# Assessing the Effectiveness of a Dynamic Driving Simulator for ADAS Development—A Back-to-Back Subjective Validation Study on Lateral Guidance Systems

Sijie Wei, Matthias Becker, Peter E. Pfeffer, and Johannes Edelmann

***Abstract*—Rapid developing driving simulator technologies enable the possibility of assisting the human-centered development of Advanced Driver Assistance Systems and Autonomous Driving. It boosts the efficiency of functional testing and professional subjective assessment of the system under test and thus greatly shortens the development cycles of ADAS functions. Furthermore, standardized and transparent assessment procedures contribute to the robustness and transferability of the test results. However, the application of driving simulators in the practical development process is based upon the validity of the test results in the virtual environment. The aim of this study is to determine the subjective validity of a high-fidelity dynamic driving simulator. A back-to-back study was designed to subjectively evaluate two lateral guidance systems on an objectively validated simulator and in a real vehicle respectively. The results show that the professional drivers evaluate the system characteristics similarly in the aspects of driver interaction, perceived safety and functional performance as well as in most of their subaspects in the virtual and in the physical test environment and that absolute subjective validity can be established. Although the intervention intensity of the lane departure avoidance system and the general reproducibility of the lane keeping assist system show significant differences between the test environments, relative subjective validity can also be confirmed. In addition, the results of regression analysis reveal the influencing factors of driver's subjective evaluation of the three main system characteristics and confirm the effectiveness of the evaluation methods.**

***Index Terms*—Lateral guidance system, advanced driver assistance system, dynamic driving simulator, subjective validation.**

## I. Introduction

WITH the rapid development of Advanced Driver Assistance Systems (ADAS) and Autonomous Driving (AD) functions, and the growing trend towards industrial digitalization, the need to minimize development time, cost and resource consumption is increasing. The importance of lateral guidance systems in particular is highlighted by the EU regulation, which came into force in July 2024, requiring all new vehicles to be equipped with an emergency lane departure warning system [1]. There are two types of lateral guidance systems of Automation Level 2 (according to [2]), namely Type I and Type II according to [3]. The primary goal of the Type I system is to prevent the driver from unintentionally leaving the lane. It intervenes only near the lane markings, leaving a relatively wide area around the center of the lane for the driver to steer freely. Type II, on the other hand, is designed to relieve the driver of driving tasks, such as long-distance highway driving, and keep the vehicle in its lane. Type II systems are an essential foundation for higher-level automation and full autonomy. These two types of lateral guidance systems are referred to in this paper as Lane Departure Avoidance (LDA) and Lane Keeping Assist (LKA) systems, respectively. For a comprehensive review on the topic of subjective and objective evaluation of lateral guidance systems refer to [4]. Compared to the widespread application of LDA and LKA in commercial vehicles, customer acceptance of lateral guidance systems remains limited. A study involving over 120 participants found that drivers experience significantly higher stress levels, both subjectively and objectively, when using LKA [5]. Additionally, a customer survey conducted by AllianzDirect in 2023 reported that only 19.3% of respondents felt safe using lateral guidance systems [6]. The current development process for lateral guidance systems relies heavily on the subjective judgment of professional test drivers and development engineers in the late prototype phase. This testing and tuning process is not transparent due to its subjective nature and dependence on specific development processes and requirements of the Original Equipment Manufacturer (OEM). On the other hand, it is also very time and resource consuming.

Driving simulators have become a powerful tool for research and development, with applications spanning engineering, medicine, psychology, and more. Their role is especially critical in vehicle and ADAS development, meeting today's high demands for efficiency, safety, resource conservation, and the integration of human factors. For a comprehensive review of the open-source driving simulators used in the research of AD, see [7]. Driving simulators provide a

Received 9 December 2024; revised 5 March 2025, 2 June 2025, and 7 July 2025; accepted 5 August 2025. The Associate Editor for this article was J. Fang. *(Corresponding author: Johannes Edelmann.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of MdynamiX AG under the Chairmanship of Prof. Markus Krug.

Sijie Wei and Johannes Edelmann are with the Institute of Mechanics and Mechatronics, TU Wien, 1060 Vienna, Austria (e-mail: e11946216@student.tuwien.ac.at; johannes.edelmann@tuwien.ac.at).

Matthias Becker and Peter E. Pfeffer are with the Department of Automotive Engineering, Munich University of Applied Sciences, 80335 Munich, Germany (e-mail: becker.matthias@hm.edu; peter.pfeffer@hm.edu).

Digital Object Identifier 10.1109/TITS.2025.3598503

safe and controlled testing environment with reproducible driving scenarios, in which test drives can be conducted under controlled conditions and in a shorter time period. An example is [8], where the researchers employed a high-fidelity dynamic driving simulator to analyze driver-preferred autonomous driving styles under varying weather conditions and oncoming traffic. The simulator enabled precise control over environmental variables and traffic scenarios across all participants, facilitating consistent and efficient evaluation. This approach is particularly valuable given the regulatory and safety constraints associated with real-world autonomous vehicle testing. Furthermore, driving simulators are suitable for investigating human-machine-interaction and including human factors such as distractions, mental load, interaction etc. in the technical design and evaluation. A recent review [9] examined the use of driving simulators for detecting abnormal driving behaviors across 70 studies. The authors highlighted the advantage of driving simulators in providing a controlled yet realistic environment that enables the safe exploration of risky or ethically challenging scenarios, as opposed to real-world GPS-based testing methods. Additionally, driving simulators offer a promising alternative to overcome the limitations of current real-vehicle subjective evaluations mentioned above. They can be flexibly integrated into all phases of the V-model [3], particularly in early development stages of ADAS and AD, involving the driver as early as possible. This approach could reduce or even replace the reliance on empirical real-world test drives with professional test drivers during the prototype phase. Thus, it can significantly lower development costs by shortening iteration cycles and minimizing the need for prototypes, professional test drivers, and test tracks.

However, the subjective studies focusing on human-machine-interaction and subjective driver evaluation, such as safety perception, acceptance etc., primarily remain in the research phase and are not yet fully integrated into the actual product development process of ADAS and AD. Findings from subjective studies conducted in driving simulators cannot always be directly transferred to real-world scenarios, as the validity of such studies has not been thoroughly investigated. This lack of validation hinders decision-makers from effectively incorporating subjective feedback into the product development process. Validity, along with objectivity and reliability, is one of the three essential requirements for the design, conduct and evaluation of subject research [10].

Even though driving simulators have been used in a wide range of application areas, validation studies of driving simulators are rather heterogeneous. There are validation studies regarding the individual use cases, but the results cannot be extrapolated. This is because the degree of validity required is often dependent on the purpose of the experiment. Driving simulator validity means how accurately it replicates real-world driving [11] and can be classified from different perspectives. The most common one is the distinction between absolute and relative validity [12]. Absolute validity refers to cases where the data obtained in the simulator matches the data obtained in the real vehicle to a sufficient accuracy. If the absolute value of the data from the two domains are different but have a comparable trend, it can be referred to as relative validity. The data being compared in the two domains could be variables such as driving parameters, psychophysiological measures or subjective evaluations etc. [13].

Other classifications are also introduced and investigated. Mullen et al. [14] distinguish physical from behavioral validity. Physical validity focuses on the characteristics of the driving simulator, while behavioral validity refers to the correspondence of the driving behavior in the two domains. For example, Klüver et al. [13] studied how the simulator characteristics and user characteristics influence the behavioral validity of the simulators. They found a significant association between simulator sickness and impaired driving performance, with gender and age also influencing performance in fixed-base simulators. In addition, some researchers have taken driver's mental workload across different environments into consideration. Lobjois et al. [15] investigated the behavioral validity of a low-cost driving simulator by comparing speed control and cognitive demand. While speed-related measures showed high validity, attentional aspects related to vehicle handling were less consistent. The joint analysis of behavioral and mental workload metrics revealed discrepancies between these two validity dimensions. In [16], the author divides validity into external and internal validity. External validity refers to the representativeness and transferability of an experiment and its results to real-life conditions. Internal validity, on the other hand, is a question of whether a found effect is really unambiguously due to the influence of the independent variable [16].

Moreover, objective and subjective validity are also mentioned in literature. Objective validity, which by definition is equivalent to physical validity, can be established by comparing selected objective measures, and subjective validity by considering drivers' perceptions, obtained through survey [17]. Physical or objective validation of a driving simulator typically addresses the system's components, layout, and dynamic characteristics, such as braking and steering response [12], [14], and is generally considered a prerequisite for establishing subjective validity. While validated track and vehicle models are mandatory requirements, they do not automatically ensure subjective validity. Factors such as actuator behavior, ergonomic design, screen dimensions, and motion cueing algorithms can all significantly influence the perceived realism and, thus, the overall subjective validity of the simulator [18]. Llopis-Castelló et al. [17] conducted an objective validation of a low-cost driving simulator based on continuous speed profiles, complemented by a subjective evaluation of the virtual environment quality and task similarity between the simulator and the real world through questionnaires. In addition to external validity, Hussain et al. [19] investigated the subjective validity of a static driving simulator setting utilizing a questionnaire comparing the general comfort, speed perception and performance of the simulator with the real vehicle. Li et al. [20] evaluated subjective validity in straight and large-radius cornering sections using both physiological measures and questionnaires addressing perceived similarity to real-world driving in various aspects, including steering wheel, clutch, gear, scenarios, and speed perception. Research [21] also assessed subjective validity for speed research by asking participants to rate the perceived realism of the simulation. Andriola et al. [22] similarly examined perceived realism and motion sickness in an immersive simulator compared to the real-world scenario. These studies have primarily evaluated subjective validity based on participants' perceptions of similarity between the simulator and real-world driving. However, the definition of "subjectively valid" remains vague, often interpreted as being achieved when the majority of participants rate the realism above a medium threshold. More comprehensive studies have extended this approach by integrating both descriptive and statistical analyses of subjective evaluations collected in both simulated and real environments, such as [23], [24], and [25]. Gómez et al. [23] studied the objective and subjective validity of a dynamic driving simulator regarding steering feel and handling by comparing the objective metrics and the subjective assessment between the driving simulator and the real road test respectively. Due to the limited number of professional test drivers – only two completed the simulator trials – the analysis focused on comparing directional trends in subjective ratings across different vehicle setups in different test environments. In contrast, studies [24] and [25] included larger sample sizes ($N > 70$), allowing for statistical analysis of subjective ratings using methods such as ANOVA. Further aspects of simulator validity from the point of view of psychology and behavioral sciences can be found in [26].

