Assessment of Simulator Fidelity and Validity in Simulator and On-the-road Studies

Andreas Riener
Institute for Pervasive Computing, Johannes Kepler University Linz, Austria
Tel. +43/(0)732/2468–1432, Fax. +43/(0)732/2468–8426
Email: riener@pervasive.jku.at

Abstract*—A lot of research groups all over the world have tried to relate results from driving simulator studies to real driving behavior. A solution, e. g. in form of a conversion table, would be of great value. Unfortunately, status quo is that even with expensive, high fidelity simulators the validity of results cannot be guaranteed. One reason for this is that a person's behavior cannot be described by mathematical rules and depends, beside the task of interaction, on several subsidiary influence factors. Starting with an elaborate review of driving validity and fidelity constraints, the aim of this paper is to summarize on our research responding to the question to what extent driving simulators can be used to serve as cheap and easy realizable environments for simulating on-the-road behavior. The purpose of the studies was to determine (i) whether or not it is in general possible to approximate real driving with simulator studies, (ii) situation and modality dependent correction or scale factors to deduce real reaction times from simulation, and (iii) further requirements, parameters, and restrictions to be satisfied for succeeding high fidelity studies. Two user studies were conducted, a low fidelity trace-driven simulation in a lab environment and a on the road driving experiment. Recorded reaction times were compared in order to assess the validity of data generated in these experimental series. The events were, in the case of simulation, triggered trace-driven or, in the real driving experiment, manually activated by the experimenter and notifications were forwarded to the driver using a random assignment of one of the modalities vision, hearing, or touch. Results indicate that drivers responds faster to steering requests in the driving simulator compared to real driving. The explanation for this difference can most likely be derived from the fact that test persons were less demanded in the first (artificial) compared to the second (real) setting. When analyzing data on individual notification channel basis, it can be observed that (i) the order of channels with respect to average response times is the same in both settings (vibro-tactile, visual, auditory) and (ii) the reaction time differences are almost uniformly distributed. Prospective work in comparative studies is projected to happen in two directions, on improvements in the behavior of the low fidelity simulator to become as close to reality as possible and in the utilization of high fidelity driving simulators to directly relate real driving results to.

***Keywords*—Simulator validity/fidelity, Real-driving studies, Trace-driven simulation, Driver-vehicle interaction (DVI), Feedback modalities, Performance evaluation.**

I. Motivation: Simulation in the Car Domain

Simulation or driving simulator studies have become, for several reasons, state-of-the-art methods in the development process of vehicles. One factor for this is *economically justified* – research and development expenses for a new generation of vehicles is just about 5% on the overall costs of car production; however, car manufacturers are increasingly requested to operate as efficient in terms of cost as possible – for the area of development and design this would only be possible when applying computer assisted simulation techniques to all stages of development. Particularly in user interface design, the strong interrelationship between the driver, his/her personal preferences and the different control and assistance systems in a vehicle poses problems to be considered, and necessitates, in excess of pure simulation, a more detailed treatment by application of user studies and/or driving experiments.

Another factor is *time driven* – the car domain is today requested to shorter and shorter time-to-market cycles. The design life cycle for the automotive market has continuously decreased from over 4 years, approaching now 15 to 18 months – a delay of only one or two months can cost a manufacturer up to 30% of market loss [2]. This shorter cycles are caused by (i) driver assistance systems and control instruments catching on more and more into the dashboard, (ii) new forms of driver-vehicle interfaces established from one vehicle generation to the next, and (iii) trade rivalry by car manufacturers to outsell competitors which forces developers to make their own cars more attractive with regard to functionality, comfort, gadgets, etc., to a particular target market.

The third is *technology driven* – beside the before mentioned production-centric issues other challenges are expected to arise in the near future. Damiani *et al.* [p. 95][3] stated in the context of a discussion on attributes of future cars "*[..] Maybe new, unexpected needs and fashions will arise, but in any case the design and development of new technologies and devices will have to face the challenges opened by the new paradigms*". Overlooking the current research in the field of driver-vehicle interaction shows that it is moving toward enabling a full interconnection between drivers, vehicles, and infrastructure in order to increase driving safety and efficiency [4], [5], and in a next step to vehicles moving fully autonomic ("*[..] It is expected that men and women of the future when moving will continue their normal life, leisure and work while the car will take care of their safety*") [3].

With respect to these three pillars a performance and/or usability evaluation of user interfaces for (future generations of) vehicles is already today (and will be so even more in the future) infeasible in on-the-road experiments. This is to a lesser extent due to pure economical reasons, but most likely caused by (i) safety problems, may affecting test persons and other road participants, (ii) confirmed too long development, preparation and execution times to adhere to the strict production plans, and (iii) unknown real-time behavior of new interaction/communication concepts. Therefore, we (and many other research groups all over the world) recommend to use driving simulators to fulfill all of the specified requirements and basic conditions. In this way manufacturer should achieve both a successful launch of new vehicle models as well as a optimum, “relaxed” use while in operation.

Simulation in Automotive Applications

Our main motivation followed in this work stems from the third issue, *technological challenges*, as we are, for instance in the project SOCIONICAL [6], interested in the development of socio-technical systems on large scale. One of the investigated domains is the field of transportation, where we are amongst other things dealing with new concepts for driver-vehicle-infrastructure operation. (The task of driving can be described as a highly dynamic and local feedback loop between a driver and a car or its operational controls and assistance systems [7, p. 5]; the interaction between one person and the technology integrated into a single car can be denominated as the basic building block of a socio-technical system.) Within this field of research it would almost be impossible to conduct on-the-road studies evaluating the collective effect of hundreds or thousands of drivers operating their car in a certain region of interest.

Research branches currently under development and products already introduced into market confirm the need for simulation to avoid safety issues for both test persons and other road participants as well. The head-up display (HUD) available for years, for example in BMW’s premium type cars [8], serves as one instance. Ward *et al.* [9] discovered that the head-up display requires both a higher mental effort and a higher cognitive load from the drivers while the results of Nakamura *et al.* [10] and others showed reduced workload, decreased response times, more consistent speed, and increased driving comfort. A second example legitimating simulation is the emerging interest in olfactory interaction. Such systems have only be used rarely in the car to date [11], [12, p. 30]; however, it would be feasible to display a scent of burning oil in the passenger compartment to warn the driver or even to systematically employ odors to calm down the driver or strengthen his/her energy to increase driving safety – as it has been evidenced that the odor of jasmine or lavender elicits sedative or relaxant effects [13]. Nevertheless, the application of the same is questionable – on the one side as it still remains, most likely due to its subtle and imprecise perception, a developing research branch [14], [15] and at the other side, as particular fragrances won’t “work” for everyone [11]. Aside from this, it is known that the emotional state of healthy subjects has a clear effect on olfaction [16, p. 287]. As one example, a negative emotional state would reduce olfactory sensitivity (emotional states are likely to change quickly and uncontrolled during vehicle operation, e. g., in congested situations or on vehicles cutting in).

Research Hypotheses

The main reason arguing for simulation in the car domain is, that it can be applied in many critical areas and allow issues to be addressed before they actually become problems. Therefore, it is not “just a technology” because it forces one to consider global terms of system behavior, most frequently represented by complex models behaving in more than the sum of their components [17, p. 1].

Within this work we have engaged in monitoring, recording, and interpreting dynamic driving activities applied to a trace-driven simulation approach, and afterwards repeated in a real driving experiment. The focus of our research attempt was the investigation and classification of disparities using different feedback channels in the two settings. Moreover, we wanted to give answers to the question to what extent driving simulators can be used to serve as cheap and easy realizable environment for simulating on-the-road behavior, thus facilitating later user experiments to be laboratory based only. The declared aim of this work was to provide a metric for response time differences between simulation and the real world to be used as a conversion table to replace future on-the-road studies with simulation experiments. This suggestion can be assumed feasible as it has been shown for the automotive domain that simulation is already today a useful approach for data collection and driver behavior analysis [18], [19].

In particular, we hypothesize that

(i) Reaction times detected in real driving tests as well as in simulation studies are in the same order of magnitude for notifications with a certain sensory channel, so that a situation dependent correction factor can be tabulated first (separately for each feedback modality), and later used for deducing real reaction times from simulation (*relative validity*).

(ii) The different notification modalities available in the car (only the three main sensory channels vision, hearing, and touch were analyzed) behave, with regard to a single driver-vehicle interaction cycle, similar in their position/rank in on-the-road studies and simulator experiments.

As in detail elaborated below, the quality of a simulator is normally defined by its *validity* and *fidelity* – the stated research hypotheses are in line with these two qualitative simulator characteristics.

Outline

The rest of this paper is organized as follows: Section II gives an overview of the state-of-the-art of driving simulators, required for the development of the simulation environment used within our studies and furthermore summarizes predeterminations and regulations for both the simulated and the real driving experiments. Section III examines several aspects of user interface evaluation in the car domain while Section IV gives insight into both the driving simulator and the on-the-road experiments, their execution, and details on the logging of user data. Section V presents and discusses the findings of the comparison between these two experimental series with regard to qualitative aspects. The concluding Section VI summarizes the paper and gives suggestions for future work to improve the quality of the driving simulator in order to successfully replace real driving studies later.

