This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Acceptance and Trust: Drivers' First Contact With Released Automated Vehicles in Naturalistic Traffic

Sarah Schwindt-Drews, Kai Storms, Steven Peters, and Bettina Abendroth

***Abstract*— This study investigates the impact of initial contact of drivers with an SAE Level 3 Automated Driving System (ADS) under real traffic conditions, focusing on the Mercedes-Benz Drive Pilot in the EQS. It examines Acceptance, Trust, Usability, and User Experience. Although previous studies in simulated environments provided insights into human-automation interaction, real-world experiences can differ significantly. The research was conducted on a segment of German interstate with 30 participants lacking familiarity with SAE Level 3 ADS. Pre- and post-driving questionnaires were used to assess changes in acceptance and trust. Supplementary metrics included post-driving ratings for usability and user experience. Findings reveal a significant increase in acceptance and trust following the first contact, confirming results from prior studies. Factors such as Performance Expectancy, Effort Expectancy, Facilitating Condition, and Perceived Safety were rated higher after initial contact with the ADS. However, inadequate communication from the ADS to the human driver was detected, highlighting the need for improved communication to prevent misuse or confusion about the operating mode. However, it's worth noting that most participants already had a high affinity for technology. Although overall reception was positive and showed an upward trend post first contact, the ADS was also perceived as demanding as manual driving. Future research should focus on a more diverse participant sample and include longer or multiple real-traffic trips to understand behavioral adaptations over time.**

***Index Terms*— Automated driving, acceptance, trust.**

I. Introduction

WITH the introduction of automated vehicles, drivers will be able to travel more comfortable in the future [1], [2]. Drivers in vehicles with SAE Level 3 are currently authorized to temporarily shift their focus from driving to other non-driving related tasks (NDRTs) [3]. Nevertheless, they must always be capable of resuming the driving task when requested to do so by the ADS. Automated vehicles have the potential of increasing road safety by minimizing human error [4]. Still, drivers must use these systems properly to achieve these benefits. A more comfortable and safer journey is only achievable if drivers accept the automated systems accordingly in order to actually use the functions [5], [6]. Adequate trust also is a critical variable for the appropriate use of automated vehicles [5], [6]. A lack of trust results in the non-utilization of the functions. Conversely, overtrust can be hazardous, as drivers overestimate the system's capabilities and are unable to react appropriately [7].

Manuscript received 18 January 2024; revised 29 July 2024; accepted 4 August 2024. The Associate Editor for this article was S. C. Wong. *(Sarah Schwindt-Drews and Kai Storms contributed equally to this work.) (Corresponding author: Kai Storms.)*

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

Sarah Schwindt-Drews and Bettina Abendroth are with the Institute of Ergonomics and Human Factors, Technische Universität Darmstadt, 64287 Darmstadt, Germany.

Kai Storms and Steven Peters are with the Institute of Automotive Engineering, Technische Universität Darmstadt, 64287 Darmstadt, Germany (e-mail: kai.storms@tu-darmstadt.de).

Digital Object Identifier 10.1109/TITS.2024.3443927

A. Acceptance and Trust

While acceptance according to [8] describes whether or why a technology is used or not, trust describes the attitude towards the technology without including its use [9].

Current studies on acceptance and trust in the context of automated driving focus in particular on which factors influence acceptance and trust of automated vehicles in society [1], [4], [10], [11], [12], [13], [14], [15], [16], [17], [18]. The research examines how the actual experience of one [11], [13], [14], [18], [19], [20] or more drives [21], [22] affects the development of acceptance and trust. The focus is often on takeover scenarios in SAE Level 3 automated vehicles, where responsibility for Object and Event Detection and Response (OEDR) switches from the ADS to the human driver [6], [19], [23]. Furthermore, the influence of acceptance and trust on driving behavior is currently under examination [6], [22]. Especially the engagement in NDRTs is analyzed based on drivers gaze behaviour [24].

Recent experiments are mainly conducted in simulators [6], [11], [14], [19], [20], [21], [22], [23], or utilizing a Wizard of Oz (WOz) approach [13], [18] on a test track. Furthermore, studies have employed the use of vehicle prototypes that are secured with additional safety drivers on the passenger seat [25] or in a manually driven vehicle situated behind them [24]. Other studies are restricted to examining the society's acceptance and trust, both from a drivers and an outside perspective, through online surveys without considering actual experience [1], [4], [10], [12].