When it comes to developing ADAS and AD and analyzing their interaction with drivers, driving simulators come in especially handy. They enable the testing of features that are not yet ready for production or that would not yet meet regulatory requirements, in user-orientated or extreme scenarios. However, in comparison to the number of simulator studies on ADAS and AD, the number of studies that have confirmed the validity of the results is limited. Some examples are [24], which investigated the driving simulator validity for evaluating driving comfort of autonomous driving, and [25], which conducted a validation study of an autonomous emergency braking system using a static driving simulator. Not only is there a paucity of research, but there is also a lack of standardized methods for assessing simulator validity, as mentioned in [11]. Furthermore, behavioral and subjective validity of a driving simulator is always limited to a specifically defined research question or driving task [24], [27]. Therefore, this study is dedicated to investigating the potential of a high-fidelity dynamic driving simulator in terms of its validity for the human-centered development and tuning of lateral guidance systems. The study offers three key benefits: First, it identifies weaknesses in current lateral guidance systems by benchmarking commercial LDA and LKA systems, which, as noted, do not fully meet customer expectations. Second, if subjective validity is established alongside objective validity, it will prove that this simulator can be utilized not only in the early development phase, but also consistently throughout the product development process. In particular, it can add value in the late tuning phase by potentially reducing the dominance of the road test with test drivers. Finally, the findings provide insights for the development of other ADAS requiring similar levels of human monitoring and interaction.

Fig. 1. The Advanced Vehicle Driving Simulator (aVDS) at IFM of Kempten University of Applied Sciences.

The objective of this study is to examine the absolute validity of professional test drivers' subjective perception and evaluation of lateral guidance systems in a high-fidelity dynamic driving simulator, in order to investigate the potential and limitations of developing such systems in a virtual environment. This study aims to answer the research question: Do professional test drivers perceive and evaluate LDA and LKA system similarly in a dynamic diving simulator compared to a real-world environment? Furthermore, this study discusses which sub-aspects influence professional test drivers' overall impression of LDA and LKA in terms of interaction, safety and performance.

The remainder of this paper is organized as follows: Section II presents the experiment methods and study design, as well as how the collected data are processed. Section III demonstrates the detailed results of the validation study, which include the comparison of the main subjective assessment criteria, the sub-criteria and the transferibility of the assessment of LDA and LKA in the two domains respectively. In addition, the regression models of the subjective assessments are also analyzed. The results are then discussed in Section IV, based on which key conclusions are drawn and the research question is answered in Section V.

## II. Methods and Data

### *A. Dynamic Driving Simulator and Lateral Guidance Functions*

The study is conducted on a high-fidelity dynamic driving simulator from AB Dynamics at the Institute for Driver Assistance and Connected Mobility (IFM) associated with the University of Applied Sciences Kempten, Germany (see Fig 1). The high performance motion platform with 8 linear actuators provides 6 degree of freedom (DOF) motion. 7 LED projectors provide visualization of the driving scenario with a diameter of 8.8 m, a height of 4 m and a field of view of 250° for the driver. A half-cockpit is installed on the motion platform to increase driver immersion. A steering rig with a linear actuator that applies simulated rack forces allows the installation of the original steering system, including the electrical control unit [28], [29]. A vehicle model is parameterized and validated on the simulator based on the test vehicle on which the road test is performed. The LDA and LKA algorithms, identical to those from the actual vehicle ECU, are provided by the

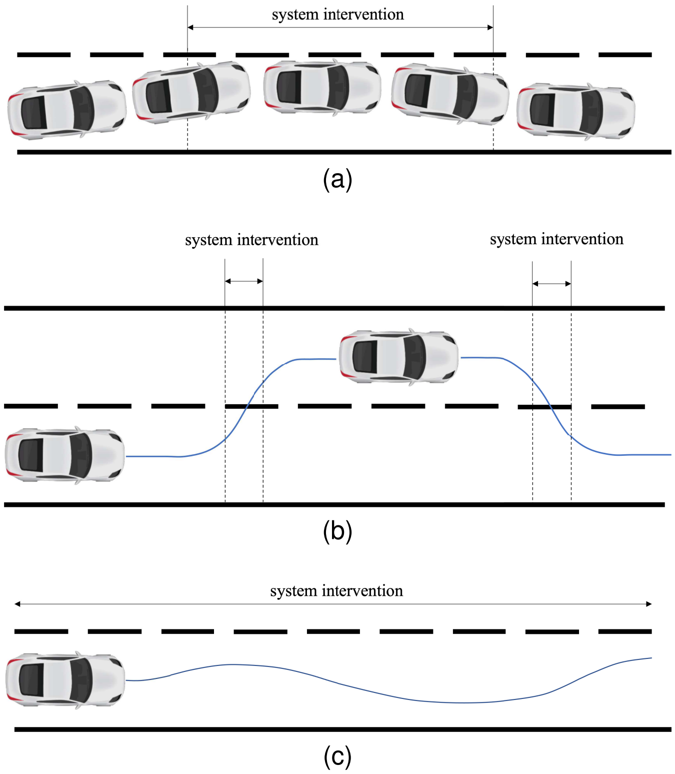

Fig. 2. Test scenarios. (a) Drifting off (DO) on a straight section. (b) Lane change (LC) on a straight section. (c) Free run (FR) on a straight section.

industrial project partner as C-code and integrated into the validated vehicle model in IPG CarMaker. This Software-in-the-Loop (SiL) simulation is executed on a real-time PC running RedHawk Linux. The LDA and LKA functions are objectively validated prior to the subjective study.

### B. Study Design

For the validation of the LDA function, the maneuver of Drifting Off (DO) and Lane Change (LC) were performed on a straight road at the speed of 90 km/h. The LKA function was validated by hands-off Free Run (FR) in the middle of the lane on a straight section at 100 km/h. The test scenarios are depicted in Fig 2. The same scenarios with the same initial conditions were simulated both in offline simulation and on the simulator. The objective validation was accomplished by comparing the characteristic objective metrics (OM) such as intervention time and position, maximal lane deviation, lateral acceleration, lateral jerk, and intervention torque etc.

22 male professional drivers (median age: 33 years, mean age: 36.9 years, standard deviation: 9.4 years) were recruited for the back-to-back subjective validation study, with informed consent obtained. Of these, 20 drivers had experience with more than 5 test drives, while the remaining 2 had participated in 2–5 test drives prior to the study. The subjects were asked to test the LDA and LKA function of the Vehicle under Test (VUT) on the proving ground and then on the dynamic driving simulator in Fig 1. They were asked to conduct the maneuvers described in Fig 2, with both hands on and hands off, 5 times in the case of each system respectively. After 5 repetitions, the drivers rated the System under Test (SUT) based on their subjective perceptions. Then they repeated the same process on the dynamic driving simulator.

This study has inherited the questionnaire from study [30], which was designed to evaluate LDA on a static driving simulator, with minor modifications. The subjects were asked to evaluate the LDA and LKA system in three major subjective aspects, i.e. *Driver Interaction*, *Perceived Safety* and *Functional Performance*. Two additional criteria, *Lateral Jerk* and *Handover Behavior on Intervention Termination*, have been incorporated into the original questionnaire in [30] for the evaluation of the systems' dynamic characteristics and the smoothness of the handover process. While the main focus of the subjective evaluation of both systems is similar, there are some minor differences in the sub-criteria of the two systems to account for their different use cases. For instance, due to the short intervention duration of LDA, the primary focus is on the overall intervention intensity. In contrast, less emphasis is placed on the driver's interaction with the steering wheel during intervention, such as *Co-* and *Counter-Steering* or *Steering Wheel Motion*. This limited timeframe does not allow drivers sufficient opportunity to adequately assess these factors. As for LKA, the system employs a constant intervention strategy on the straight sections to maintain the vehicle in the center of the lane. Consequently, criteria such as *Disengagement Position* and *Angle* cannot be evaluated during continuous intervention and are excluded. The sub-criteria of each category are listed in Table I. For detailed definition of each criterion, refer to [30]. Subjects rated each criterion on a 1-10 ATZ-scale with the option of rating the direction of the deficiency. For example, for LDA-SA1.1 *General Intervention Strength*, 1 is too weak, 10 is optimal, and -1 is too strong.

After all test runs were completed, the subjects answered summary questions about their perception of the two SUT on the simulator compared to the proving ground: i) Have you recognized the functional characteristics of the LDA/LKA system on the dynamic simulator? ii) Is it conceivable to transfer the LDA/LKA development procedure to the dynamic driving simulator? The subjects answered these questions according to their experience in the two test environments on a scale of 1-5, where 1 represents "not recognizable/conceivable at all" and 5 represents "very well recognizable/conceivable".