II. Driving Simulators: Review of the State-of-the-Art

Driving simulators have a long history in the automotive domain, providing a means of precise design and control of scenarios, allowing for instrumentation and logging, and ensuring repeatability, e. g., for re-running situations with anomalous behavior. Simulators have been successfully applied for the evaluation of driving safety ([20], [21]) and driver behavior, such as influence on fatigue ([22], [23], [24]) or drunk driving ([25], [26]), education in driving schools ([27], [28], [29], [30]), for the design of new functions of vehicles ([31], [32], [33], [34]), and in particular for user interface evaluation ([35], [36], [37], [38], [39], [40], [41], [42], [43], [44]). To directly address the issue of ever decreasing production cycles, simulation has also been successfully applied for a long time to crash ([45], [46], [47]) and wind tunnel tests ([48], [49], [50], [51], [52]).

One of the greater problems in simulation is, that driver behavior (cognitive resources that drivers devote to both steering tasks and side activities) in simulated driving experiments may differ significantly compared to those deployed in real studies [20, p. 590], [53, p. 89] – most likely due to the fact of (i) complex person behavior representation ([54], [55], [56]), (ii) a discoverable lower focus on the tasks due to the risk free environment [1, p. 30], [53], and (iii) vehicle movement dynamics ([57], [58]) or motion cuing ([59], [60], [61]).

A. Qualitative Simulator Development

In the simulation domain the two aspects (i) fidelity and (ii) validity are used for characterizing the quality of simulators. Engströem *et al.* [62] indicated that there is a discoverable relationship between these two metrics as high fidelity simulators offer, per definition, a more realistic driving environment and, thus, a higher validity of obtained results as compared to low fidelity driving simulators with its accredited lower validity of data. For the implementation of a concrete simulator a trade-off between fidelity (the better the higher the development and running costs) and validity (greater validity is attributed to provide results more close to reality; however, some effects observed in low fidelity simulators may be are not appearing in high fidelity simulators) has to be set based on trade-off between the requirements of the studies to be conducted later and the budget available for implementing/constructing the simulator.

(A) Fidelity: Fidelity is often equated to the level of realism represented in the simulation [53, p. 89] (it has been reported that the closer a simulator approximates the real world in terms of design and layouts of control, realism of the shown scene, etc. the greater is the fidelity of this simulator). For a better classification of simulators, Rehmann [63] has proposed to use four interrelated fidelity dimensions.

As stated above, high fidelity often goes along with high driving simulator costs; particularly for human factors research the costs of driving simulators are often very high, prohibiting their application so that automotive human factors research related to the evaluation of interface design is usually to date still done in on-the-road experiments [64].

Consequences (for this work): With respect to the designated fidelity and the high costs expected for simulators required for human factors research, we designed the driving simulator as used in the studies explained below in a opposite direction, achieving relatively high fidelity at low cost: High *equipment fidelity* was achieved by using a real car for our simulator tests (the same as later used in the real driving experiments), *environmental fidelity*, such as motion cues and sensory information obtained from the real world environment, was, however, fulfilled to a lesser extent (motion cues were not percepted at all as the car was parked (jacked up) in a garage, sensory information from the outer world was either not available or not related to the scene shown in the simulator). *Objective fidelity* is said to be related to dynamic cue timing and synchronization issues, e. g., between steering input and corresponding visual cues. This point was almost fulfilled as the shown driving scene, the notification patterns delivered to the driver (visual, auditory, vibro-tactile), and the vehicle controls used by the driver to react on perceived stimuli were all connected and synchronized one with the other on a single computer (a correct timely behavior was one of the most important design requirements). The last dimension of Rehmann's four-staged classification proposal, *perceptual fidelity*, is concerned with (i) the degree to which the driver perceives the simulation to be a reproduction of the real driving task and (ii) the degree to which a driver's interaction with the driving environment corresponds to real-world driving [63]. Both sub-points are fulfilled to a medium to high degree as the scene used in the simulator was a prerecorded video trace of a real journey through the city which was synchronized to stimulations (vehicle feedback). The corresponding feedback was collected from the user via the real control instruments in the car (i. e., turn indicators, light switch).

(B) Validity: The second qualitative aspect to describe a simulator is its validity [65]. It typically refers to the degree to which behavior in a simulator corresponds to behavior in real-world environments under the same conditions and maybe is, as stated before, affected by the level of fidelity. Validity is much harder to achieve as fidelity as it is dependent on many (human) factors like rewards and punishments for "appropriate" driving behavior [66], cognitive workload levels and psychological environment [67, p. 315], different levels of stress [68], etc. Furthermore, it has been shown that individuals can experience symptoms of sickness (like dizziness or headache) in driving simulators [66], making it difficult to relate results to driving behavior observed in reality.

According to Young *et al.* [53], the best method for determining the validity of a simulator is to compare driving performance in the simulator to driving performance in real vehicles under the same driving tasks. For validity, two types can be differentiated, (i) absolute and (ii) relative validity. The former is achieved if the numerical values obtained from driving simulator studies and on-the-road tests are identical (or near identical), the latter when driving tasks in the simulator and in real driving studies have a similar affect (e. g., a similar direction of change) [67]. The feasibility of this recommendation was shown by Lee [69], who confirmed (i) the validity of driving simulator studies and it's on-the-road driving counterpart and (ii) that it is a safer and more economical method than the on-road testing to assess the driving performance of (older adult) drivers.

Consequences (for this work): In order to determine the validity of our approach on simulated driving we have conducted a second experiment using an on-the-road driving study. For that the same (physical) car, the same notification patterns, and the same operational tasks to be completed by the test drivers were used. Beside these basic factors we selected a similar length of the route (with regard to driving time) for the real driving experiment and used days with low traffic to avoid distraction from other road participants as good as possible.

To come to the point, the results obtained in our studies when comparing the two experimental series are satisfying (a detailed investigation is given in Section IV). Although absolute validity of the simulation environment has not been demonstrated, relative validity has been achieved. In order to further strengthen the plausibility of a relationship between the two approaches, several improvement potentials have been identified during experimental conduction – these are summarized in the last section of this paper.

III. Vehicle UI's: Evaluation using Simulators

Research on new automotive systems currently relies on car driving simulators, as they are a cheaper, faster, and safer alternative compared to tests on real tracks. However, there is increasing concern about the fidelity of results provided from the simulator and their influence on the validity of these studies in the reality. Especially for motion cuing, and here for high-speed curve driving, the provision of large sustained acceleration would be difficult to be reproduced in the limited motion space of simulators.

User interface (UI) composition and evaluation is a challenging task in the design phase of a new vehicle generation, particularly today where the time-to-market cycles decreases steadily, reaching 15 to 18 months after up to 48 months a few years ago [70]. This reduction in development time would only be possible when applying simulation to all stages of the vehicle manufacturing cycle, even to the design and evaluation of user interfaces [71]. Solutions like Virtual Product Development (VPD) and Computer Aided Engineering (CAE) techniques [70] are applicable to "hardware design", nevertheless these approaches are not suitable for the development of user interfaces without further considerations regarding the user as they are highly dependent on users preferences – and a person's behavior does simply not follow mathematical rules or physical laws.

However, (full-motion) flight simulators, which are situated between "model-based" simulation and tests in the real world, have been successfully showed its applicability for pilot training in the past decades (see, for instance, [72], [73], [74] or [75]). Following this approach, driving simulators should be established for user interface testing in the automotive domain as well. Like in the flight simulator, tests with vehicle simulators can be conducted in a quiet, controlled test environment with the following advantages: (i) mistakes can be reviewed immediately, (ii) a failed task can be repeated by rewinding and replaying the scenario, (iii) the "driver" is secured from accidents and other road participants do not need to be endangered, and (iv) user's concentration is on the task, not on the noisy, stressful environment [76].

It is supposed, and has already been shown for the evaluation of single driver-single vehicle interaction issues, that driving simulators can be successfully applied for user interface evaluation (at least for a certain level of validity) [77], [78], [79] or in the automotive electronics Journal [80].

Summarizing the considerations regarding simulation poses the problem of a still missing "reality effect", as in real situations aggregated, e. g., from motor noise, environmental sound (raindrops pattering on the front windshield, honking cars, etc.), road vibrations (providing implicit knowledge about driving behavior, such as drifting on gravel or snowy roads), penalties for driving violations, danger of road accidents, etc. – even if the replayed scenery in the simulator looks highly realistic and the experiments are processed in a "real" physical car (as done in the trace-driven simulation experiment described later in this work). For driving experiments using a simulator instead of real driving studies it

Figure 1. Experimental setting for the trace-driven experiment with vehicle control signals taped via Atmel μC and forwarded to the control application (leftmost image), screen shot of the prerecorded test run as shown the participants during the experiment (rightmost picture).

would be rather important to know these missing parameters or at least their joint impact on the distortion of results. If it is possible to provide evidence for a common or individual "correction factor(s)" or to build a "correction-model", it would no longer be required to conduct expensive (in terms of cost and time) real-driving studies. Nevertheless, in any case, and according to findings obtained so far, it will be almost impossible to cover the full range of errors, and thus to fully replace on-the-road studies with simulation.