Research findings reveal that experiencing automated driving firsthand through simulated rides increases both acceptance [21] and trust [23], but that driving without experiencing the system's limitations quickly leads to an overestimation of the system's capabilities and thus to careless behavior [7], [21]. Experiencing system limits, for example in the form of unexplained takeover requests, does however not lead to a decrease in acceptance [6]. Research indicates that individual factors, such as gender [1], age [11], [25], the general attitudes towards technological innovations [4], as well as people's feelings of ease [25] and their perceived usefulness [4], [25] influence acceptance and trust as well as the willingness to use ADS [25]. Another factor that affects acceptance and trust is the prevailing social norm. Inexperienced individuals may be particularly susceptible to the influence of the opinions of third parties in shaping acceptance [4]. Media reporting on accidents involving automated vehicles also has influence on the acceptance and trust of individuals [16]. Not least for that reason, according to the study by [10], one of the most common reasons for lack of acceptance is safety concerns. Furthermore, the velocity and driving behavior of the ADS, depending on traffic density, have an important influence on the acceptance of and trust in automated vehicles [11], [13], [18], which is, for example, reflected by the drivers' gaze behaviour and propensity to engage in NDRTs [24].

B. Aim of the Study

In the German traffic environment, until the end of 2022, only series systems up to SAE Level 2 automation had road approval. However, the Mercedes-Benz Drive Pilot is the first SAE Level 3 ADS in series production to be approved for German road use according to UN-R157 [26]. Prior, only the Honda SENSING Elite has had a limited SAE Level 3 approval for Japan [27]. This has changed the traffic environment and highlighted a critical point of interest for the field of ADS research. Current methods, such as driving simulators, WOz tests, where a hidden human controls part of the ADS that appears to be automated, or driving an unreleased ADS with a safety driver [25], provide informative insights on the interaction between humans and ADS. However, they do not fully encapsulate the human interaction with live ADS in a real-world traffic environment without any additional safety mechanisms such as safety drivers. In particular, studies conducted in simulators or using WOz concepts show significant variation in the factors that influence acceptance and trust, such as risk perception [4], driving style, and velocity [11], [13], [18] compared to actual driving in real traffic. Therefore, it is imperative to validate the results described above through real traffic studies [28]. The launch of the Drive Pilot enables researchers to conduct evaluations in the final Operational Design Domain (ODD) of the SAE Level 3 ADS, rather than in controlled, supervised or simulated environments. This provides a new opportunity to study human-automation interactions in the context of everyday use and real-world traffic conditions.

Therefore, this study will evaluate the development of acceptance and trust in the Level 3 Drive Pilot system, with a particular focus on drivers' first contact with Level 3 ADS. It will further address self-reported satisfaction and the emotional experience of using the system. During the study, the following research questions will be answered by testing the EQS with participants who have no prior driving experience with automated vehicles:

1) How does the first contact with an SAE Level 3 ADS in the form of driving in real traffic affect trust in automated driving systems?
2) How does the first contact with an SAE Level 3 ADS in the form of driving in real traffic affect the acceptance of automated driving systems?
3) How is usability and user experience rated after the first contact with an SAE Level 3 ADS?

II. Methods

The following delineates the method used to quantify the effect of first contact with ADS on the constructs of acceptance, trust, usability, and user experience. For this, the dependent variables, that will be quantified to describe the constructs, are outlined. For better readability and to better distinguish the constructs and quantified variables from the abstract concepts they represent, the **constructs** and quantified dependent variables are highlighted. Further, the study environment, procedure, and participants are described to give insight into how the quantified variables were obtained.

A. Dependent Variables and Questionnaires

Acceptance and **Trust** of the participants towards the ADS are evaluated as dependent variables. Additionally, **Usability** and **User Experience** are recorded.

To assess trust in automated systems, the German translation [29] of the Scale of Trust [30] was used. This questionnaire represents the current standard for evaluating **Trust** in the context of interaction with automated vehicles [28] and includes the subscales Trust and Mistrust as independent dimensions of trust. It consists of twelve items that participants rate on a 7-point Likert scale that ranges from 1 ("strongly disagree") to 7 ("strongly agree").

Acceptance is measured with the questionnaire according to [31], which is an adaptation of the UTAUT [32] and the CTAM [33] for automated vehicles. The questionnaire is suitable for evaluating a specific SAE Level 3 ADS and includes the scales of the Autonomous Vehicle Acceptance Model: Performance Expectancy, Effort Expectancy, Attitude Toward Technology, Social Influence, Facilitating Conditions, Self-Efficacy, Anxiety, Behavioral Intention, and Perceived Safety. The survey contains 26 items that are rated on a 7-point Likert scale from 1 ("strongly disagree") to 7 ("strongly agree"). Additionally, the importance of Hands, Feet, and Eyes as method of control for the ADS is assessed on a similar 7-point scale.