### C. Data Analysis

The subjective assessment (SA) scores, which ranged from 1-10 with the introduced sign representing the direction of deficiency mentioned above, were first converted to a linear scale of 1-19 to deal with the negative sign [30]. In order to find out whether the subjects evaluate the SUT on the dynamic simulator in a similar way as in the real vehicle, the SAs were statistically tested. Since the SAs are not normally distributed, Wilcoxon signed rank test for paired samples was selected. The null hypothesis is that there is no difference between the subjects' SAs of the LDA/LKA system on the dynamic driving simulator and in the vehicle. A significance level of 0.1 [30] was selected for the test. After testing the hypothesis, the effect size $r$ was calculated to evaluate the practical relevance of the

TABLE I
MAIN AND SUB-CRITERIA FOR THE SUBJECTIVE ASSESSMENT OF LDA AND LKA SYSTEM RESPECTIVELY

| Lane Departure Avoidance | | | |
|---|---|---|---|
| # | Main Criteria | # | Sub-Criteria |
| SA1 | Driver Interaction | LDA-SA1.1 | General Intervention Strength |
| | | LDA-SA1.2 | Override Capability at Lane Boundary |
| SA2 | Perceived Safety | LDA-SA2.1 | Maximum Lane Overshoot |
| | | LDA-SA2.2 | General Reproducibility |
| SA3 | Functional Performance | LDA-SA3.1 | Control-Free Corridor |
| | | LDA-SA3.2 | Returning Behavior |
| | | LDA-SA3.3 | Disengagement Position |
| | | LDA-SA3.4 | Disengagement Angle |
| | | LDA-SA3.5 | Lateral Jerk |
| | | LDA-SA3.6 | Handover Behavior on Intervention Termination |
| Lane Keeping Assist | | | |
| # | Main Criteria | # | Sub-Criteria |
| SA1 | Driver Interaction | LKA-SA1.1 | Counter-Steering |
| | | LKA-SA1.2 | Co-Steering |
| | | LKA-SA1.3 | General Intervention Strength |
| | | LKA-SA1.4 | Steering Wheel Torque Build-up |
| | | LKA-SA1.5 | Steering Wheel Motion |
| SA2 | Perceived Safety | LKA-SA2.1 | Maximum Lane Deviation |
| | | LKA-SA2.2 | General Reproducibility |
| SA3 | Functional Performance | LKA-SA3.1 | Control-Free Corridor |
| | | LKA-SA3.2 | Oscillation Behavior |
| | | LKA-SA3.3 | Lateral Jerk |
| | | LKA-SA3.4 | Handover Behavior on Intervention Termination |

results. It is calculated with the $z$-score and the sample size $n$ as follows:

$$r = |\frac{z}{\sqrt{n}}|. \quad (1)$$

The value of $r$ ranges from 0 to 1, with 0 representing no effect and 1 maximum effect. According to [31], the effect size is small if the absolute value of $r$ varies around 0.1, medium if around 0.3, and large if more than 0.5.

If the results of the Wilcoxon test are not significant, but have a medium to large effect size $r$, it was then examined whether conspicuously high variances of the SA ratings may have diluted the difference between the vehicle and the simulator. The variance of the SA ratings in the two domains was examined using the coefficient of variation ($CV$) and then a Levene's test was performed to see if the variances are significantly different. The $CV$ is defined as the ratio of the standard deviation $\sigma$ to the mean $\mu$:

$$CV = \frac{\sigma}{\mu}. \quad (2)$$

In addition, it was also investigated which sub-criteria influence the rating of the main criteria. The collinearity of the sub-criteria ratings were tested first. After that, single and multiple linear regression models were established to examine the relationship of the sub-criteria and main criteria of each evaluation category. The regression equation is generalized as

$$y = b_0 + \boldsymbol{b}' \cdot \boldsymbol{x} + u, \quad (3)$$

where $y$ represents the score of the main criterion, $\boldsymbol{x}$ the sub-criterion or several sub-criteria. $\boldsymbol{b}$ represents the regression coefficients and $u$ the stochastic unobservable disturbance.

## III. VALIDATION RESULTS

### A. Results of Subjective Validation

The distribution of the SAs for the vehicle and the simulator is shown in Fig 3. They are not normally distributed and right-skewed. Wilcoxon test was conducted to examine the difference between all the SAs from the two environments. The result with a $p$-value of 0.1687 failed to reject the null hypothesis, meaning that the subjects generally evaluate the SUT in a similar way in the vehicle and on the dynamic driving simulator. The effect size $r = 0.0573$ indicates a very weak effect of the impact of the driving test environment. In the following, SA criteria are analyzed individually for each SUT.

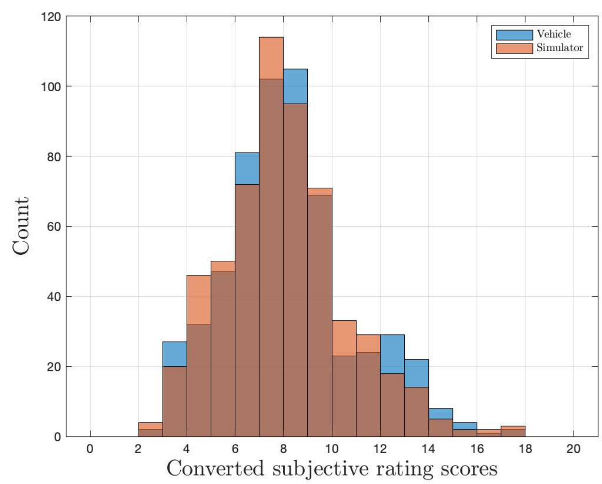


Fig. 3. Distribution of SAs for the vehicle and the dynamic driving simulator.

*1) Main Subjective Criteria:* The main SAs of both LDA and LKA are comparable in the road test and on the dynamic simulator. The results of the Wilcoxon tests show no significant difference between the two domains. This means that the drivers do not perceive the interaction, safety and functional performance differently in the simulator as in the real vehicle. The comparison of the the SAs is shown in Fig 4 and the results of the Wilcoxon test for each main criterion as well as the effect size $r$ are listed in Table II. Medium to large effect

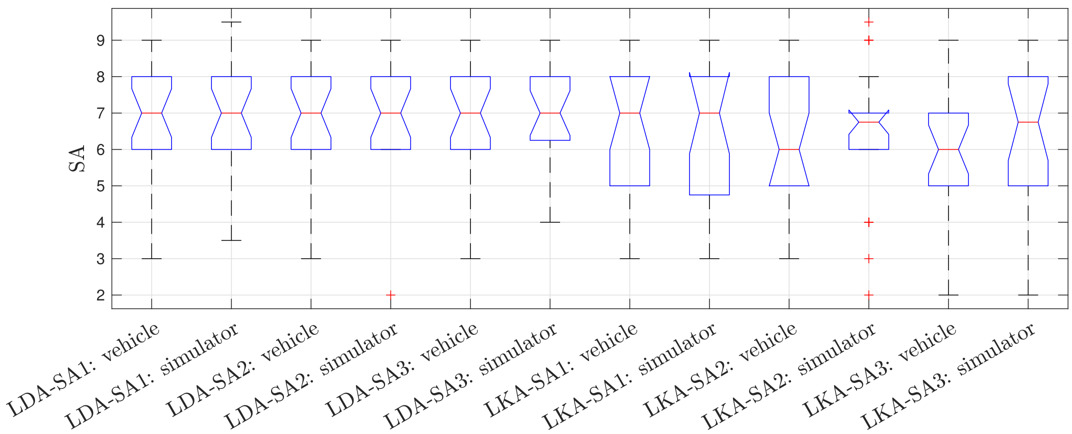


Fig. 4. Comparison of drivers' rating scores of 3 main SAs for LDA and LKA.

TABLE II
WILCOXON TEST RESULTS AND EFFECT SIZE OF LDA AND LKA MAIN SAs

| # | Main Criteria | $p$-Value | Effect Size $r$ |
|---|---|---|---|
| LDA-SA1 | Driver Interaction | 0.18 | 0.07 |
| LDA-SA2 | Perceived Safety | 0.70 | 0.08 |
| LDA-SA3 | Functional Performance | 0.31 | **0.23** |
| LKA-SA1 | Driver Interaction | 0.54 | 0.13 |
| LKA-SA2 | Perceived Safety | 0.79 | 0.06 |
| LKA-SA3 | Functional Performance | 0.16 | **0.32** |

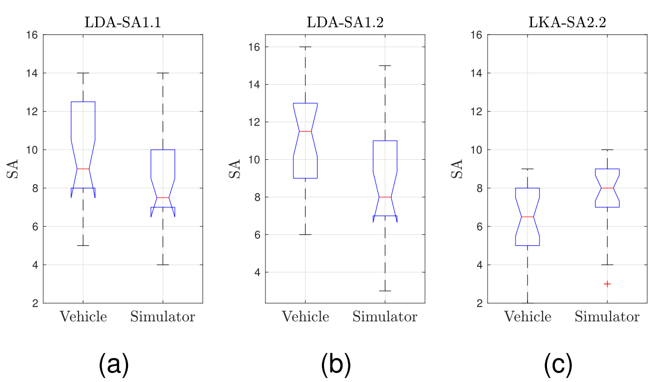


Fig. 5. Box plot of subjective ratings in vehicle vs in simulator. (a) LDA-SA1.1 General Intervention Strength. (b) LDA-SA1.2 Override Capability at Lane Boundary. (c) LKA-SA2.2 General Reproducibility.

sizes $r$ are in bold. Even though not statistically significant, SA3 *Functional Performance* of both LDA and LKA show a medium effect of $r = 0.23$ and $r = 0.32$ between the two domains. This can still show practical relevance and should be taken into consideration in further studies.

*2) Lane Departure Avoidance System:* The statistical test results of the drivers' ratings of each sub-criterion for LDA system are shown in Table III. In this and the following tables, $p$-values smaller than 0.1 and 0.01 are marked with * and ** respectively. It can be seen that LDA-SA1.1 *General Intervention Strength* and LDA-SA1.2 *Override Capability at Lane Boundary* are evaluated significantly differently in the vehicle compared with the simulator. The box plots of the comparison between the subjects' ratings in vehicle and in the simulator for each sub-criterion are displayed in Fig 5a and 5b. It shows that drivers perceived a significantly weaker LDA intervention in the simulator and that it was subjectively easier to override the system in the simulator compared to the real vehicle.

The rest of the sub-criteria do not show significant differences between the two test environments, which means the subjects perceive these system characteristics similarly in vehicle and in the dynamic simulator. However, even though the null hypothesis cannot be rejected, the comparison of LDA-SA2.1 *Maximum Lane Overshoot* and LDA-SA3.5 *Lateral Jerk* in the two test environments also have a medium effect size $r$.

*3) Lane Keeping Assist System:* The results of the Wilcoxon signed rank test as well as effect size $r$ are shown in Table IV. The only sub-criterion that shows significant difference between real vehicle and the dynamic simulator is LKA-SA2.2 *General Reproducibility*. The comparison of the SAs in vehicle and in the dynamic simulator is shown in Fig 5c. It indicates that the subjects experienced higher reproducibility with less variation in the simulator. The comments of the drivers are analyzed. 7 drivers commented on the real system in the vehicle and 1 on the simulator for this criterion. 6 comments mention that the LKA system in the vehicle "intervenes differently on the left and on the right side of the lane" and that it is "asymmetric". 1 comment on the simulator implies that the system intervention is "more symmetric".