Authenticity: Example of alcohol impaired driving

Before replacing user studies with simulation experiments on a broader basis it has to be approved that the results of the latter are equal or comparable to that of the former, at least for the key issues. If authenticity for both test persons and achieved results can be confirmed it would be much easier (or even yet feasible) to perform specific tests by simulation, and to give expressive statements on the results for the reality. For instance, it is not recommended to study alcohol impaired driving or the influence of fatigue in on-the-road studies, however, an investigation by using a driving simulator would be possible at no risk and is assumed to provide meaningful results directly transferable "back to the real traffic". To substantiate this implication, the research conducted by Robbe [81] can be quoted, performing one laboratory and three on-road driving studies (restricted highway, normal highway, urban traffic) concerning the effects of marijuana smoking on actual driving performance, and detecting almost the same results.

IV. THE EXPERIMENTS

A driving simulator was developed initially, measuring response times from a driver on a limited number of vehicle control operations, notified either visually, auditory or vibro-tactile. The experiment was processed in a real car parked in a garage, with "real" vehicle controls such as turn indicator or light switch connected to a microcontroller (see Figures 1, 4). Instead of a computer-generated scenery, a prerecorded journey of about $30min.$ in length across the city of Linz was replayed on a large projection screen placed in front of the windscreen. The detailed setting of the experimental system as well as a in-depth description of the conducted experiments including evaluation and results is given in Riener *et al.* [31].

After this first experiment – which has already been defined considering a later reuse in a real driving study – a similar test series (using the same car, the same test participants (a subset of them), the same tasks, the same notifications) was prepared and performed in an on-the-road scenario as described below. This second series was conducted several month later so that dependencies between the two experiments like learning effects of test participants should not have been observed.

As the detailed results of both the simulation experiment and the on-the-road study are published ([31], [1]), we will subsequently focus more on the specialties of the two settings, accounting differences between the two approaches (wherever applicable), and conclude with a elaborate interpretation of achieved results – the last two Sections V and VI are dedicated to this objective. The two experiments will be there compared in detail and, more interesting, possible effects will be discussed.

Geographic Similarity of Experiments

Both the simulation experiment (actually the scene shown in the simulator) and the on-the-road driving tests were conducted in the greater area of Linz, Austria. For the former, a $21km$ (original driving time approximately $30min.$) long trip cross-town the city of Linz was recorded, containing controlled and uncontrolled crossings, road tunnels and freeway components (the video camera was mounted in the car in order to tape the view of the driver). Waiting times on crossings and unsubstantial or pointless driving sections were removed from the taped run (in a way indiscernible for the test driver), afterwards the remaining video was tagged with action/notification points. On the other side, the real driving study was conducted in and around the city of Perg, $25km$ east of Linz, mainly due to the lower volume of traffic and thus, a reduced risk of incidents. All the test runs were processed here on a predefined, fixed course (see Figure 2) with a circuit length of $25.79km$; an entire run lasts on average $34min.$ It has to be noted that vehicle specific data from the controller area network (CAN) bus (via ElmScan5 USB ELM327 on board diagnostics (OBD) interface) and global positioning system (GPS) position as well as vehicle acceleration data (from RaceTechnology [82] DL2 data logger with integrated $20Hz$ high accuracy GPS tracker and IMU06 six degree of freedom inertial measurement unit) were gathered for a different purpose, and not for evaluation or data interpretation in the context of this work (to be correct, they were only be used for visual inspection; additional data correctness verification was not applied).

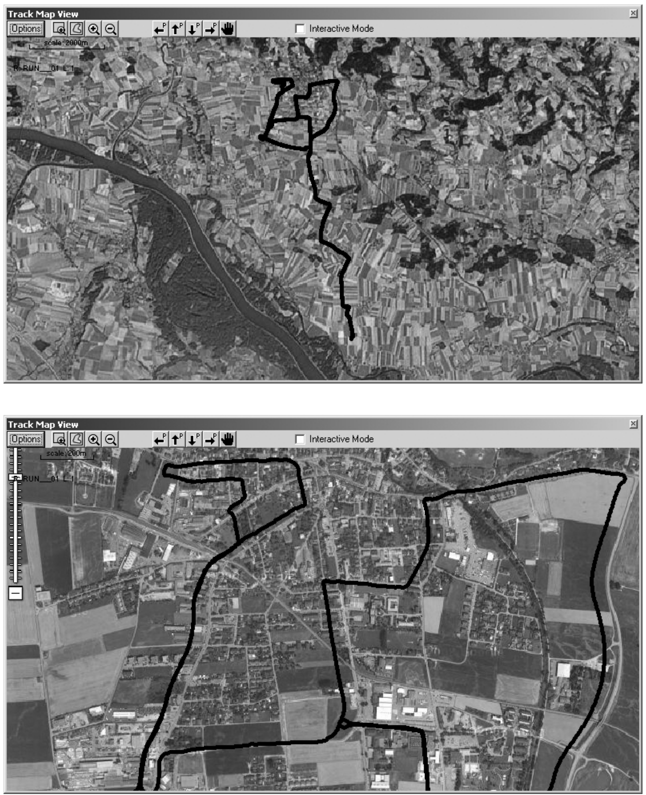

Figure 2. GPS traces of the predefined driving route (length of $25.79km$) with subjacent satellite maps (the upper image shows the overall journey (similar initial and final points), the lower displays a detailed view of the trip across the city of Perg).

Differences in Data Acquisition

The main difference between the two settings is grounded in the initiation of notifications. In the experiment using simulation it is trace-driven and time aligned to the video of a prerecorded journey, in the real study it is executed manually by the experimenter according to predefined positions in the driven route (a person initiating the feedback was seated on the back seat behind the driver so that he/she – and the actual task of activation – could not be seen, neither be guessed by the driving person; once the activation key was pressed (=task notification) one sensory channel out of the three available modalities was chosen randomly by the software).

For both the trace-driven experiment and the on-the-road series exactly the same setting was used. Notifications about required driving activities were delivered to a particular driver using either a visible, audible or vibro-tactile feedback signal. Reaction times from the driving person were then collected from the real control instruments of a car (e. g., turn indicator, light switch) connected to a Atmel AVR ATMEGA 8 microcontroller and forwarded to the capturing software.

```
------------------------------------------------------------
Real-Driving Journey started at 30 March 2009 10:04:51
------------------------------------------------------------
Visual Task TurnRight 0 created. (10:05:01734)
Visual Task TurnRight 0 finished: 1,141 ms (10:05:02875)
Visual Task TurnLeft 0 created. (10:06:1393)
Visual Task TurnLeft 0 finished: 935 ms (10:06:02328)
Vibro-tactile Task TurnLeft 1 created. (10:07:42734)
Vibro-tactile Task TurnLeft 1 finished: 1,016 ms (10:07:43750)
Auditory Task LightsOn 0 created. (10:08:05234)
Auditory Task LightsOn 0 finished: 1,412 ms (10:08:06646)
.
.
.
Visual Task TurnRight 14 created. (10:37:01343)
Visual Task TurnRight 14 finished: 1,094 ms (10:37:02437)
Auditory Task TurnLeft 13 created. (10:38:15640)
Auditory Task TurnLeft 13 finished: 860 ms (10:38:16500)
------------------------------------------------------------
Real-Driving Journey stopped at 30 March 2009 10:38:35
Stopping unfinished tasks...
Total number of tasks: 31
------------------------------------------------------------
```

Figure 3. Abstract of the log file of one particular driving experiment. Time flags in parenthesis are used for synchronization with other vehicle-specific data on basis of GPS time.

Logging/Recording: Log files were compiled for each trip in the two experimental series for evaluation purpose. Figure 3 exemplarily shows an abstract of such a log file for a real driving trip employing the following list descriptors.

(i) Selected notification channel: In the actual studies one out of *visual*, *auditory* or *vibro-tactile* (unimodal feedback). The software was designed modularly expandable, so that this list could later be extended, for instance with notifications using the olfactory channel (*smell*) or by combining notifications from more than one sensory modality at a time to a multimodal feedback system.

(ii) Kind of activity: This is an indicator for the action to be performed by the test person. Within the here discussed experiments we used the activities (i) turn right, (ii) turn left, (iii) lights on, and (iv) lights off. Each activity is followed by an enumerator counting the actual number of occurrences of that activity (starting from zero).

(iii) Indicated command: This field is one of either "start" (*created*) where user notification is initiated, or "stop" (*finished*) where the response from the user was recognized.

(iv) Response time: This field is defined only for the *finished* command and indicates a driver's reaction

Table I
Statistics on reaction times using a 5% confidence interval (time values given in ms).