To investigate self-reported satisfaction, the study utilizes the System Usability Scale (SUS) as described in [28] and [34]. The SUS consists of 10 items, rated on a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). After analysis, the **System Usability** is displayed on a scale from 0 to 100. To also assess the emotional aspect of **User Experience** of system use, the User Experience Questionnaire (UEQ) [35] is assessed. This includes the six subscales Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty with a total of 26 opposing adjective pairs rated on a 7-point scale from -3 to 3.

Fig. 1. (Top) used vehicle for study (Mercedes-Benz EQS (V297), (bottom) visual Human-Machine-Interface for the SAE Level 3 Drive Pilot while in operation.

TABLE I
DRIVING PROCESS

Segment	Action	Length
A67 ↖	Approach from Darmstadt	25 km
A60 ←	ADS activation if possible	4 km
⤹	Turn around	1.5 km
A60 →	ADS activation if possible	4 km
A67 ↗	Construction site	3 km
↶	Turn around	1.5 km
A67 ↙	Approach	3 km
A60 ←	ADS activation if possible	4 km
⤹	Turn around	1.5 km
A60 →	ADS activation if possible	4 km
A67 ↘	Return to Darmstadt	25 km
		Σ 76.5 km

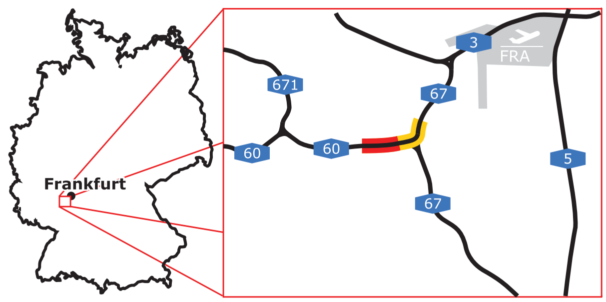


Fig. 2. Map of study route. Red designates the used section of highway for ADS activation, while yellow is the auxiliary section for creating a circular route. The respective ends of both sections have off- and on ramps for turning around.

Affinity for Technology and General Trust in technology are also recorded as confounding variables. Affinity for Technology is measured using the German ATI scale according to [36]. It consists of 9 items, which the participants rate on a 6-point Likert scale from 1 ("completely disagree") to 6 ("completely agree"). The three items "One should be careful with unknown automated systems", "I trust a system rather than mistrust it" and "Automated systems generally work well" from the questionnaire of [9] were used to measure General Trust in technology. These are rated on a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree").

B. Study Environment

For the study, the subject vehicle was a Mercedes-Benz EQS (V297), and the subject ADS was the enabled SAE Level 3 Drive Pilot. The Drive Pilot enables automated driving on German highways up to a speed of 60 km/h and outside construction zones, while another vehicle is driving in front. To enable activation, generally good environmental conditions have to be given. For a more detailed description of the ODD see [37]. Figure 1 displays both the EQS and the HMI of the Drive Pilot while in operation.

To maintain consistency throughout all trials, the driving within the study was carried out on a designated section of German highway during specific times of the day. Evaluations of the SAE Level 3 ADS were completed on the 4 km section of the A60 marked in red (see Table I and Fig. 2). This particular section was deemed optimal for the purposes of this study, as roadworks in the easterly direction on the section marked in yellow have been observed to reliably result in congestion, thus providing an environment conducive to the activation of the SAE Level 3 ADS. Furthermore, it was possible to demonstrate limitations of the system by driving into the roadworks. Afterwards, the yellow-highlighted portion was utilized for making directional changes through the succeeding ramp, permitting smooth re-entry into traffic flow for subsequent rounds. The full study round course encompasses the 17 km of the red and yellow highlighted sections, allowing for two complete runs of the evaluation section. Each participant completed the round course at least once, depending on whether it was possible to activate the ADS or not. The 25-kilometer route on the A67, originating from Darmstadt, was used to familiarize the participants with the vehicle.