For LKA-SA1.3 *General Intervention Strength*, LKA-SA3.2 *Oscillation Behavior* and LKA-SA3.3 *Lateral Jerk* the effect size is medium. There are no comments from the subjects on LKA-SA1.3 or LKA-SA3.3. For LKA-SA3.2, however, 4 out of 4 comments on the real system in the vehicle mention that the LKA system "oscillates differently" dependent on the side of the lane. 2 out 5 comments on the simulator mention that the LKA system oscillates "less" and "more symmetric". For all other sub-criteria, the subjective ratings are comparable between the two test environments.

*4) Relative Validity:* For subjective ratings that do not show a significant difference between the vehicle and the simulator, but the effect sizes are still moderate, their variances are further examined to see if the variability of the data has masked the

TABLE III

WILCOXON SIGNED RANK TEST RESULTS AND EFFECT SIZE OF SA SUB-CRITERIA FOR LDA SYSTEM ($p^* < 0.1$, $p^{**} < 0.01$)

| # | Main Criteria | # | Sub-Criteria | $p$-Value | Effect Size $r$ |
|---|---|---|---|---|---|
| LDA-SA1 | Driver Interaction | LDA-SA1.1 | General Intervention Strength | 0.0762* | **0.3781** |
| | | LDA-SA1.2 | Override Capability at Lane Boundary | 0.0025** | **0.6442** |
| LDA-SA2 | Perceived Safety | LDA-SA2.1 | Maximum Lane Overshoot | 0.7136 | **0.2861** |
| | | LDA-SA2.2 | General Reproducibility | 0.4781 | 0.1512 |
| LDA-SA3 | Functional Performance | LDA-SA3.1 | Control-Free Corridor | 0.9434 | 0.0151 |
| | | LDA-SA3.2 | Returning Behavior | 0.5838 | 0.1168 |
| | | LDA-SA3.3 | Disengagement Position | 0.9496 | 0.0135 |
| | | LDA-SA3.4 | Disengagement Angle | 0.4490 | 0.1614 |
| | | LDA-SA3.5 | Lateral Jerk | 0.2315 | **0.2551** |
| | | LDA-SA3.6 | Handover Behavior on Intervention Termination | 0.4010 | 0.1790 |

TABLE IV

WILCOXON SIGNED RANK TEST RESULTS AND EFFECT SIZE OF SA SUB-CRITERIA FOR LKA SYSTEM ($p^* < 0.1$, $p^{**} < 0.01$)

| # | Main Criteria | # | Sub-Criteria | $p$-Value | Effect Size $r$ |
|---|---|---|---|---|---|
| LKA-SA1 | Driver Interaction | LKA-SA1.1 | Counter-Steering | 0.4130 | 0.1786 |
| | | LKA-SA1.2 | Co-Steering | 0.7970 | 0.0561 |
| | | LKA-SA1.3 | General Intervention Strength | 0.2077 | **0.2749** |
| | | LKA-SA1.4 | Steering Wheel Torque Build-up | 0.2334 | **0.2600** |
| | | LKA-SA1.5 | Steering Wheel Motion | 0.8923 | 0.0295 |
| LKA-SA2 | Perceived Safety | LKA-SA2.1 | Maximum Lane Deviation | 0.3568 | 0.1965 |
| | | LKA-SA2.2 | General Reproducibility | 0.0032** | **0.6289** |
| LKA-SA3 | Functional Performance | LKA-SA3.1 | Control-Free Corridor | 0.6972 | 0.0870 |
| | | LKA-SA3.2 | Oscillation Behavior | 0.1886 | **0.2869** |
| | | LKA-SA3.3 | Lateral Jerk | 0.1123 | **0.3465** |
| | | LKA-SA3.4 | Handover Behavior on Intervention Termination | 0.6419 | 0.1040 |

TABLE V

CORRELATION ANALYSIS OF THE CRITERIA THAT ARE SIGNIFICANTLY DIFFERENT BETWEEN TWO DOMAINS OR HAVE A MEDIUM TO LARGE EFFECT

| # | Criteria | $r$ | $p$-Value |
|---|---|---|---|
| LDA-SA1.1 | General Intervention Strength | 0.60 | 0.0033** |
| LDA-SA1.2 | Override Capability at Lane Boundary | 0.52 | 0.0131* |
| LKA-SA2.2 | General Reproducibility | 0.65 | 9.4752e-4** |

difference between the domains. The *CV* of these SAs are between 0.19 and 0.44, which are considered low to moderate. The results of the Levene's test show that none of the SAs have siginificantly different variances in the vehicle and the simulator. Their low to moderate variance and the similarity between the two domains further support the reliability of the Wilcoxon test results. However, the medium effect size $r$ of the non-significant difference cannot yet be explained with the present dataset and requires further study.

In the next step, for subjective criteria whose absolute validity has not been confirmed, it is necessary to examine whether there is at least relative validity. For this purpose, correlation analysis is conducted for all the SAs that are significantly different between the vehicle and the simulator. The results are presented in Table V. Note that $r$ here refers to *Pearson's* correlation coefficient and is to be differentiated from the effect size. The results show that all three SAs show a moderate to strong significant positive correlation. This indicates that the drivers' subjective assessment in the real vehicle and in the simulator are at least consistent and shows a comparable trend. Thus, at least relative subjective validity of the simulator in terms of LDA-SA1.1 *General Intervention Strength*, LDA-SA1.2 *Override Capability* and LKA-SA2.2 *General Reproducibility* can be confirmed.

5) *Transferibility of the SUT on a Dynamic Simulator:* The drivers were asked to rate the following questions and were encouraged to leave a comment in addition after all the test runs of both SUT were completed: i) Have you recognized the functional characteristics of the LDA/LKA system on the dynamic simulator? ii) Is it conceivable to transfer the LDA/LKA development procedure to the dynamic driving simulator? The histogram and the statistical information of the results are shown in Fig 6 and Table VI. The medians of the responses to both questions with respect to both SUT are 4 (well recognizable/conceivable) and the maxima are 5 (very well recognizable/conceivable). The minima for LDA system for both questions are 3 (partly recognizable/conceivable). Only one driver rated both questions for LKA system with a minimum of 2 (barely recognizable/conceivable). In total, more than 68% of the drivers find that the LDA and LKA functional characteristics are well to very well recognizable on the dynamic simulator and it is well to very well conceivable, that the system development procedure can be transferred to the dynamic driving simulator.

There are a total of 64 comments on the 4 summary questions, namely question i) and ii) with respect to LDA and LKA. 35 of those are for LDA and 29 for LKA. These comments can be summarized into 7 categories, which are listed in Table VII. 20 comments in total mentioned explicitly that the system characteristics are recongnizable in the dynamic simulator. 16 comments in total explicitly approved the practicality of the driving simulator. They mentioned that it is meaningful to use a dynamic driving simulator, at least in the early stages of system design, because it provides a safe and reproducible test environment. It should be noted that these comments are open additional comments to questions i) and ii) without many

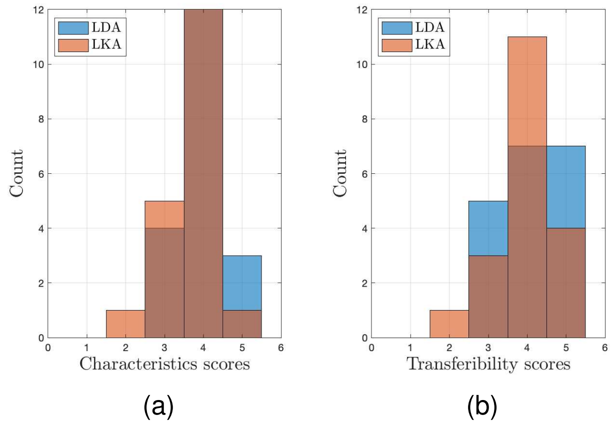


Fig. 6. Histograms of the subjects' responses to the summary questions. (a) Functional characteristics recognition. (b) Conceivability of the system transferibility.

TABLE VI
STATISTICAL INFORMATON OF THE SUMMARY QUESTIONS

| i) Functional characteristics recognition | | |
|---|---|---|
| | median | >= well recognizable |
| **LDA** | 4 (well recognizable) | 78.9% |
| **LKA** | 4 (well recognizable) | 68.4% |
| ii) Conceivability of system transferibility | | |
| | median | >= well conceivable |
| **LDA** | 4 (well conceivable) | 73.7% |
| **LKA** | 4 (well conceivable) | 78.9% |

predetermined conditions. This means that the number of each comment category and the comparison of these numbers does not necessarily have a statistical significance. The purpose of summarizing and analyzing these comments is simply to sort out any notable features that were noticed by more subjects and are less likely to be biased opinions.

## *B. Regression Analysis of Subjective Criteria*

In order to find out the relationship between the main criteria and their sub-criteria, which can also be seen as the customer level and expert level criteria, and examine which aspects influence the rating of the main criteria the most, regression analysis is conducted for each SUT in each subjective category in both test environments. Before that, collinearity of the sub-criteria was tested to avoid redundant predictors. The LDA subcriterion LDA-SA1.1 *General Intervention Strength* shows a high collinearity with LDA-SA1.2 *Override Capability at Lane Boundary*, and LKA-SA2.1 *Maximum Lane Deviation* with LKA-SA2.2 *General Reproducibility* for LKA system. Both collinear relationships can be identified in the vehicle and the simulator assessment with a correlation coefficient larger than 0.7. The complete collinearity matrices of LDA and LKA subjective sub-criteria are presented in Table XI.

Both single and multiple linear regression (SLR and MLR) were used to model the relationship between the main criteria and the sub-criteria. The significant SLR models ($p < 0.05$) with their estimated model coefficients ($b_0$, $b$), coefficient of determination $R^2$ and respective overall model $p$-value are listed in Table VIII and IX. Vehicle and simulator SA regression results are presented side-by-side for comparison.