Trait	Min x_{min}	Max x_{max}	Mean $\overline{x}$	Median $\widetilde{x}$	Std.Dev. σ	Min x_{min}	Max x_{max}	Mean $\overline{x}$	Median $\widetilde{x}$	Std.Dev. σ
	Trace-driven simulation (CI 5%, 752 datasets)					On-the-road studies (CI 5%, 353 datasets)				
Combined	281.0	1,985.0	889.2	812.0	349.9	203.0	1,750.0	1,003.2	1,000.0	331.9
Visual	391.0	1,922.0	784.3	703.0	295.8	360.0	1,750.0	978.7	940.0	328.7
Auditory	641.0	1,984.0	1,129.6	1,078.0	269.6	203.0	1,719.0	1,179.5	1,195.0	298.6
Vibro-tactile	281.0	1,625.0	690.6	641.0	255.9	265.0	1,672.0	879.9	828.0	299.3

time from one particular stimulation activity. It can be calculated simply by using the formula

$$t_{response} = t_{finished} - t_{created} \quad (1)$$

(v) Actual time: This field contains Internet synchronized time (UTC) to allow data alignment between measures recorded with the local acquisition system connected to the notebook computer with vehicle specific measurements and GPS data recorded with the RaceTechnology [82] data logging equipment.

User Studies

Two user studies were carried out to prove the research hypotheses (i) reaction times are in the same order of magnitude and (ii) ranks of channels are similar. The lab-based simulation experiment was conducted with eighteen volunteers, the later executed field study with twelve voluntary test persons.

Trace-driven simulation: Eighteen persons (15 male, 3 female subjects) in the age range from age_{min}=18 to age_{max}=38 years ($\overline{age}$=25.00 years, σ_{age}=5.12 years) participated in the first experiment on trace-driven simulation (Figure 4). All of the test persons were relatives, colleagues at the university and students with a valid driving license and a driving experience of on average more than seven years. Not a single one was in a relationship with our department, had previous knowledge about the aim of the experiments nor experiment related interaction with participants driven previously. Male subjects vary in age from age_{min}=18 to age_{max}=38 years ($\overline{age}$=24.80 years, σ_{age}=5.41 years). Female test persons vary in age from age_{min}=22 to age_{max}=30 years ($\overline{age}$=26.00 years, σ_{age}= 4.00 years).

On-the-road study: In contrast, the second study was conducted on a smaller group of twelve test persons whereof 7 subjects were male (58.33%) and 5 participants were female (41.67%). Most of the attendees of this second experiment already participated in the first experiment. Male subjects vary in age from age_{min}=25 to age_{max}=55 years ($\overline{age}$=33.43 years, σ_{age}=10.63 years). Female test persons vary in age from age_{min}=26 to age_{max}=52 years ($\overline{age}$=39.20 years, σ_{age}= 11.21 years), the overall mean age is 35.83 years and the standard deviation 10.78 years.

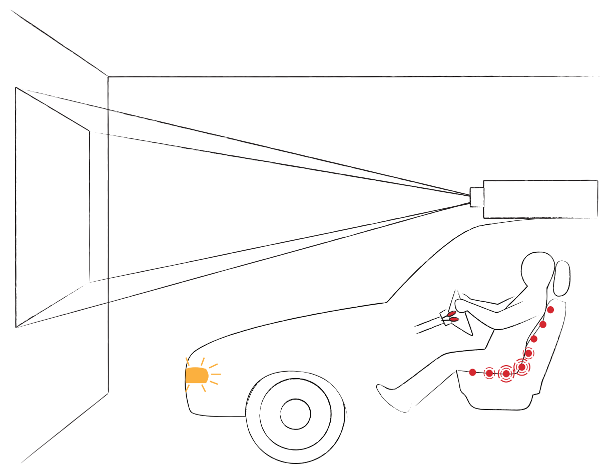

Figure 4. Scribble of the driving simulator setting as used for trace-driven simulation. The car was parked in a darkened garage with vehicle controls connected to a microcontroller. A video projector was used to generate a close-to-reality street view in the front shield.

Although the first experiment was conducted on a larger group compared to the second, results of the two series should be directly comparable. Nevertheless, one particular issue, age dependency, has to be considered accessorily as it has been evidenced earlier that the age of test persons has a major impact on the measured reaction time [7, p. 224] ("*[..] the ability to respond decreases with age.*"). From the difference in the mean age of the two test series (35.83-25.00=10.83 years) we can already prior to experimental results draw the conclusion that the average reaction time is higher in the real driving study (=slower reaction) compared to the simulator experiment conducted with younger people.

Trace-Driven Simulation versus Real Driving Studies

According to [53, p. 91] our simulator belong, as the bigger part of all simulators used in this domain, to the class of simulators with *relative* validity (*absolute* validity is the "ultimate goal" to achieve, but can only be reached using high fidelity, high priced driving simulators – and even for those, this criterion has only been demonstrated rarely to date). Nevertheless, the success of any simulation model or simulator is based on how effective a simulator can translate real-world situations and the manner that physical elements for the real world that plays an active role in the choice process are represented [83]. For the comparative study in this work a trace-driven simulation approach was chosen, operating, rather than as a common driving simulator, as tool for real-time driving decision-making.

The "simulator" was deployed to investigate the performance of driver response times using different notification channels on a prerecorded, typical trip across Linz – detailed findings from the tests are given in [7], [31].

Beside the accepted motivation for processing simulation experiments, such as safety, feasibility, independency, repeatability, or comparability, the aim in these series was more to assess a "kind of baseline" for the on-the-road studies to be performed later.

After extensive tests with the simulated driving environment and several modifications in the driver-vehicle interaction loop with the aim that the simulator was intuitively understood by all test participants, the final setting of the simulation experiment, as used in the first study, was transferred to and repeated in a similar designed on-the-road scenario. The goal in the real driving experiment was to provide evidence for similar system behavior embodied by a comparable reaction performance or workload of the driver. On successful proof this would legitimate further engagement in improving the driving simulator by considering parameters influencing the real driving performance to behave as realistic as possible in terms of cognitive workload, distraction, and reaction time in order to finally replace any real driving study with an equivalent simulation run.

Tests in real traffic situations requires a substantially higher effort for preparation compared to experiments with a simulator. Before starting to drive each test person got a detailed initial training ("dry simulation") in order to avoid (or at least reduce) the probability of later accidents or danger situations due to misconceived action triggers to a minimum. This preparatory stage lasts about $20min.$ per person, the driving experiment started immediately afterwards and took, on average, $34.3min.$ for the $\approx 26km$ round trip.

Basically, the setting was similar to that of the simulation environment; however, the number of action points was with 35 lower than in the first series (44). Each test person had to drive exactly the same predefined route (which was also different to that used in the trace-driven simulation) with notifications delivered on specific points of the route using random feedback channels. Visual notifications were given on small displays ("jumbo LEDs") placed left and right on top of the dashboard, auditory information was delivered via headphones, and vibro-tactile information was transmitted via sixteen tactor elements (two strips of eight each) integrated into the car seat (Figure 5). For measuring reaction times, the signals from the control elements activated in reality (light switch, turn indicator) were captured by a microcontroller and forwarded to and post processed in the data analysis unit (standard notebook).

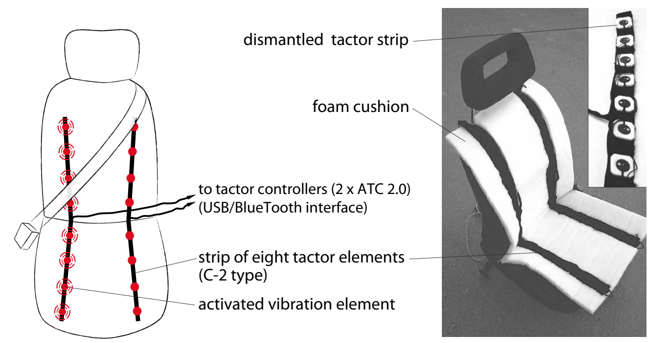

Figure 5. Vibro-tactile notifications were delivered via the driver seat. Therefore, two strips of eight voice coil transducers each were embedded into the seat. The foam cushion was applied to ensure comfortable sitting.

V. Findings

Table I gives a summary of data analysis separated for the two experimental series. Particularly the mean reaction time ($\overline{x}$) and the SD (σ) are of interest for further examinations with respect to the comparability of the experiments (see Table II).

Table II
Results show both increased average reaction times and standard deviation for real-driving journeys compared to trace-driven simulation experiments.

Attribute	Reaction Time, Increase in%			Order
$TD \rightarrow R$	$\overline{x}$	$\widetilde{x}$	σ	TD, R
CI 5% (752, 353 datasets)				
Combined	12.8	23.2	-5.1	-, -
Visual	24.8	33.7	11.1	2, 2
Auditory	4.4	10.9	10.8	3, 3
Vibro-tactile	27.4	29.7	16.9	1, 1

A. Variance

From Table I it can be assessed that the standard deviation of reaction times is similar for both test series ($349.9ms$ trace-driven, $331.9ms$ on-the-road; difference of 5.14% in favor of real-driving studies). Data inspection on modality level (Table II) shows more representative results – the differences are here 10.8% for auditory notifications, 11.1% for visual, and 16.9% for vibro-tactile stimulation, in each case in favor of trace-driven simulation.