The study was conducted between May 8th and October 27th, from Monday to Friday during two intervals from 06:00 to 10:00 and from 15:00 to 19:00. Each subject participated in either the first or second interval. The selected time frames matched commonly anticipated commuter traffic patterns, thus reflecting a realistic scenario in which end-users could engage an ADS (Fig. 3). The test drives were cancelled if rain was forecast for the entire interval. This scheduling permitted observing the performance of the ADS under varying levels of traffic density, which enhances the applicability of the study to everyday driving situations. The number and duration of sections that could be driven with an activated ADS varied as a consequence of the influence of different environmental conditions, including traffic density and the associated speed, as well as insufficient weather or lighting. On average, the ADS was activated slightly more than five times per participant. The average active duration of the ADS was 3.2 minutes, with a

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

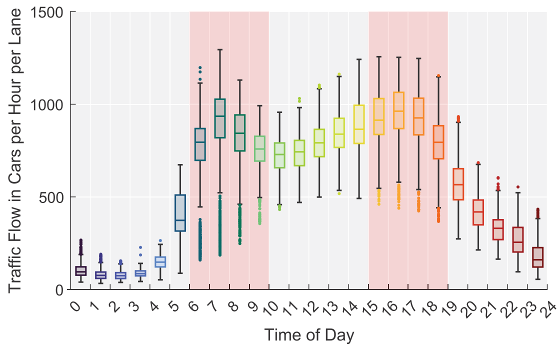


Fig. 3. Lane based hourly traffic flow for the Study route on A60. The used time intervals are marked in red. The data was aggregated from hourly measurements from the workdays of 2021 based on the Bundesanstalt für Straßenwesen (BASt) measurement sites *"6991 AD Rüsselsheim (W)"* and *"6993 AS Königstädten (W)"*. [38].

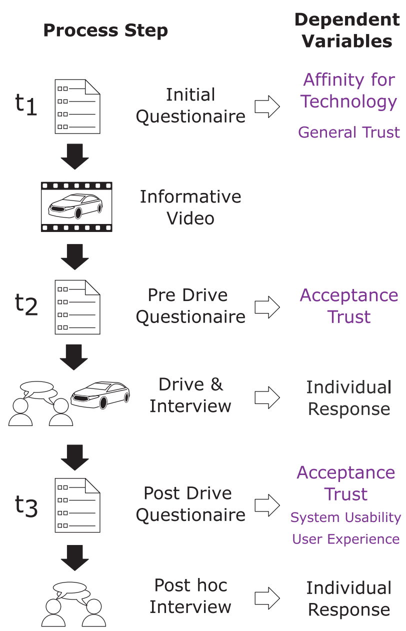


Fig. 4. Flowchart for study procedure.

maximum of 28 minutes and a minimum of approximately 30 seconds.

C. Procedure

The procedure of the study is shown in figure 4. Before driving, the participants were informed about the goal, content and the voluntary nature of the study. After consenting the participants' socio-demographic data, experience with SAE Level 1 and 2 ADS, Affinity for Technology and General Trust in technology, as well as prior knowledge of automated vehicles are recorded at time t_1. This is followed by a 2-minute video demonstrating the functionality, activation conditions, and handling of the SAE Level 3 ADS in the form of the Drive Pilot system of the Mercedes-Benz EQS as well as the required takeover behavior of the driver. Afterwards participants' questions about the system were answered. The video can be accessed via the following link: https://www.dom.daimler.com/.

To capture the impact of the first contact with the SAE Level 3 ADS on **Acceptance** and **Trust**, participants completed pre-journey (t_2) and post-journey (t_3) surveys regarding both dimensions (see II-A). During the test drive, participants were instructed to activate level 3 whenever the system indicated that activation was possible via the illuminated buttons on the left and right of the steering wheel. The activated system was indicated by the turquoise illuminated elements on the right and left of the steering wheel and a turquoise "A" in the instrument cluster display (see Fig. 1). Additionally, the participants were informed that the use of NDRT was permitted when the system was activated, though they were not explicitly asked to perform NDRT. This procedure was selected due to the absence of a safety driver and the aim to avoid any potentially dangerous situations. During the ride, the the subjects' reactions were filmed as well as documented by the study supervisor, sitting in the front passenger seat. After the ride, a post hoc interview was conducted to collect further qualitative data on the influence of the first contact with the ADS on **Acceptance** and **Trust**. The participants were asked about their experience while driving, their ability to understand the behavior of the ADS, the usefulness of the feature, their comfort in distracting themselves from the driving task, and whether the vehicle would increase road safety. Since usability and user experience can both only be evaluated properly after the SAE Level 3 ADS has been experienced, the subjects answered both the SUS and the UEQ at time t_3 after driving.

To ensure high-quality data, all participants were required to complete the questionnaires in a controlled and uniform manner at designated time points. For this purpose, the responsible investigators were provided with a comprehensive guideline that outlined the instructions for each time point and provided the wording of the questions and potential answers.