The LDA main criterion LDA-SA1 *Overall Driver Interaction* does not show dependency on any of the sub-criteria in either domain. LDA-SA2 *Overall Perceived Safety* shows a similar relationship with the sub-criterion LDA-SA2.2 *General Reproducibility* in both domains ($R^2 = 0.56$ in vehicle, $R^2 = 0.60$ in simulator). *Maximal Lane Overshoot* of LDA does not directly influence the assessment of drivers' perceived safety in vehicle ($R^2 = 0.00, p = 0.9810$), but significantly in the driving simulator ($R^2 = 0.42, p = 0.0012$). LDA-SA3 *Overall Functional Performance*, however, shows inconsistency in the two domains. This main criterion is explained by LDA-SA3.6 *Handover Behavior* in vehicle and not by LDA-SA3.4 *Disengagement Angle* ($R^2 = 0.17, p = 0.0565$), and the other way around in simulator (*Handover Behavior* $R^2 = 0.02, p = 0.5794$).

The regression results of LKA are more consistent throughout the test domains compared to LDA. Apart from the fact that LKA-SA1 *Overall Driver Interaction* has a significant dependency on the sub-criterion LKA-SA1.5 *Steering Wheel Motion* only in the driving simulator but not in the real vehicle ($R^2 = 0.17, p = 0.565$ in vehicle), all other significant models can be found in both domains with similar model coefficients. Furthermore, except for LKA-SA3.1 *Control-Free Corridor*, all the other sub-criteria are significant in explaining the respective main criteria. Especially the sub-criteria LKA-SA1.2 *Co-Steering*, LKA-SA2.1 *Maximal Lateral Deviation* and LKA-SA2.2 *General Reproducibility*, they explain 44-59% of the variation of the respective main criteria both in vehicle and in simulator with a high significance of $p < 0.001$.

The above results show that co- and counter-steering and the intervention intensity play an essential role in driver's perception of driver-vehicle interaction when driving with LKA. LDA-SA2 and LKA-SA2 *Overall Perceived Safety* are both highly positively dependent on the sub-criterion *General Reproducibility*, which means a higher reproducibility of the assistant system leads to a better safety feeling. This relationship stands both in the road tests and in the dynamic driving simulator. The drivers relate the SA3 *Overall Functional Performance* of LDA and LKA with different sub-aspects. But SA3.1 *Control-Free Corridor* cannot directly explain drivers' perception functional performance of neither system in neither domains. Besides, it can be noted that the drivers subjectively strongly relate the functional performance of LKA to lateral jerk, whereas this is not the case for LDA.

MLR is utilized in the next step to explore more complex relationships of subjective assessment using more than one predictors. To maintain the interpretability of the models and avoid overfitting, only MLR models with a lower value of Akaike Information Criterion (AIC) than the SLR models are selected. The optimal MLR models are illustrated in Fig 7 and the model coefficients as well as quality criteria are listed as in Table X. *Overall Functional Performance* of LDA in the simulator can be better explained by adding an additional predictor LDA-SA3.2 *Returning Behavior* with a higher $R^2$ and lower $p$-value. Similarly, the MLR models of *Overall Perceived Safety* and *Overall Functional Performance*

TABLE VII
SUMMARY OF COMMENTS ON LDA AND LKA FROM SUBJECTS

| Category | Comments | LDA | LKA | Sum |
|---|---|---|---|---|
| 1 | System characteristics are recognizable in the dynamic simulator. | 10/35 | 10/29 | 20/64 |
| 2 | Reproducibility/Symmetry of the function is different in the simulator than in the vehicle. | 4/35 | 4/29 | 8/64 |
| 3 | Two-step intervention is not recognizable in the simulator. | 4/35 | 1/29 | 5/64 |
| 4 | Oscillation behavior when returning in the lane is different in the simulator than in the vehicle. | 3/35 | 3/29 | 6/64 |
| 5 | Different intervention duration. | 2/35 | 3/29 | 5/64 |
| 6 | Different intervention strength. | 5/35 | 4/29 | 9/64 |
| 7 | Dynamic simulator is meaningful/sufficient for the basic design of the function, it is practical, safe and reproducible. | 8/17 | 8/15 | 16/32 |

TABLE VIII
SIGNIFICANT SLR MODELS FOR LDA ($p^* < 0.05, p^{**} < 0.01$)

| | | | Vehicle | | | | Simulator | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Main Criteria | # | Sub-criteria | $b_0$ | $b$ | $R^2$ | $p$-value | $b_0$ | $b$ | $R^2$ | $p$-value |
| LDA-SA2 | LDA-SA2.1 | Maximal Lane Overshoot | - | - | - | - | 3.98 | 0.41 | 0.42 | 0.0012** |
| | LDA-SA2.2 | General Reproducibility | 1.94 | 0.70 | 0.56 | 5.7575e-5** | 2.08 | 0.65 | 0.60 | 2.2008e-5** |
| LDA-SA3 | LDA-SA3.4 | Disengagement Angle | - | - | - | - | 4.54 | 0.32 | 0.29 | 0.0146* |
| | LDA-SA3.6 | Handover Behavior | 3.72 | 0.44 | 0.22 | 0.0296* | - | - | - | - |

TABLE IX
SIGNIFICANT SLR MODELS FOR LKA ($p^* < 0.05, p^{**} < 0.01$)

| | | | Vehicle | | | | Simulator | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Main Criteria | # | Sub-criteria | $b_0$ | $b$ | $R^2$ | $p$-value | $b_0$ | $b$ | $R^2$ | $p$-value |
| LKA-SA1 | LKA-SA1.1 | Counter-Steering | 0.88 | 0.78 | 0.37 | 0.0025** | 3.20 | 0.48 | 0.21 | 0.0378** |
| | LKA-SA1.2 | Co-Steering | -0.21 | 0.91 | 0.49 | 0.0003** | -0.49 | 0.97 | 0.59 | 4.5673e-5** |
| | LKA-SA1.3 | General Intervention Strength | 3.04 | 0.46 | 0.53 | 0.0001** | 4.27 | 0.35 | 0.27 | 0.0155* |
| | LKA-SA1.4 | Steering Wheel Torque Build-up | 0.66 | 0.82 | 0.53 | 0.0001** | 2.67 | 0.58 | 0.34 | 0.0053** |
| | LKA-SA1.5 | Steering Wheel Motion | - | - | - | - | 2.48 | 0.56 | 0.35 | 0.0045** |
| LKA-SA2 | LKA-SA2.1 | Maximal Lateral Deviation | 3.10 | 0.41 | 0.44 | 0.0008** | 2.25 | 0.54 | 0.59 | 2.8358e-5** |
| | LKA-SA2.2 | General Reproducibility | 2.58 | 0.59 | 0.49 | 0.0003** | 1.05 | 0.71 | 0.54 | 9.9500e-5** |
| LKA-SA3 | LKA-SA3.2 | Oscillating Behavior | 1.83 | 0.72 | 0.43 | 0.0009** | 3.06 | 0.54 | 0.23 | 0.0318* |
| | LKA-SA3.3 | Lateral Jerk | 2.73 | 0.33 | 0.26 | 0.0159* | 2.88 | 0.41 | 0.32 | 0.0093** |
| | LKA-SA3.4 | Handover Behavior | 2.40 | 0.57 | 0.20 | 0.0415* | 2.10 | 0.67 | 0.26 | 0.0206* |

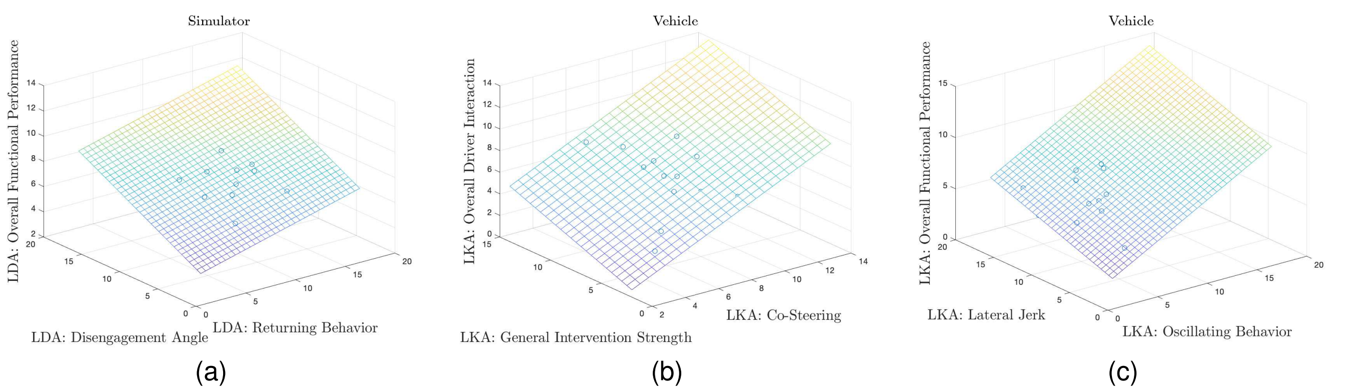


Fig. 7. Optimal multi linear regression results.

for LKA evaluation in vehicle present a better model quality than the SLR models by combining 2 predictors. Although these optimal MLR models are not consistently found in both test environments, they still provide relevant insights for more complex modelling as more data becomes available in the future. For example, just because *Returning Behavior* does not explain the variation in the main criterion alone, it does not necessarily mean that it does not contribute to the evaluation of system performance at all.

It should be kept in mind that only linear relationships were investigated because of the limited data points. The potential non-linear relationships are not revealed and the results should not be extrapolated.