Interpretation: The variance of reaction time over all participants for a certain experiment is relatively stable at $300ms$ (see Table I). Cross *et al.* [84] observed increased stability of (finger) reaction times as an inverse function of age – examining the individual notification channels confirms this finding as the mean age in the case of simulation is 25.0 years compared to 35.8 years for the real driving study.

It can be assumed that the variance of reaction time is rather independent from the channel of notification and whether the experiment is processed in the real or as simulation – the only reason for the larger variability should be based on the age difference of person groups.

For definitive confirmation it would be essential to conduct further studies with test persons in a broader range of age, e. g., age group 18 to 65. Initial evaluations with respect to age were presented in Riener [85] for the trace-driven experiment – the order of variance ($364.94ms$ for the group of persons aged 25 years or below, $337.37ms$ for the group older than 25 years) follows the findings presented by Cross and Luper [84], and are confirmed within this work.

B. Reaction Time

Reaction time is attributed to cover (i) the time required to perceive the need for an action, (ii) thoughts about how to solve the problem, (iii) the selection of a solution, and (iv) the initiation of motor actions. The mean reaction time for the two experiments differ by 12.8% (Table II) in favor of simulation. When comparing the on-the-road studies to the simulation experiment based on individual modalities, the reaction time increase is 4.4% for auditory, 24.8% for visual, and 27.4% for vibro-tactile delivered notifications (visually represented in Figure 6).

Interpretation: The significant increase in reaction time for on-the-road studies, discoverable over all modalities, can be explained by several reasons and is confirmed, at least to some extent, by surveys carried out in connection with the experiments.

The driver, while steering a "real car", is (or should be, for safety reasons) focused on the main task of driving. Therefore, Erp [86, p. 2] proposed the safety strategy "*[..] vehicle operation should allow the driver to have both eyes on the road, both hands on the steering wheel, and both feet on or near the pedals*". In the last time it can be increasingly observed that he/she is distracted by secondary (operation of driver assistance and/or information systems) or tertiary tasks (communication with passengers, adjusting car stereo, and operation of other comfort/entertainment services) in the car [87], [88]. Even if these additional distraction factors were limited as good as possible for the conducted experiments (e. g., no car passengers, car stereo switched off, air conditioning system fixed, etc.), the driver is distracted to some extent, and has therefore limited capacity free for perceiving and reacting to actions.

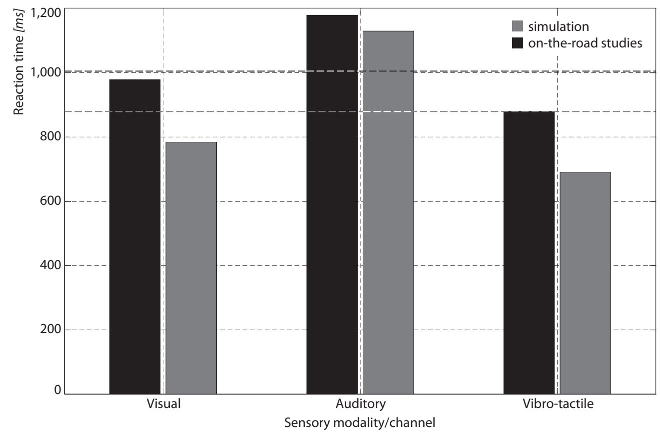

Figure 6. Mean reaction times for the three feedback channels visual, auditory, and vibro-tactile, isolated for simulation (gray) and real journeys (black). Dotted lines at 889.2 and 1,003.2ms indicate the overall mean reaction times.

The large increase in the mean reaction time between simulation and on-the-road study for the visual notification channel (24.8%) (Table II) can be explained by the fact that driving is mostly a visual task, demanding much higher attention when driving in real traffic compared to controlling a simulator. The behavior of the driver in the simulation environment has no impact to the real world (there is no "real danger" – neither for the driver, nor for other road participants, pedestrians or the infrastructure), so that the driver can completely focus on the task of vehicle control. Furthermore, visual notifications were overlayed to the replayed video in the simulation while this information was provided around the dashboard for the real experiment. This requires glances with re-focusing the eye for the latter study, which is well known to require some extra time. Results would be definitely better comparable when showing information in the on-the-road experiment in a more natural way, using, for instance, a head-up display to provide the driver with information without taking his/her eyes from the windscreen.

The vibro-tactile stimulation channel, ascribed to be uninfluenced from the cognitive load of visual and auditory senses, was added accessorily. It is supposed that the increase in the mean reaction time between simulation and real driving tests (27.4%, see Table II) results from the uncommonness of using the sense of touch as information carrier. For the trace-driven simulation, test persons are willing to trust this modality, but in real-driving studies users are more cautious as operating errors are prone to run unnecessary risks. The histograms shown in Figure 7 reflect these assumptions as the reaction time is on average lower with less variance for the first (simulation) compared to the latter experiment (on-the-road studies). We are confident that the great difference in reaction time declines (or even disappears) with common utilization of the sense of touch

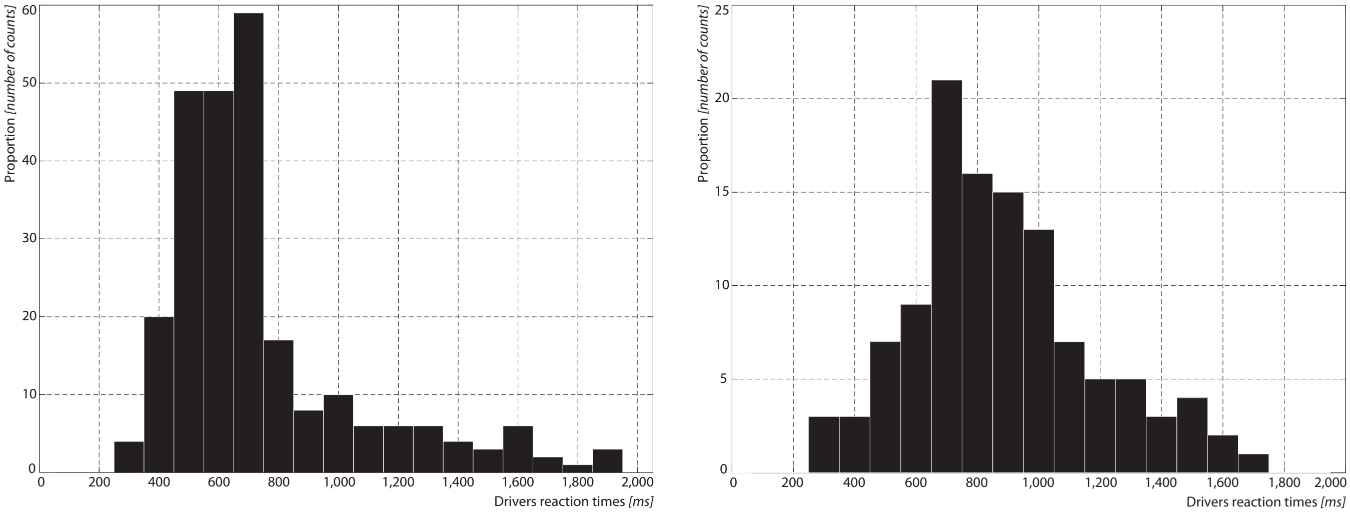

Figure 7. Histograms for trace-driven simulation (left) and on-the-road experiments (right) for the vibro-tactile modality. The shape of the plot shows that reaction times are lower with less variance for the case of simulation compared to real-driving studies.

in driver-vehicle interfaces (the characteristic of touch-based feedback was already discussed before, e. g., in [7, p, 194]).

C. Further Results

The surreal behavior of a driving simulator is another factor influencing the performance compared to on-the-road studies. For simulation, we found in particular that (i) a discoverable lower concentration is on the task of driving due to the risk-free environment, (ii) traffic rules (road signs such as speed limits) and road traffic regulations can be (and are) ignored by reason of no punishment on delicts, and (iii) the general unreal behavior of the simulation environment (absence of engine and environmental noise, road vibrations, etc.) has a negative impact on the validity of results.

It should also be noted that the comparison lacks of equal preconditions, for instance in the simulated experiment – contrary to the real driving study – test persons were actually not involved in a driving task (they had only to watch the video and react on a requested feedback using the real control instruments of the car). Future simulation settings, desired to provide a more realistic behavior, should be designed under the guidelines to cover the identified issues.

VI. Conclusion and Future Work

With increasing pressure regarding production cost and time, automobile manufacturer are requested to apply simulation to all stages of product development including user interface evaluation. The application of driving simulators is, particularly for the last item, a great challenge as a person's behavior cannot be described by mathematical equations or physical rules. Moreover, and beside the primary task of interaction with the vehicle, driver behavior depends on several subsidiary factors of influence.