D. Participants

Participants were selected from an existing participant pool from institutes at TU Darmstadt, providing a foundational demographic. A key inclusion criterion for the study was participants' lack of professional experience with SAE Level 3 ADS. Experts with advanced knowledge and familiarity in this area were excluded, as their perspectives and behaviors may not accurately reflect those of the general public without experience. In addition, the participants were required to have no prior SAE Level 3 ADS driving experience. Thirty participants were selected to take part in the study. The age of the 21 males and 9 females ranged from 21 to 69 years with a mean age of 37.7 years (SD = 14.9). While most of them regularly use Cruise Control, the Lane Keeping Assist, Lane Departure Warning, Adaptive Cruise Control and Blind Spot Assists are rarely used. Figure 5 displays the characteristics of all 30 participants. As the figure further indicates, both the participants' General Trust in Technology (M = 3.57, SD = 0.52) as well as their Affinity for Technology (M = 4.43, SD = 0.94) were rather high. Although the selection criteria were stringent, it is acknowledged that the participant pool remains subject to bias. A majority of the participants came from the Rhine-Main Area, with a notable concentration from Darmstadt. Darmstadt, also known as the "city of science," is characterized by a distinguished academic background and a

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

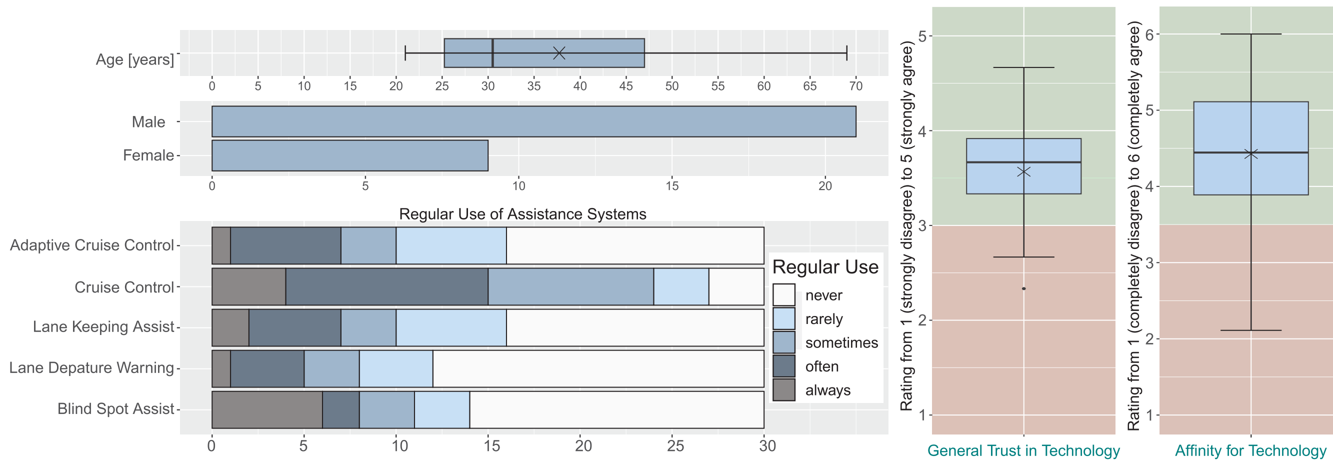


Fig. 5. Characteristics of the 30 participants: Age, Sex, Regular Use of Assistance Systems as well as General Trust in Technology and Affinity for Technology. The green background signifies positive responses and the red background negative responses towards ADS.

robust high-tech industrial focus. This aspect inevitably shapes the demographics and cultural inclinations of its residents. Notably, the study's sample is dominated by younger participants due to the disproportionately young demographic of the city.