## IV. DISCUSSION

The drivers' assessment of the main criteria in the aspects of driver interaction, perceived safety and functional performance show no significant difference between the road test with the real vehicle and in the virtual environment in the driving simulator. The comparison of most sub-criteria in the two test environments also show no significant difference for both systems except for 3 criteria. The drivers perceived the general

TABLE X
OPTIMAL MLR MODELS FOR LDA AND LKA ($p^{**} < 0.01$)

| Domain | Main Criterion | # | Sub-criteria | $b$ | $R^2$ | $p$-value |
|---|---|---|---|---|---|---|
| Simulator | LDA-SA3 | <br>LDA-SA3.2<br>LDA-SA3.4 | (intercept)<br>Returning Behavior<br>Disengagement Angle | 2.53<br>0.21<br>0.33 | 0.45 | 0.0060** |
| Vehicle | LKA-SA1 | <br>LKA-SA1.2<br>LKA-SA1.3 | (intercept)<br>Co-Steering<br>General Intervention Strength | -1.52<br>0.72<br>0.37 | 0.81 | 1.1711e-7** |
| Vehicle | LKA-SA3 | <br>LKA-SA3.2<br>LKA-SA3.3 | (intercept)<br>Oscillating Behavior<br>Lateral Jerk | 0.24<br>0.54<br>0.27 | 0.55 | 0.0011** |

intervention strength of LDA to be weaker in the driving simulator and found it easier to override the system around lane boundaries in the simulator. As for LKA, the drivers assessed the reproducibility of the system to be significantly higher in the simulator compared to the real vehicle. The drivers' comments indicated that the interventions in the road test were dependent on the side of the road, whilst the interventions in the simulator were more symmetrical. However, the comments were not unanimous or did not mention which side was stronger than the other. One potential reason for this phenomenon is the road bank angle, which was not included in the road model in the simulation environment. The drivers were not given a specific driving direction during the road tests and could test the system on a straight section in both directions for a more efficient test procedure. This might have caused the inconsistency in the comments regarding the stronger side. However, the rating of these 3 criteria, i.e. LDA-SA1.1 *General Intervention Strength*, LDA-SA1.2 *Override Capability at Lane Boundary* and LKA-SA2.2 *General Reproducibility*, all show a moderate to strong significant correlation between the two test environments. This indicates that the relative subjective validity with respect to these SAs can be confirmed.

Other than the statistical significance, the effect size of each test is also analyzed to examine the practical meaning of the domain difference of each subjective criterion. The difference of the subjectively perceived lateral jerk is found to show a medium effect for both LDA and LKA. The perceived lateral jerk in the simulator is smaller than in the road test, which could be traced back to the motion cueing of the dynamic simulator. The motion platform provides acceleration cues that stimulate the driver's vestibular system, enhancing the perception of vehicle behavior [32]. However, physical limitations in cockpit displacement and actuator power restrict the simulator's ability to replicate the full vehicle dynamics [33]. The state-of-the-art dynamic simulator from AB Dynamics used in this study operates with a classical washout motion cueing algorithm (MCA), incorporating a sway gain scaling of 0.75 and a high-pass filter cut-off frequency of 0.1 Hz. These parameters result in attenuated and more gradual motion cues compared to real-world driving conditions, potentially influencing the driver's perception of lateral jerk. The test environment also showed a medium effect in the subjective assessment of LKA's general intervention strength and torque build-up. In addition, the perceived oscillation behavior of LKA is slightly more intense in the simulator than in the real vehicle and this difference has a medium effect. For LDA, the difference of the subjectively perceived maximum lane overshoot in the simulator compared to the real vehicle has a medium effect size, even though the difference is not significant. However, the low to moderate variance of the SAs and the similarity of the variance between the two domains confirmed the reliability of the Wilcoxon test results. This suggests that the medium effect of the domain difference with respect to these SAs is not caused by data variability. Considering the results of the Wilcoxon test (equal medians) and the Levene's test (equal variances), the absolute subjective validity in terms of these SAs can be confirmed based on this study. Nevertheless, the medium effect size of these SAs is valuable for further research and should be taken into account in future study design, even though the effects are not statistically significant in this study. Possible influencing factors that could be explored in more detail include subtle simulator latencies or visual rendering nuances. While the simulator used is a state-of-the-art, high-fidelity system with high motion bandwidths and a 240 Hz visual setup, minimal latencies, though likely negligible, may still arise from motion cueing algorithms, multi-PC rendering synchronization, or actuator response times. These could subtly affect human driver perception, such as through sensorimotor mismatch, and may warrant closer examination in advanced follow-up studies.

When comparing the common sub-criteria of LDA and LKA, there are also some interesting findings. First of all, as mentioned above, the perceived *General Intervention Strength* and *Lateral Jerk* is smaller in the simulator than in the real vehicle for both systems. Furthermore, the drivers' subjective assessment of *Control-Free Corridor* and *Handover Behavior on Intervention Termination* do not show significant difference or noticeable effect between the driving simulator and the real vehicle for both LDA and LKA. As for differences, the drivers noticed a significant difference in *General Reproducibility* between the test environment when driving with LKA, but not with LDA.

The regression analysis is then conducted to explore the sub-aspects that actually influence the drivers' perception of interaction, safety and functional performance of the tested lateral guidance system. It is found out that drivers' assessment of interaction with LDA is not dependent on either LDA-SA1.1 *General Intervention Strength* or LDA-SA1.2 *Override Capability at Lane Boundary* in both test environments. This also explains why the subjective assessments of the two sub-criteria are significantly different between the simulator and the real vehicle, but LDA-SA1 *Overall Driver Interaction* is not. This means that drivers do not directly relate the interaction with LDA to the intensity of the intervention, which may be caused by the short intervention duration of LDA and the safety-oriented character of the system. The factors that actually influence the interaction evaluation need to be further studied. Some potential factors could be drawn from the collected comments in Table VII, such as intervention duration or the actual course of the assist torque build-up. On the other hand, the drivers' assessment of the interaction with LKA is positively dependent on LKA-SA1.1 - LKA-SA1.4 in both domains. This shows that a more harmonious co- and counter-steering and a stronger intervention lead to a better perceived driver-vehicle interaction when driving with LKA.

The drivers' perceived safety shows a clear tendency across systems and domains and is positively dependent on *General Reproducibility*, i.e. a higher system reproducibility leads to a greater safety feeling. For LKA, a smaller lane deviation also has a positive influence on the perceived safety in both test environments, whereas for LDA, a smaller lane overshoot only has a positive influence in the simulator. It must be emphasized that the evaluation of the lane deviation is subjective. Therefore, this result has to be interpreted as an interpersonal difference in lateral deviation tolerance leading to different feelings of safety. The regression results of the two systems *Functional Performance* are not comparable. This is to be expected as the performance criteria for LDA and LKA are defined differently due to their different primary purpose. However, there are still some common aspects that can be derived. For both systems, *Handover Behavior* is a critical factor in assessing system functional performance. On the other hand, drivers do not relate *Functional Performance* directly to *Control-Free Corridor*, or at least it is not the dominant influencing factor. And *Lateral Jerk* plays a more important role for the perception of system performance for LKA as for LDA.

In general, more SLR models are found for LKA than for LDA. These LKA models are also more consistent in different test environments. However, it is rash to conclude that the sub-criteria of LDA assessment are mostly irrelevant. Underlying relationships that are more complex than a simple linear relationship, such as multivariate or nonlinear, cannot yet be excluded. The results of the MLR show the complexity of the driver's perception and evaluation of the system characteristics and indicate the potential for more complex modeling methods in the future.

Based on these results, it can be concluded that integrating the subjective evaluation method in the early phases of lateral guidance system development is feasible. However, it should still be mentioned that this study reflects a common limitation of subjective test drives and subject studies: balancing practical feasibility with statistical robustness. While the sample size is limited from a statistical standpoint, it aligns with established practice and is justified by constraints such as the availability of expert drivers, project duration, and confidentiality. Prior studies such as [34], [35] [36], and [37] support the adequacy of such sample sizes for subjective evaluations. The use of professional test drivers further strengthens the results by ensuring consistent, expert assessments. Nevertheless, the constrained sample size inherently limits statistical power, which should be considered when interpreting the results. Furthermore, for research objectives focused on broader user perspectives, such as user acceptance, it would be essential to include naïve drivers and ensure a more representative participant demographic.

In addition, several challenges remain for practical implementation. One key aspect is ensuring the subjective validity of the simulator across all necessary test scenarios, such as curves with different radii at various velocities, to guarantee representative and reliable results before practical application. Another critical factor is the seamless integration of the simulator-based evaluation into existing development workflows. This includes assessing their alignment with traditional vehicle testing methods and determining whether methodological adjustments are necessary. From a cost-benefit perspective, an essential consideration is identifying the minimum required simulator fidelity that ensures reliable subjective evaluations. Extremely high-fidelity simulations may not always be necessary if a lower-fidelity setup can still provide meaningful and accurate assessments. For example, Himmels et al. [38] found fewer differences than expected in driving behavior, as well as in speed and distance perception, across six simulators with varying levels of fidelity. In several scenarios, lower-fidelity simulators achieved absolute validity in speed and distance perception. The authors highlighted the importance of aligning simulator fidelity with specific use cases. Moreover, higher fidelity does not inherently guarantee validity [11]. A structured assessment comparing the cost savings and efficiency gains of simulator use against potential trade-offs would be valuable in evaluating its practical viability. Addressing these challenges will help refine the role of simulators in the development process and ensure their cost-effective application in lateral guidance system evaluation.