In order to gain insight in vehicle steering performance, two driver-car interaction experiments of similar type have been designed and processed, first, a *labor study* (trace-driven simulation) and second, a *real driving experiment* (on-the-road journey). Comparing the average reaction times of drivers based on notifications using different sensory channels confirmed that the two instances perform similar; for both studies, response time from vibro-tactile delivered commands were lowest, followed by responses from visual and auditory stimulation. Furthermore, results have shown that the reaction time in real driving situations is on average 13% higher compared to simulated driving. The main reason for this large difference is, that the simulation environment is only a imprecise, low detailed copy of the real environment (static car in a garage with a video of the prerecorded track displayed on the front shield; no real driving task). Therefore, the difference cannot be considered as universal, linear correction factor to obtain real response times from driving simulator studies.

According to the detailed elaboration of simulator *validity* and *fidelity*, there are two strategies for future improvements. The first is to use a more sophisticated simulator for upcoming experiments, providing an immersive environment with close-to-reality behavior (road vibrations, engine noise, penalty models for speeding, etc.). With such a simulator, using enhanced settings to cover the discussed problems, it should be feasible to estimate the differences in reaction time between simulator and on-the-road study in a more precise way. Such high end simulators must not be developed in-house. There are several research institutions owning (and leasing) simulators, such as, for instance, simulators at TNO Netherlands [89], DLR Germany [90], TRW Automotive [91], or IZVW Wurzburg/Germany [92]. Nevertheless, high fidelity simulators are very expensive (even to rent) and it has been shown that also these systems cannot guarantee validity. From this state of knowledge, it should be clear to put not more effort into the development of high fidelity simulators.

The second option is to improve both quality and comparability of the existent systems based on the findings from the initial experiment series summarized in this work. One issue to cope with is, for example, to adjust the cognitive workload the driver is exposed to in both settings to be similar, e. g., by adding a task to the simulation experiment or by increasing the complexity of the simulator to behave more realistic. It is supposed, that simulator outcome would then be better comparable to real driving results. If approved, a metrics table could be provided, containing rows for notification channels and columns for the level of simulator fidelity, e. g. low fidelity simulator, visual notification, add $x\%$ to get real world behavior or high-tech simulator, visual channel, add $y\%$ (with $y < x$); values should be provided for all the modalities.

ACKNOWLEDGMENTS

This work is supported under the FP7 ICT Future Enabling Technologies program of the European Commission under grant agreement No. 231288 (SOCIONICAL).

REFERENCES

[1] Andreas Riener. Simulating On-the-Road Behavior Using a Driving Simulator. In *Third International Conference on Advances in Computer-Human Interactions (ACHI '10)*, pages 25–31, February 2010. ISBN: 978-1-4244-5693-2.

[2] W. Fornaciari, M.V. Angiolillo, and R. Farina. Seamless design space exploration for automotive systems. In *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on*, pages 241–244, 2007.

[3] Sergio Damiani, Enrica Deregibus, and Luisa Andreone. Driver-vehicle interfaces and interaction: where are they going? *European Transport Research Review*, 1(2):87–96, July 2009.

[4] Alois Ferscha and Andreas Riener. Pervasive Adaptation in Car Crowds. In *First International Workshop on User-Centric Pervasive Adaptation (UCPA) at MOBILWARE 2009, Berlin, Germany*, page 6. Springer-Verlag Berlin Heidelberg, April 27 2009.

[5] Simon Nestler, Marcus Tnnis, and Gudrun Klinker. Common interaction schemes for in-vehicle user-interfaces. *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction*, pages 159–168, 2009.

[6] SOCIONICAL. Complexity Science based modeling, prediction and simulation methods for large scale socio-technical systems. [Online]. last retrieved January 9, 2011. http://www.socionical.eu.

[7] Andreas Riener. *Sensor-Actuator Supported Implicit Interaction in Driver Assistance Systems*. Vieweg+Teubner Research, Wiesbaden, Germany, 1st (January 14, 2010) edition, January 2010. ISBN-13: 978-3-8348-0963-6.

[8] BMW. BMW Technology Guide: Head-Up Display. [Online]. last retrieved January 9, 2011. http://www.bmw.com/com/en/insights/technology/technology_guide/articles/head_up_display.html?source=index&article=head_up_display.

[9] N.J. Ward and A. Parkes. Head-up displays and their automotive application: An overview of human factors issues affecting safety. *Accident Analysis & Prevention*, 26(6):703–717, 1994.

[10] K. Nakamura, J. Inada, M. Kakizaki, T. Fujikawa, S. Kasiwada, H. Ando, and N. Kawahara. Windshield Display for Safe and Comfortable Driving. *SAE World Congress on Intelligent Vehicle Initiative (IVI) Technology 2005, Advanced Controls and Navigation Systems*, 2005.

[11] Tony Swan. 2009 Tokyo Auto Show: At Nissan, Driving Pleasure Includes Olfactory Gratification. [Online.], October 22, 2009. last retrieved January 17, 2011. http://blog.caranddriver.com/2009-tokyo-auto-show-at-nissan-driving-pleasure-includes-olfactory-gratification/.

[12] Gran Kecklund, Torbjrn Akerstedt, David Sandberg, Mattias Wahde, Tania Dukic, Anna Anund, and Magnus Hjlmdahl. State of the art review of driver sleepiness. Deliverable 1.1, Intelligent Vehicle Safety Systems (IVSS) program project, April 2007. [Online]. last retrieved January 17, 2011. http://www.vti.se/11934.epibrw.

[13] Kyoko Kuroda, Naohiko Inoue, Yuriko Ito, Kikue Kubota, Akio Sugimoto, Takami Kakuda, and Tohru Fushiki. Sedative effects of the jasmine tea odor and (r)-(–)-linalool, one of its major odor components, on autonomic nerve activity and mood states. *European Journal of Applied Physiology*, 95(2):107–114, October 2005.

[14] Joseph Nathaniel Kaye. Symbolic Olfactory Display. Master thesis, School of Architecture and Planning, Program in Media Arts and Sciences, MIT, May 2001. [Online]. last retrieved January 11, 2011. http://alumni.media.mit.edu/~jofish/thesis/symbolic_olfactory_display.html.

[15] Adam Bodnar, Richard Corbett, and Dmitry Nekrasovski. AROMA: ambient awareness through olfaction in a messaging application. In *Proceedings of the 6th international conference on Multimodal interfaces (ICMI'04)*, pages 183–190, New York, NY, USA, 2004. ACM.

[16] Olga Pollatos, Rainer Kopietz, Jennifer Linn, Jessica Albrecht, Vehbi Sakar, Andrea Anzinger, Rainer Schandry, and Martin Wiesmann. Emotional Stimulation Alters Olfactory Sensitivity and Odor Judgment. *Chem. Senses*, 32(6):583–589, 2007.

[17] Alan M. Christie. Simulation: An Enabling Technology in Software Engineering. In *The Journal of Defense Software Engineering*, page 6. Software Engineering Institute, Carnegie Mellon University, 1999.

[18] Jeffrey Adler, Michael McNally, and Will Recker. Interactive Simulation for Modeling Dynamic Driver Behavior in Response to ATIS. TR UCI-ITS-TS-WP-93-4, Institute for Transportation Studies, UCI, 1993.

[19] J. Baujon *et al.* A new low-cost tool for driver behavior analysis and studies. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 569–574, 2000.

[20] Ying Wang, Wei Zhang, Su Wu, and Yang Guo. Simulators for Driving Safety Study – A Literature Review. *Virtual Reality*, pages 584–593, 2007.

[21] Yi-hu Wu, Dan Yu, and Xiaomei Yin. A simulation of car-following driving based on coordinated model. In *ICICTA '08: Proceedings of the 2008 International Conference on Intelligent Computation Technology and Automation*, pages 219–224, Washington, DC, USA, 2008. IEEE Computer Society.

[22] S.F. Zhao, G.H. Xu, and T.F. Tao. Modeling human vehicle driving which integrate models of fatigue. In *4th IEEE Conference on Industrial Electronics and Applications (CIEA 2009)*, pages 3493–3497, 2009.

[23] T. Von Jan, T. Karnahl, K. Seifert, J. Hilgenstock, and R. Zobel. Dont sleep and drive – VW's fatigue detection technology. In *Proceedings of 19th International Conference on Enhanced Safety of Vehicles, Washington, DC*, 2005.

[24] Aladino Amantini and Pietro Cacciabue. A simple simulation predicting driver behavior, attitudes and errors. *Digital Human Modeling*, pages 345–354, 2009.

[25] Jiangpeng Dai, Jin Teng, Xiaole Bai, Zhaohui Shen, and Dong Xuan. Mobile phone based drunk driving detection. In *4th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pages 1–8, March 22–25, 2010. ISBN: 978-963-9799-89-9.

[26] M.E. Rakauskas, N.J. Ward, E.R. Boer, E.M. Bernat, M. Cadwallader, and C.J. Patrick. Combined effects of alcohol and distraction on driving performance. *Accident Analysis & Prevention*, 40(5):1742–1749, 2008.

[27] Y. Wang and W. Zhang. A Driver Training Platform Prototype Based on Distributed Simulation. In *International Conference on Cyberworlds 2008*, pages 506–510. IEEE, 2008.