III. Results

A. Questionnaire Evaluation and Statistical Analysis

For the evaluation of the **Acceptance** questionnaire according to [31], the mean values of the respective items were calculated for each participant for all twelve subscales at both time points t_1 and t_2. The subscales Performance Expectancy, Effort Expectancy, Social Influence, Facilitating Conditions, Attitude Towards Using Technology, Self-Efficacy, Anxiety, and Perceived Safety each include three items, while the subscale Behavioral Intention contains only two items. The subscales Method of Control for Hands, Feet, and Eyes each consist of only one item. For the evaluation of the **Trust** questionnaire according to [29], the mean values of the six trust items and the six mistrust items were calculated for each participant at both time points t_1 and t_2. The evaluation of the SUS was conducted according to the instructions by [34]. Accordingly, the responses were initially recoded. For the positive items, the response value was reduced by one. For the negative items, the response value was subtracted from five. Subsequently, the recoded values of all items were summed. The resulting raw score was then multiplied by 2.5 in order to calculate the SUS score, which ranges from 0 to 100. The result is illustrated in form of a box plot in Fig. 8. The evaluation of the UEQ followed the procedure described by [35]. Consequently, the mean value of all items associated with the subscales Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty was calculated for all test subjects. The results are illustrated in form of bar charts in Fig. 8. Outliers were not excluded from the analysis to prevent any further reduction in the validity of the results, given the relatively small number of participants. The statistical analysis was then conducted at the level of the subscales.

To statistically examine whether the initial contact with SAE Level-3 ADS in the form of a drive in real traffic has a significant impact on **Acceptance** and **Trust**, the difference values of the data from time points t_2 to time point t_3 were first tested for normal distribution. The Shapiro-Wilk test confirmed a normal distribution of the difference values for the subscales Trust (p = 0.165) and Mistrust (p = 0.447). In contrast, the difference values of the 12 subscales of **Acceptance** were not normally distributed.

Given the use of a standardized 7-point Likert scale for both questionnaires, tests for interval-scaled data can be applied to all data [39]. Consequently, the hypothesis that the initial contact with SAE Level-3 ADS in the form of a drive in real traffic has a significant impact on the subscales of **Acceptance**, the Wilcoxon signed-rank test was used due to the lack of normal distribution [40]. The effect size was calculated according to Cohen [41]. The medians, as well as the results of the Wilcoxon signed-rank test, including test statistics, p-values, and effect size calculations (Cohen's r), are presented in Table II and illustrated as box plots in the Fig. 6

To test the hypothesis that the initial contact with SAE Level-3 ADS in the form of a drive in real traffic has a significant impact on **Trust** paired-samples t-tests were used [40]. The effect size was subsequently calculated according to Cohen [41]. The means and standard deviations as well as the results of the t-tests, including test statistics, p-values, and effect size calculations (Cohen's d), are presented in Table III and illustrated as box plots in the Fig. 7

Since a total of 14 statistical tests were calculated, the alpha level was adjusted using the Bonferroni method to avoid Type I errors [42]. Based on an initial confidence interval of 95%, the corrected alpha level was calculated as follows:

$$\alpha_{\text{corrected}} = \frac{0.05}{14} = 0.0036$$

In the following, all results with $p < 0.0036$ are therefore assumed to be statistically significant. The statistical analyses were conducted using IBM SPSS Statistics software (Version: 28.0.1.1 (14)). The raw data can be requested via the following link: https://doi.org/10.48328/tudatalib-1487

B. Acceptance

Figure 6 displays the outcomes of the subscales of the **Acceptance** questionnaire [31]. These results indicate that

TABLE II

MEDIAN AND RESULTS OF THE STATISTICAL ANALYSIS (WILKOXON-TEST FOR PAIRED SAMPLES) FOR **ACCEPTANCE** MEASURED AT t_2 AND t_3 (* SIGNIFICANCE LEVEL < 0.0036, ** SIGNIFICANCE LEVEL < 0.001)

	Measuring Point	Mdn	Sum of positive Ranks	Sum of negative Ranks	Z	p	Cohen's r	effect size	N
Performance Expectancy	t2	4.33	435	0	-4.706**	0.0000	0.86	strong	30
	t3	6.5							
Effort Expectancy	t2	5.67	311.5	39.5	-3.467**	0.0002	0.63	strong	30
	t3	6.5							
Social Influence	t2	5.00	315.5	119.5	-2.124	0.0327	0.39	medium	30
	t3	5.67							
Facilitating Conditions	t2	5.67	301	24	-3.738**	0.0000	0.68	strong	30
	t3	6.33							
Attitude Towards Using Technology	t2	1.67	74.5	201.5	-1.937	0.0529	0.35	medium	30
	t3	1.33							
Self-Efficacy	t2	5.00	137	188	-0.688	0.5023	0.13	small	30
	t3	4.67							
Anxiety	t2	5.75	236.5	88.5	-2.02	0.0454	0.37	medium	30
	t3	6.00							
Behavioural Intention	t2	5.84	157.5	95.5	-1.009	0.3237	0.18	small	30
	t3	6.00							
Perceived Safety	t2	4.33	260	40	-3.17**	0.0007	0.58	strong	30
	t3	5.00							
Method of Control - Hands	t2	7.00	0	105	-3.33**	0.0001	0.61	strong	30
	t3	5.50							
Method of Control - Feet	t2	5.50	17.5	135.5	-2.82	0.0036	0.51	strong	30
	t3	3.00							
Method of Control - Eyes	t2	7.00	3.5	101.5	-3.109**	0.0009	0,57	strong	30
	t3	6.00							