## V. Conclusion

This paper designed and conducted a back-to-back subjective validation study, in which 22 professional drivers took part, to compare the subjective assessment of the drivers in a real vehicle and in a high-fidelity, objectively validated dynamic driving simulator. The aim of the study was to investigate the subjective validity of the driving simulator and thus to explore the possibility of transferring the human-centered development process of lateral guidance systems to the virtual environment. The analysis of the study shows promising results in general. The subjects perceived and evaluated the main aspects, i.e. *Driver Interaction*, *Perceived Safety* and *Functional Performance*, of both LDA and LKA similarly in the vehicle and in the simulator. The evaluation of most of the sub-criteria does not show significant differences between the two domains, with an exception of LDA-SA1.1 *General Intervention Strength*, LDA-SA1.2 *Override Capability at Lane Boundary* and LKA-SA2.2 *General Reproducibility*. However, the ratings of these three SAs show a moderate to

TABLE XI
COLLINEARITY OF SUBJECTIVE SUB-CRITERIA OF LDA AND LKA

| | | LDA-SA1.1 | LDA-SA1.2 | LDA-SA2.1 | LDA-SA2.2 | LDA-SA3.1 | LDA-SA3.2 | LDA-SA3.3 | LDA-SA3.4 | LDA-SA3.5 | LDA-SA3.6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Vehicle** | | | | | | | | | | | |
| LDA-SA1.1 | General Intervention Strength | 1.00 | | | | | | | | | |
| LDA-SA1.2 | Override Capability at Lane Boundary | 0.90 | 1.00 | | | | | | | | |
| LDA-SA2.1 | Maximum Lane Overshoot | | | 1.00 | | | | | | | |
| LDA-SA2.2 | General Reproducibility | | | -0.16 | 1.00 | | | | | | |
| LDA-SA3.1 | Control-Free Corridor | | | | | 1.00 | | | | | |
| LDA-SA3.2 | Returning Behavior | | | | | 0.45 | 1.00 | | | | |
| LDA-SA3.3 | Disengagement Position | | | | | 0.26 | 0.58 | 1.00 | | | |
| LDA-SA3.4 | Disengagement Angle | | | | | 0.06 | -0.03 | -0.02 | 1.00 | | |
| LDA-SA3.5 | Lateral Jerk | | | | | -0.07 | -0.12 | 0.04 | 0.28 | 1.00 | |
| LDA-SA3.6 | Handover Behavior on Intervention Termination | | | | | -0.02 | 0.04 | -0.19 | 0.22 | 0.26 | 1.00 |
| **Simulator** | | | | | | | | | | | |
| LDA-SA1.1 | General Intervention Strength | 1.00 | | | | | | | | | |
| LDA-SA1.2 | Override Capability at Lane Boundary | 0.94 | 1.00 | | | | | | | | |
| LDA-SA2.1 | Maximum Lane Overshoot | | | 1.00 | | | | | | | |
| LDA-SA2.2 | General Reproducibility | | | 0.61 | 1.00 | | | | | | |
| LDA-SA3.1 | Control-Free Corridor | | | | | 1.00 | | | | | |
| LDA-SA3.2 | Returning Behavior | | | | | 0.05 | 1.00 | | | | |
| LDA-SA3.3 | Disengagement Position | | | | | 0.09 | 0.36 | 1.00 | | | |
| LDA-SA3.4 | Disengagement Angle | | | | | 0.54 | -0.09 | 0.06 | 1.00 | | |
| LDA-SA3.5 | Lateral Jerk | | | | | 0.64 | 0.40 | 0.03 | 0.07 | 1.00 | |
| LDA-SA3.6 | Handover Behavior on Intervention Termination | | | | | 0.39 | -0.07 | 0.01 | 0.58 | 0.09 | 1.00 |

| | | LKA-SA1.1 | LKA-SA1.2 | LKA-SA1.3 | LKA-SA1.4 | LKA-SA1.5 | LKA-SA2.1 | LKA-SA2.2 | LKA-SA3.1 | LKA-SA3.2 | LKA-SA3.3 | LKA-SA3.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Vehicle** | | | | | | | | | | | | |
| LKA-SA1.1 | Counter-Steering | 1.00 | | | | | | | | | | |
| LKA-SA1.2 | Co-Steering | 0.60 | 1.00 | | | | | | | | | |
| LKA-SA1.3 | General Intervention Strength | 0.29 | 0.11 | 1.00 | | | | | | | | |
| LKA-SA1.4 | Steering Wheel Torque Build-up | 0.23 | 0.60 | 0.51 | 1.00 | | | | | | | |
| LKA-SA1.5 | Steering Wheel Motion | 0.14 | 0.57 | 0.15 | 0.65 | 1.00 | | | | | | |
| LKA-SA2.1 | Maximum Lane Deviation | | | | | | 1.00 | | | | | |
| LKA-SA2.2 | General Reproducibility | | | | | | 0.78 | 1.00 | | | | |
| LKA-SA3.1 | Control-Free Corridor | | | | | | | | 1.00 | | | |
| LKA-SA3.2 | Oscillation Behavior | | | | | | | | 0.39 | 1.00 | | |
| LKA-SA3.3 | Lateral Jerk | | | | | | | | 0.44 | 0.37 | 1.00 | |
| LKA-SA3.4 | Handover Behavior on Intervention Termination | | | | | | | | -0.08 | 0.21 | 0.42 | 1.00 |
| **Simulator** | | | | | | | | | | | | |
| LKA-SA1.1 | Counter-Steering | 1.00 | | | | | | | | | | |
| LKA-SA1.2 | Co-Steering | 0.57 | 1.00 | | | | | | | | | |
| LKA-SA1.3 | General Intervention Strength | 0.64 | 0.52 | 1.00 | | | | | | | | |
| LKA-SA1.4 | Steering Wheel Torque Build-up | 0.62 | 0.70 | 0.61 | 1.00 | | | | | | | |
| LKA-SA1.5 | Steering Wheel Motion | 0.19 | 0.62 | 0.27 | 0.56 | 1.00 | | | | | | |
| LKA-SA2.1 | Maximum Lane Deviation | | | | | | 1.00 | | | | | |
| LKA-SA2.2 | General Reproducibility | | | | | | 0.72 | 1.00 | | | | |
| LKA-SA3.1 | Control-Free Corridor | | | | | | | | 1.00 | | | |
| LKA-SA3.2 | Oscillation Behavior | | | | | | | | 0.62 | 1.00 | | |
| LKA-SA3.3 | Lateral Jerk | | | | | | | | 0.13 | 0.30 | 1.00 | |
| LKA-SA3.4 | Handover Behavior on Intervention Termination | | | | | | | | 0.22 | 0.29 | 0.21 | 1.00 |

strong correlation across domains, indicating relative validity. Considering both statistical significance and the effect size, the drivers perceived a weaker intervention intensity, including general intervention strength, override capability and torque build-up, in the driving simulator compared to in the real vehicle for both LDA and LKA. For the case of LKA, the drivers perceived a significantly higher reproducibility of the system in the simulator than in the road test. In addition, the perceived lateral jerk when driving with both systems is smaller than in the vehicle, which is caused by motion cueing and is plausible. The difference is not statistically significant, yet shows a medium effect size for both systems. The statistical analysis of the subjective ratings is supported by the results of the summary questions as well: More than 68% of the subjects, who are professionals from the field of ADAS development, found the system characteristics well recognizable and it is well conceivable to transfer the development procedure to the dynamic driving simulator.

It was then examined which of the sub-criteria actually had an impact on the overall perception and assessment of the three main characteristics of lateral guidance systems. The regression results show that *Driver Interaction* of LDA is not directly dependent on the intervention intensity, while a stronger intervention of LKA is related to a better interaction. Additionally, co- and counter-steering also play an essential role in the assessment of driver's interaction with LKA. Driver's perceived safety of both LDA and LKA is highly dependent on the general reproducibility of systems in both test environments. A smaller subjective perception of lateral deviation of the system also contributes to a better safety feeling, with an exception of the assessment of LDA in real vehicle. In terms of functional performance, the most relevant factors influencing the drivers' assessment in the vehicle and in the simulator are identical when driving with LKA, but different when driving with LDA. Furthermore, lateral jerk plays a more dominant role in the assessment of LKA functional performance than in the assessment of LDA.

In conclusion, absolute subjective validity of the dynamic driving simulator can be established for the use case of evaluating lateral guidance systems in the aspects of *Driver Interaction*, *Perceived Safety* and *Functional Performance*, and relative subjective validity can be established in terms of the LDA sub-aspects of *General Intervention Strength* and *Override Capability*, as well as the LKA sub-aspect of *General Reproducibility*. It can be concluded that the high-fidelity dynamic driving simulator has sufficient potential for the subjective assessment and fine-tuning of lateral guidance systems. The effectiveness of the utilization of the driving simulator is convincing based on both the subjective impression of the test drivers and the statistical analysis. The regression results also confirmed the validity of the assessment methodology of lateral guidance systems.

## Appendix

See Table XI.

## Acknowledgment

The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Programme.

## References

[1] *Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on Type-approval Requirements for Motor Vehicles and Their Trailers, and Systems, Components and Separate Technical Units Intended for Such Vehicles, As Regards Their General Safety and the Protection of Vehicle Occupants and Vulnerable Road Users*, European Union, Brussels, Belgium, Dec. 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019R2144

[2] Surface Vehicle Recommended Practice–Taxonomy and Definitions for Terms Related To Driving Automation Systems for On-Road Motor Vehicles, Standard J3016, SAE International, Warrendale, PA, USA, 2021.

[3] H. Winner, S. Hakuli, F. Lotz, and C. Singer, *Handbook of Driver Assistance Systems*. Cham, Switzerland: Springer, 2016. [Online]. Available: http://link.springer.com/10.1007/978-3-319-12352-3

[4] S. Wei, P. E. Pfeffer, and J. Edelmann, "State of the art: Ongoing research in assessment methods for lane keeping assistance systems," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 9, pp. 5853–5875, Sep. 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10106476/

[5] B. Schick, C. Seidler, S. Aydogdu, and Y.-J. Kuo, "Driving experience vs. mental stress with automated lateral guidance from the customer's point of view," in *Proc. 9th International Munich Chassis Symposium*, P. Pfeffer, Ed. Wiesbaden, Germany: Springer, 2018, pp. 27–44. [Online]. Available: http://link.springer.com/10.1007/978-3-658-22050-1_5

[6] AllianzDirect. *Selbstfahrende Autos: 20 % der Befragten sind skeptisch*. Accessed: Feb. 20, 2025. [Online]. Available: https://www.allianzdirect.de/kfz-versicherung/selbstfahrende-autos-umfrage/

[7] Y. Li et al., "Choose your simulator wisely: A review on open-source simulators for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 5, pp. 4861–4876, May 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10461065/

[8] J. Haselberger, M. Böhle, B. Schick, and S. Müller, "Exploring the influence of driving context on lateral driving style preferences: A simulator-based study," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 4, pp. 5448–5466, Apr. 2025.