[28] L.P. Lonero. Trends in driver education and training. *American journal of preventive medicine*, 35(3S):316–323, 2008.

[29] J.M. Leitão, A. Moreira, J.A. Santos, A.A. Sousa, and F.N. Ferreira. Evaluation of driving education methods in a driving simulator. *Computer Graphics and Visualization Education'99*, 1999.

[30] Chao Sun, Feng Xie, Xiaocao Feng, Mingmin Zhang, and Zhigeng Pan. A Training Oriented Driving Simulator. *Entertainment Computing (ICEC 2007)*, pages 1–9, 2007.

[31] Andreas Riener and Alois Ferscha. Simulation Driven Experiment Control in Driver Assistance Assessment. In David Roberts et al., editor, *12th IEEE International Symposium on Distributed Simulation and Real Time Applications, Vancouver, BC, Canada*, page 10. IEEE CS Press, P3425, October 27–29 2008. ISBN: 978-0-7695-3425-1.

[32] Paul F. Smith, Jie Chen, and Hongxing Hu. Model-Based Design Study and Evaluation of New HMI Concepts for Vehicle Multimedia, Climate Control and Navigation Systems. *SAE World Congress & Exhibition, April 2007, Detroit, MI, US*, page 9, 2007.

[33] Wan Jian, Chu Xiumin, Wu Yong, and Zhang Rui. The Design of Autonomous Smart Car Used in Simulation of Vehicle Platoon. In *Computational Intelligence and Industrial Application, 2008. PACIIA '08. Pacific-Asia Workshop on*, volume 1, pages 885 –890, 19-20 2008.

[34] P. Cheevarunothai, Y. Wang, and N.L. Nihan. Development of a loop detector simulator (LOOPSIM) for in-laboratory traffic research and education. *Computer Applications in Engineering Education*, 16(1):45–54, 2008.

[35] J. Santos *et al.* The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F: Psychology and Behaviour*, 8(2):135–146, 2005.

[36] F. Panerai *et al.* Speed and safety distance control in truck driving: comparison of simulation and real-world environment. In *Proceedings of Driving Simulation Conference*, pages 91–107, 2001.

[37] C. Lange, M. Wohlfarter, and H. Bubb. Vergleichbarkeit von Usability Lab und Realversuch zur Bestimmung der ergonomischen Guete und der Ablenkungswirkung von Nebenaufgaben im Kfz. *Proceedings GfA Gesellschaft fuer Arbeitswissenschaften Kongress, Stuttgart*, 2006.

[38] T.Y. Koo, K.J. Park, B.Y. Kim, H.J. Kim, and M.W. Suh. A study on drivers workload of telematics using a driving simulator: A comparison among information modalities. *International Journal of Precision Engineering and Manufacturing*, 10(3):59–63, 2009.

[39] J. Sodnik, C. Dicke, S. Tomazic, and M. Billinghurst. A user study of auditory versus visual interfaces for use while driving. *International journal of human-computer studies*, 66(5):318–332, 2008.

[40] G. Costagliola, S. Martino, and F. Ferrucci. A Simulation Environment to Assess Driving Performances while Interacting with On-board Telematics Systems. *Enterprise Information Systems*, pages 439–451, 2008.

[41] T.Y. Koo, C.H. Bae, B.Y. Kim, Z. Rowland, and M.W. Suh. Development of a driving simulator for telematics human–machine interface studies. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 222(11):2077–2086, 2008.

[42] D.D. Salvucci. Rapid prototyping and evaluation of in-vehicle interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16(2):1–33, 2009.

[43] J. Barceló, J.L. Ferrer, and R. Martin. Simulation assisted design and assessment of vehicle guidance systems. *International Transactions in Operational Research*, 6(1):123–143, 1999.

[44] H.J. Bullinger and M. Dangelmaier. Virtual prototyping and testing of in-vehicle interfaces. *Ergonomics*, 46(1):41–51, 2003.

[45] P. Spethmann, C. Herstatt, and S.H. Thomke. Crash simulation evolution and its impact on R&D in the automotive applications. *International Journal of Product Development*, 8(3):291–305, 2009.

[46] H. Burg, A. Moser, and H. Steffan. *Handbuch Verkehrsunfallrekonstruktion*, chapter Simulation und Animation, pages 585–592. Springer, 2009.

[47] R. Thomson, M. Edwards, T. Martin, C. van der Zweep, R. Damm, and G. della Valle. Car-car crash compatibility: development of crash test procedures in the VC-Compat project. *International Journal of Crashworthiness*, 12(2):137–151, 2007.

[48] E. Guilmineau and F. Chometon. Effect of side wind on a simplified car model: Experimental and numerical analysis. *Journal of Fluids Engineering*, 131:021104, 2009.

[49] E. Mercker and H.W. Knape. Ground simulation with moving belt and tangential blowing for full-scale automotive testing in a wind tunnel. In *Society of Automotive Engineers international congress and exposition*, volume 27, 1989.

[50] W. Eckert, N. Singer, and J.D. Vagt. The Porsche Wind Tunnel Floor-Boundary-Layer Control–A Comparison With Road Data and Results From Moving Belt. 1992.

[51] K. Burgin, P.C. Adey, and J.P. Beatham. Wind tunnel tests on road vehicle models using a moving belt simulation of ground effect. *Journal of Wind Engineering and Industrial Aerodynamics*, 22(2-3):227–236, 1986.

[52] H. Kwon, Y.W. Park, D. Lee, and M.S. Kim. Wind tunnel experiments on Korean high-speed trains using various ground simulation techniques. *Journal of Wind Engineering and Industrial Aerodynamics*, 89(13):1179–1195, 2001.

[53] Kristie L. Young, Michael A. Regan, and John D. Lee. *Driver Distraction: Theory, Effects, and Mitigation*, chapter Measuring the Effects of Driver Distraction: Direct Driving Performance Methods and Measures, pages 85–105. CRC Press, 2009. ISBN: 978-0-8493-7426-5.

[54] M. Pantic, A. Pentland, A. Nijholt, and T. Huang. Human computing and machine understanding of human behavior: a survey. *Artifical Intelligence for Human Computing*, pages 47–71, 2007.

[55] B.G. Silverman, M. Johns, J. Cornwell, and K. O'Brien. Human behavior models for agents in simulators and games: part I: enabling science with PMFserv. *Presence: Teleoperators & Virtual Environments*, 15(2):139–162, 2006.

[56] Giuseppe Andreoni, Marco Rabuffetti, and Antonio Pedotti. Simulation of Complex Human Movement Through the Modulation of Observed Motor Tasks. *Digital Human Modeling*, pages 3–12, 2007.

[57] F. Aparicio, F. Jiménez, and J. Sánchez. Development and use of vehicle dynamics simulation software as support for road vehicles theory teaching. *Computer Applications in Engineering Education*, 17(4):467–478, 2009.

[58] Dieter Ammon, Michael Gipser, Jochen Rauh, and Jrgen Wimmer. High performance system dynamics simulation of the entire system tire-suspension-steering-vehicle. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility*, 27(5):435–455, 1997.

[59] F. Colombet, M. Dagdelen, G. Reymond, C. Pere, F. Merienne, and A. Kemeny. Motion Cueing: what's the impact on the driver's behaviour. In *In Proceeding of the Driving Simulator Conference (DSC 2008)*, pages 171–181, 2008.

[60] G. Reymond and A. Kemeny. Motion cueing in the Renault driving simulator. *Vehicle System Dynamics*, 34(4):249–259, 2000.

[61] A. R. Valente Pais, M. Wentink, M. M. van Paassen, and M. Mulder. Comparison of three motion cueing algorithms for curve driving in an urban environment. *Presence: Teleoper. Virtual Environ.*, 18(3):200–221, 2009.

[62] J. Engström, E. Johansson, and J. Östlund. Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2):97–120, 2005.

[63] Albert J. Rehmann. A Handbook of Flight Simulation Fidelity Requirements for Human Factors Research. Technical Report DOT/FAA/CT-TN95/46, U.S. Department of Transportation Federal Aviation Administration, December 1995.

[64] Matthew P. Reed and Paul A. Green. Validation of a Low-Cost Driving Simulator Using a Telephone Dialing Task. Final Report UMTRI-95- 19, University of Michigan, Transportation Research Institute, June 1995.

[65] Nico Kaptein, Jan Theeuwes, and Richard Van Der Horst. Driving simulator validity: Some considerations. *Transportation Research Record: Journal of the Transportation Research Board*, 1550(-1):30–36, January 1996.

[66] Gary Burnett, A. Irune, and A. Mowforth. Driving Simulator Sickness and Validity: How Important Is It to Use Real Car Cabins? *International Conference on Road Safety and Simulation (RSS 2007), Rome, Italy*, page 6, November 7–9 2007.

[67] B. Reimer, L.A. D'Ambrosio, J.F. Coughlin, M.E. Kafrissen, and J. Biederman. Using self-reported data to assess the validity of driving simulation data. *Behavior research methods*, 38(2):314, 2006.