TABLE III

MEAN, STANDARD DEVIATION AND RESULTS OF THE STATISTICAL ANALYSIS (T-TESTS FOR PAIRED SAMPLES) FOR **TRUST** MEASURED AT t_2 AND t_3 (* SIGNIFICANCE LEVEL < 0.0036, ** SIGNIFICANCE LEVEL < 0.001)

	Measuring Point	M	SD	T	df	p	Cohen's d	effect size
Trust	t2	4.84	1.08	-6.20 **	29.00	0.0000	1.13	strong
	t3	6.12	0.60					
Mistrust	t2	2.55	0.79	3.21 *	29.00	0.0032	0.89	strong
	t3	2.03	0.93					

initial exposure to driving with a SAE Level 3 ADS in real traffic significantly affects Performance Expectancy, Effort Expectancy, Facilitating Conditions as well as Perceived Safety and the importance of Hands and Eyes as methods of control for the ADS. Performance Expectancy increases significantly from $Mdn = 4.33$ prior to first contact to $Mdn = 6.5$ afterwards ($z = -4.706, p = 0.0000$, Cohen's $r = 0.86, n = 30$) with a strong effect according to [41]. Similarly, Effort Expectancy increases significantly from $Mdn = 5.67$ prior to first contact to $Mdn = 6.5$ afterwards ($z = -3.467, p = 0.0002$, Cohen's $r = 0.63, n = 30$), also indicating a strong effect. Facilitating Conditions exhibit a significant rise from $Mdn = 5.67$ prior to first contact to $Mdn = 6.33$ afterwards ($z = -3.738, p = 0.0000$, Cohen's $r = 0.93, n = 30$) with a medium effect and Perceived Safety increases significantly from $Mdn = 4.33$ prior to first contact to $Mdn = 5.0$ afterwards (z=-3.17, $p = 0.0007$, Cohen's $r = 0.58, n = 30$) with a strong effect. In contrast, the importance of Hands as methods of control decreases significantly from $Mdn = 7.0$ prior to first contact to $Mdn = 5.5$ afterwards ($z = -3.33, p = 0.0001$, Cohen's $r = 0.61, n = 30$) with a strong effect. Similarly, the importance of Eyes as methods of control decreases significantly from $Mdn = 7.0$ prior to first contact to $Mdn = 6.0$ afterwards ($z = -3.109, p = 0.0009$, Cohen's $r = 0.57, n = 30$) also showing a strong effect.

However, no significant affection of the initial exposure to driving with a SAE Level 3 ADS was observed for Social Influence, Attitude Towards Using Technology, Self-Efficacy, Anxiety, Behavioural Intention, and the importance of Feet as

