lin, and B. K. Jaeger, "The effect of driving environments on simulator sickness," in *Proc. Human Factors Ergonomics Society 51st Annu. Meeting*, 2007, pp. 1232–1236.

[35] G. Moussa, E. Radwan, and K. Hussain, "Augmented reality vehicle system: Left-turn maneuver study," *Transp. Res. Part C*, vol. 21, no. 1, pp. 1–16, 2012.

[36] S. Mudd, "Assessment of the fidelity of dynamic flight simulators," *Human Factors*, vol. 10, no. 4, pp. 351–358, 1968.

[37] N. Mullen, J. Charlton, A. Devlin, and M. Bédard, "Simulator validity: Behaviors observed on the simulator and on the road," in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, J. D. Lee, Eds. Boca Raton, FL: CRC Press, 2011, pp. 1–18.

[38] K. von N-Cosel, M. Dupuis, and C. Weiss, "Virtual test drive–provision of a consistent toolset for [D,H,S,V]-in-the-loop," in *Proc. Driving Simulation Conf. 2009*, pp. 1–8.

[39] J. T. Reason and J. J. Brand, *Motion Sickness*. London, U.K.: Academic Press, 1975, pp. 1–321.

[40] L. E. Richardson and K. B. Shaw. (2009). Traffic safety in the European Union: The role of interest groups. [Online]. Available: http://www.unc.edu/euce/eusa2009/papers/shaw_03F.pdf

[41] D. R. Silverman and R. A. Slaughter, "An exploration of simulator sickness in the MH-60G operational flight trainer, and advanced wide field of-view helicopter trainer," Aircrew Train. Res., Mesa, AZ, Tech. Rep. AL/HR-TR-1994-0173, 1995.

[42] K. M. Stanney and R. S. Kennedy, "The psychometrics of cybersickness," *Commun. ACM*, vol. 40, no. 8, pp. 67–68, 1997. 3.

[43] K. M. Stanney, R. R. Mourant, and R. S. Kennedy, "Human factors issues in virtual environments: A review of the literature," *Presence*, vol. 7, no. 4, pp. 327–351, 1998.

[44] A. Starke and F. Hänsel, "Vehicle in the loop–fahrerassistenzsysteme mit virtual reality im realen fahrzeug testen, entwickeln und erleben," in *Proc. 12. Braunschweiger Symp.AAET*, 2011, pp. 230–244.

[45] H. A. Stoner, D. L. Fisher, and M. Mollenhauer, Jr., "Simulator and scenario factors influencing simulator sickness," in *Handbook of Driving Simulation For Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, and J. D. Lee, Eds. Boca Raton, FL: CRC Press, 2011, pp. 1–24.

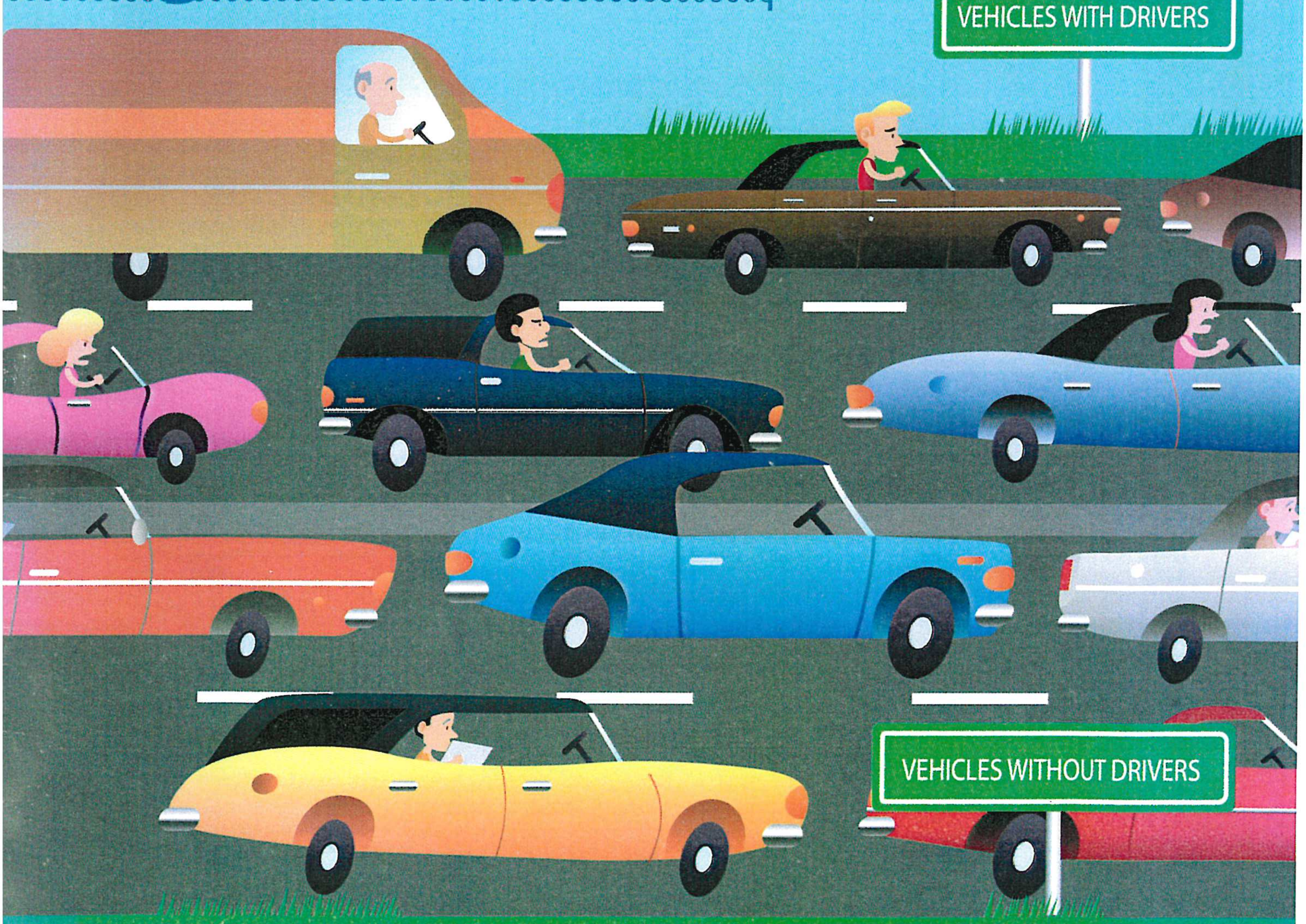
[46] J. Törnros, "Driving behavior in a real and a simulated road tunnel–a validation study," *Accident Anal. Prevent.*, vol. 30, no. 4, pp. 497–503, 1998.

ITS

IEEE

Intelligent Transportation Systems MAGAZINE

VOLUME 5 NUMBER 1
SPRING 2013



Automated Vehicles

On-Road and Off-Road

INTELLIGENT TRANSPORTATION SYSTEMS SOCIETY
ITSS
IEEE

IEEE