[68] H. Ashton, R.D. Savage, J.W. Thompson, and D.W. Watson. A method for measuring human behavioural and physiological responses at different stress levels in a driving simulator. *British Journal of Pharmacology*, 45(3):532, 1972.

[69] H.C. Lee. The validity of driving simulator to measure on-road driving performance of older drivers. *Transport Engineering in Australia*, 8:89–100, 2003.

[70] Hurshkumar G. Donde, Rakesh K. Lad, Prabal Kumar Biswas, Praveen B. Patil, and Gordon Stace. Reduction in Time to Market of Automotive Seating System Using LS-DYNA . In *Proceedings of the 7th International LS-DYNA Users Conference, Detroit*, pages 9–17 – 9–30 (14), Electronics City, Bangalroe 561 229, India, July 2002. Infosys Technologies Limited.

[71] Larry Gould. GM Speeds Time to Market Through Blistering Fast Processors. *Automotive Design & Production, Gardner Publications, Inc.*, July 2004. [Online]. last retrieved January 11, 2011. http://www.autofieldguide.com/articles/070407.html.

[72] Alfred T. Lee. *Flight Simulation: Virtual Environments In Aviation*. Ashgate Publishing Limited, Hampshire, England, 2005. ISBN: 0-7546-4287-9.

[73] Federal Aviation Administration. Airplane Simulator Qualification. Advisory circular 120-40b, U.S. Department of Transportation, Washington, D.C., July 1991.

[74] Michael D. Proctor, Maria Bauer, and Thomas Lucario. Helicopter Flight Training through Serious Aviation Gaming. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology (JDMS)*, 5(1):18, January 2008.

[75] Thomas Longridge, Judith Brki-Cohen, Tiauw H. Go, and Andrew J. Kendra. Simulator fidelty considerations for training and evaluation of today's airline pilots. In *Proceedings of the 11th International Symposium on Aviation Psychology*, page 7, March 2001.

[76] Dave Hirschmann. Simulation nation: Flight simulator training for general aviation. *Aircraft Owners and Pilots Association (AOPA) Flight Training*, page 3, January 2009.

[77] Samantha Jamson, Frank Lai, Hamish Jamson, Anthony Horrobin, and Oliver Carsten. Interaction between Speed Choice and the Environment. Road safety research report 100, Department for Transport, Great Minster House, 76 Marsham Street, London SW1P 4DR, 2008. ISBN: 978-1-906581-39-8.

[78] Andrew Kun, Tim Paek, and Zeljko Medinca. The effect of speech interface accuracy on driving performance. In *Proceedings of 8th Annual Conference of the International Speech Communication Association (INTERSPEECH 2007)*, Antwerp, Belgium, August 27–31 2007. Microsoft Research, USA and University of New Hampshire, USA.

[79] Susan Chisholm, Jeff Chaird, Lulie Lockhart, Lisa Fern, and Elise Teteris. Driving Performance while Engaged in MP3 Player Interaction: Effects of Practice and Task Difficulty on PRT and Eye Movements. In *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, pages 238–245 (8), Skamania Lodge, Stevenson, Washington, USA, July 9–12 2007.

[80] N. N. Making a driving simulator even better – Improved imagery and simulation have increased safety and performance at Ford. *automotive electronics*, pages 10–12, August/September 2008.

[81] Hindrik W.J. Robbe. Marijuana use and driving. *Journal of the International Hemp Association*, 1:44–48, 1994.

[82] Race Technology Ltd. Automotive Technical Excellence – Products. [Online]. last retrieved January 9, 2011. http://www.race-technology.com/products_2_3.html.

[83] Reiner Suikat. The new Dynamic Driving Simulator at DLR. In *Driving Simulator Conference, Orlando, Florida*, pages 374–381 (8), November 30 – December 2 2005.

[84] Douglas E. Cross and Harold L. Luper. Relation between finger reaction time and voice reaction time in stuttering and nonstuttering children and adults. *J Speech Hear Res*, 26(3):356–361, September 1983.

[85] Andreas Riener. Age- and Gender-Related Studies on Senses of Perception for Human-Vehicle-Interaction. In *Proceedings of the 2nd Workshop on Automotive User Interfaces and Interactive Applications, September 7-10, 2008, Lübeck, Germany*, page 8. GI-Edition 2008, September 2008.

[86] J.B.F. Van Erp and H. Van Veen. Vibro-tactile information presentation in automobiles. In *Proceedings of Eurohaptics*, volume 2001, pages 99–104, 2001.

[87] Juergen Schwarz et al. Code of Practice for the Design and Evaluation of ADAS. PReVENT Report 11300, v1.6, Response 3, a PReVENT Project, August 2006. Preventive and Active Safety Applications, Integrated Project, Contract number FP6-507075.

[88] Marcus Toennis, Verena Broy, and Gudrun Klinker. A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers. In *Proceedings of the 3D User Interfaces (3DUI'06)*, pages 127–134, Washington, DC, USA, 2006. IEEE Computer Society.

[89] Netherlands Organisation for Applied Scientific Research (TNO). Automotive. [Online]. last retrieved January 9, 2011. http://www.tno.nl/groep.cfm?context=markten&content=markt&laag1=59&item_id=59.

[90] German Aerospace Center (DLR). Institute of Transportation Systems – Dynamic Driving Simulator. [Online]. last retrieved January 9, 2011. http://www.dlr.de/fs/en/desktopdefault.aspx/tabid-1236/1690_read-3257.

[91] TRW Automotive. Cognitive Safety Systems. [Online]. last retrieved January 9, 2011. http://www.trwauto.com.

[92] Center for Traffic Sciences (IZVW). Driving simulator with motion platform. [Online]. last retrieved January 9, 2011. http://www.psychologie.uni-wuerzburg.de/methoden/forschung/technik/fahrsimulator.php.en.

[93] T. Bock, M. Maurer, and G. Farber. Validation of the Vehicle in the Loop (VIL) – A milestone for the simulation of driver assistance systems. In *2007 IEEE Intelligent Vehicles Symposium*, pages 612–617, 2007.

[94] Yaman Barlas. Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, 12(3):183–210, 1996.

[95] Gian Beeli, Susan Koeneke, Katja Gasser, and Lutz Jancke. Brain stimulation modulates driving behavior. *Behav Brain Funct.*, 4(34):7, August 6, 2008. [Online].

[96] Monica Bordegoni, Umberto Giraudo, Giandomenico Caruso, and Francesco Ferrise. Ergonomic interactive testing in a mixed-reality environment. In *Proceedings of the 2nd International Conference on Virtual Reality (ICVR 2007), Held as part of HCII'07, Beijing, China*, pages 431–440, July 22–27 2007.

[97] N. Dahlstrom, S. Dekker, R. van Winsen, and J. Nyce. Fidelity and validity of simulator training. *Theoretical Issues in Ergonomics Science*, 10(4):305–314, 2009.

[98] Paul Green. How Driving Simulator Data Quality Can Be Improved. In *Driving Simulation Conference North America 2005, Orlando, Florida*, page 11, November 2005.

[99] Yaron Hollander and Ronghui Liu. The principles of calibrating traffic microsimulation models. *Transportation*, 35(3):347–362, May 2008.

[100] A.H. Hoskins and M. El-Gindy. Technical report: Literature survey on driving simulator validation studies. *International Journal of Heavy Vehicle Systems*, 13(3):241–252, 2006.

[101] Maura Houtenbos. *Expecting the unexpected – A study of interactive driving behaviour at intersections*. Hs-043 950, Technical University Delft, The Netherlands, January 2008. pp. 216.

[102] Jack P. C. Kleijnen. Verification and validation of simulation models. *European Journal of Operational Research*, 82(1):145–162, 1995.

[103] Michael A. Regan, John D. Lee, and Kristie L. Young. *Driver Distraction: Theory, Effects, and Mitigation*, chapter Measuring the Effects of Driver Distraction, pages 85–106. CRC Press, 2009. ISBN: 978-0-8493-7426-5.

[104] Kwon Son, Kyung-Hyun Choi, and Sung-Sook Eom. Virtual prototyping simulation for a passenger vehicle. *Journal of Mechanical Science and Technology*, 15(4):448–458, April 2001.

[105] Tomer Toledo and Haris Koutsopoulos. Statistical validation of traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board*, 1876(-1):142–150, January 2004.

[106] Marcel B.F. Uhr, Daniel Felix, B.J. Williams, and Helmut Krueger. Transfer of Training in an Advanced Driving Simulator: Comparison Between Real World Environment and Simulation in a Manoeuvring Driving Task. In *Driving Simulation Conference, North America 2003 (DSC-NA 2003)*, page 11, 2003.

[107] Ying Wang, Bruce Mehler, Bryan Reimer, Vincent Lammers, Lisa A. D'Ambrosio, and Joseph F. Coughlin. The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics*, 53(3):404–420, 2010.

[108] Garrett Weinberg and Bret Harsham. Developing a low-cost driving simulator for the evaluation of in-vehicle technologies. In *AutomotiveUI '09: Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 51–54, New York, NY, USA, 2009. ACM.