[9] Y. M. Tang, D. Zhao, T. Chen, and X. Fu, "A systematic review of abnormal behaviour detection and analysis in driving simulators," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 109, pp. 897–920, Feb. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847825000026

[10] J. Bortz and C. Schuster, *Statistik für Human-und Sozialwissenschaftler*, 7th ed., New York, NY, USA: Springer, 2010.

[11] R. A. Wynne, V. Beanland, and P. M. Salmon, "Systematic review of driving simulator validation studies," *Saf. Sci.*, vol. 117, pp. 138–151, Aug. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0925753518320459

[12] G. J. Blaauw, "Driving experience and task demands in simulator and instrumented car: A validation study," *Human Factors: J. Human Factors Ergonom. Soc.*, vol. 24, no. 4, pp. 473–486, Aug. 1982. [Online]. Available: https://journals.sagepub.com/doi/10.1177/001872088202400408

[13] M. Klüver, C. Herrigel, C. Heinrich, H.-P. Schöner, and H. Hecht, "The behavioral validity of dual-task driving performance in fixed and moving base driving simulators," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 37, pp. 78–96, Feb. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1369847815001953

[14] N. Mullen, J. Charlton, A. Devlin, and M. Bédard, "Simulator validity: Behaviors observed on the simulator and on the road," in *Handbook Driving Simulation for Engineering, Medicine, and Psychology*, D. L. Fisher, M. Rizzo, J. K. Caird, and J. D. Lee, Eds., Boca Raton, FL, USA: CRC Press, Apr. 2011. [Online]. Available: http://www.crcnetbase.com/doi/10.1201/b10836-14

[15] R. Lobjois, V. Faure, L. Désiré, and N. Benguigui, "Behavioral and workload measures in real and simulated driving: Do they tell us the same thing about the validity of driving simulation?," *Saf. Sci.*, vol. 134, Feb. 2021, Art. no. 105046. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0925753520304434

[16] M. Vollrath, "Statistical methods," in *Automotive Ergonomics*, H. Bubb, K. Bengler, R. E. Grünen, and M. Vollrath, Eds., Wiesbaden, Germany: Springer Fachmedien Wiesbaden, 2021, pp. 653–673. [Online]. Available: https://link.springer.com/book/10.1007/978-3-658-33941-8

[17] D. Llopis-Castelló, F. J. Camacho-Torregrosa, J. Marín-Morales, A. M. Pérez-Zuriaga, A. García, and J. F. Dols, "Validation of a low-cost driving simulator based on continuous speed profiles," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2602, no. 1, pp. 104–114, Jan. 2016. [Online]. Available: https://journals.sagepub.com/doi/10.3141/2602-13

[18] T. Schwarzhuber, L. Wörle, M. Graf, and A. Eichberger, "Validity quantification of driver-in-the-loop simulation in motorsport," in *Proc. FISITA World Congr.-Tech. Programme*, Sep. 2021, pp. 1–16. [Online]. Available: https://api.semanticscholar.org/CorpusID:247828065

[19] Q. Hussain, W. K. M. Alhajyaseen, A. Pirdavani, N. Reinolsmann, K. Brijs, and T. Brijs, "Speed perception and actual speed in a driving simulator and real-world: A validation study," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 62, pp. 637–650, Apr. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1369847818307824

[20] J. Li, X. Zhao, S. Xu, J. Ma, and J. Rong, "The study of driving simulator validation for physiological signal measures," *Proc.-Social Behav. Sci.*, vol. 96, pp. 2572–2583, Nov. 2013. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1877042813024142

[21] S. H. de Frutos, M. S. Torner, and M. Castro, "Validation of a driving simulator for speed research on two-lane rural roads," *Transp. Res. Proc.*, vol. 71, pp. 156–163, Jan. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2352146523003514

[22] C. Andriola, G. R. Di Rado, D. S. Presta García, and C. T. Nodari, "Classification of driving simulators validation: A case study using an immersive driving simulator," *Accident Anal. Prevention*, vol. 213, Apr. 2025, Art. no. 107944. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457525000302

[23] G. L. G. Gómez et al., "Validation of a moving base driving simulator for subjective assessments of steering feel and handling," in *Proc. 13th Int. Symp. Adv. Vehicle Control*, Munich, Germany, Sep. 2016. [Online]. Available: http://www.crcnetbase.com/doi/book/10.1201/9781315265285

[24] H. Bellem, M. Klüver, M. Schrauf, H.-P. Schöner, H. Hecht, and J. F. Krems, "Can we study autonomous driving comfort in moving-base driving simulators? A validation study," *Human Factors: J. Human Factors Ergonom. Soc.*, vol. 59, no. 3, pp. 442–456, May 2017. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0018720816682647

[25] M. Jentsch and A. C. Bullinger, "Is simulation (not) enough? Results of a validation study of an autonomous emergency braking system on a test track and in a static driving simulator," in *Proc. Hum. Factors Ergonom. Soc. Europe Chapter Annu. Conf.* Lisbon, Portugal: HFES, 2014, pp. 161–174.

[26] 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," *Behav. Res. Methods*, vol. 38, no. 2, pp. 314–324, May 2006. [Online]. Available: http://link.springer.com/10.3758/BF03192783

[27] S. T. Godley, T. J. Triggs, and B. N. Fildes, "Driving simulator validation for speed research," *Accident Anal. Prevention*, vol. 34, no. 5, pp. 589–600, Sep. 2002. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0001457501000562

[28] S. Trunzer, K. Schuler, and B. Schick, "Driver behavior on a motion-base simulator for vehicle dynamics maneuvers," in *Proc. Driving Simulation Conf. Eur. (DSC)*, Antibes, France, 2023, pp. 45–52.

[29] M. Böhle, B. Schick, and S. Müller, "Steering feedback in dynamic driving simulators: The influence of steering wheel vibration and vehicle motion frequency," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 12, pp. 1–17, Dec. 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10531699/

[30] S. Wei, M. Becker, P. E. Pfeffer, and J. Edelmann, "Investigation of the suitability of a static driving simulator for the characterization of lane departure avoidance systems," *IEEE Trans. Intell. Vehicles*, early access, Jul. 24, 2024, doi: 10.1109/TIV.2024.3433369.

[31] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. New York, NY, USA: Psychology Press, 2009.

[32] I. G. Salisbury and D. J. N. Limebeer, "Motion cueing in high-performance vehicle simulators," *Vehicle Syst. Dyn.*, vol. 55, no. 6, pp. 775–801, Jun. 2017. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/00423114.2017.1280172

[33] G. Reymond and A. Kemeny, "Motion cueing in the Renault driving simulator," *Vehicle Syst. Dyn.*, vol. 34, no. 4, pp. 249–259, Oct. 2000. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1076/vesd.34.4.249.2059

[34] M. Harrer, "Characterisation of steering feel," Ph.D. dissertation, Dept. Mech. Eng., University of Bath, Bath, U.K., Mar. 2007.

[35] S. Monsma, "Feel the tire—Tire influence on driver's handling assessment," Ph.D. dissertation, Dept. Eng. Des. Prod., Aalto University, Espoo, Finland, Nov. 2015.

[36] H. Oschlies, *Komfortorientierte Regelung für die automatisierte Fahrzeugquerführung*. Wiesbaden, Germany: Springer Fachmedien Wiesbaden, 2019. [Online]. Available: http://link.springer.com/10.1007/978-3-658-25235-9

[37] M. Festner, "Objektivierte Bewertung des Fahrstils auf Basis der Komfortwahrnehmung bei hochautomatisiertem Fahren in Abhängigkeit fahrfremder Tätigkeiten: Grundlegende Zusammenhänge zur komfortorientierten Auslegung eines hochautomatisierten Fahrstils," Ph.D. dissertation, Dept. Mech. Process Eng., Univ. Duisburg-Essen, Duisburg, Germany, Dec. 2019. [Online]. Available: https://duepublico2.uni-due.de/receive/duepublico_mods_00070681

[38] C. Himmels, J. Venrooij, A. Parduzi, M. Peller, and A. Riener, "The bigger the better? Investigating the effects of driving simulator fidelity on driving behavior and perception," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 101, pp. 250–266, Feb. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847824000135


**Sijie Wei** received the B.S. degree in automotive engineering from Tongji University, Shanghai, China, in 2014, and the M.S. degree in automotive mechatronics from Munich University of Applied Sciences, Germany, in 2017. She is currently pursuing the Ph.D. degree with TU Wien. Her research interests include vehicle dynamics and human-centered development of advanced driver assistance systems.

**Matthias Becker** received the B.S. degree in physical engineering and the M.S. degree in automotive mechatronics from Munich University of Applied Sciences (MUAS) in 2014 and 2015, respectively. He is currently a Research Assistant at the Automotive Engineering Laboratory, MUAS, under the Direction of Prof. Peter E. Pfeffer. His research interests include the characterization, modeling, and simulation of automated vehicle lateral guidance systems, in particular the interaction with humans during the steering process.

**Peter E. Pfeffer** received the B.S. and M.S. degrees from TU Wien, Austria, and the Ph.D. degree, in 2007. From 2002 to 2007, he was a Lecturer at the Mechanical Engineering Department, University of Bath, where he researched the interaction of vehicle and steering system regarding handling. He has been a Professor of automotive engineering at Munich University of Applied Sciences, Germany, since 2007. He is currently the Scientific Director of the International Munich Chassis Symposium and an Editor of the *Steering Handbook* (Springer Cham, 2017). His research interests include automated driving, vehicle dynamics, steering systems, steering feel, elastomer, and hydro mounts.

**Johannes Edelmann** received the Ph.D. degree from TU Wien, Austria, in 2008, and his venia docenti for Technical Dynamics at TU Wien in 2011. From 2018 to 2022, he was an Associate Professor at the Institute of Mechanics and Mechatronics, TU Wien. Since October 2021, he has been a Professor of technical dynamics at the Institute of Mechanics and Mechatronics. His research interests include applied multi-body dynamics, stability of dynamical systems, vehicle dynamics, driver modeling, dynamic behavior of bicycles, scooters, motorcycles, and automobiles, development of driver assistance systems, and vehicle dynamics control systems.