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

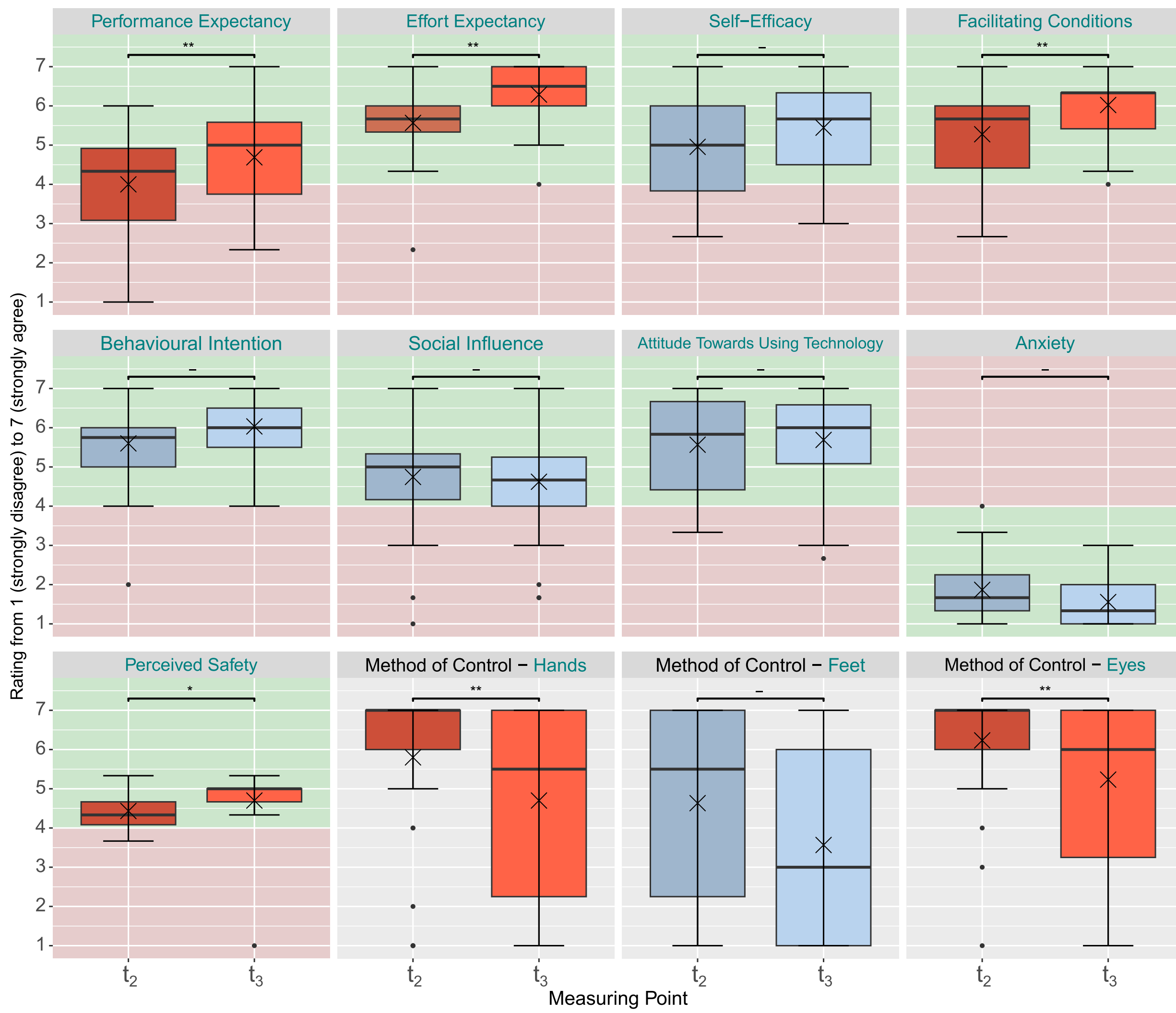


Fig. 6. Scales of **Acceptance** measured at t_2 and t_3 (* significance level < 0.0036, ** significance level < 0.001). Green background signifies positive and red background negative responses towards ADS. Blue box charts identify variables with no significant effect, while red identifies those with significant effect.

method of control (see Table II). Nevertheless, the data shows a tendency towards positive responses to AD for all subscales (see Fig. 6).

*C. **Trust***

As figure 7 shows, the mean level of Trust before initial contact was 4.84 ($SD = 1.08$) and increased to 6.12 ($SD = 0.6$) after the exposure. Meanwhile, the mean level of Mistrust before initial contact was 2.55 ($SD = 0.79$) and decreased to 2.03 ($SD = 0.93$) after the exposure. The results demonstrate that the first contact, in the form of driving in real traffic, has a significant effect on both Trust ($t(30) = -6.2$, $p < .001$, Cohen's $d = 1.13$) and Mistrust ($t(30) = -3.209$, $p = .01$, Cohen's $d = .887$) with a strong effect size for both according to [41].

*D. **Usability** and **User Experience***

The results of the System Usability Scale [34] presented in figure 8 show that **System Usability** is rated as good to excellent ($M = 84.33, SD = 13.52$) [43]. Figure 8 also presents the **User Experience** results of the UEQ [35], which exhibit good to very good results on all scales, particularly positive ratings for Attractiveness and Dependability.

E. Post Hoc Interviews

The post hoc interviews were analyzed according to the procedure for qualitative content analysis of [44]. In sum, all participants felt safe during the automated journey, although ten of them expressed feeling safe during the automated journey only after a short familiarization period. The majority of participants understood the ADS behavior and the driving style was deemed highly anticipatory. However, the reasons why the ADS deactivates itself or refused to activate again could not always be understood. For instance, participants who received the message "road section unsuitable" did not understand why it was unsuitable. Nine participants felt unable to disengage from the driving task because they believed that they needed to monitor the ADS, but assumed that it would be possible to disengage after a longer period of familiarization. The dominant opinion of the participants was that regular use of the ADS would increase trust and lead to the ability to turn

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

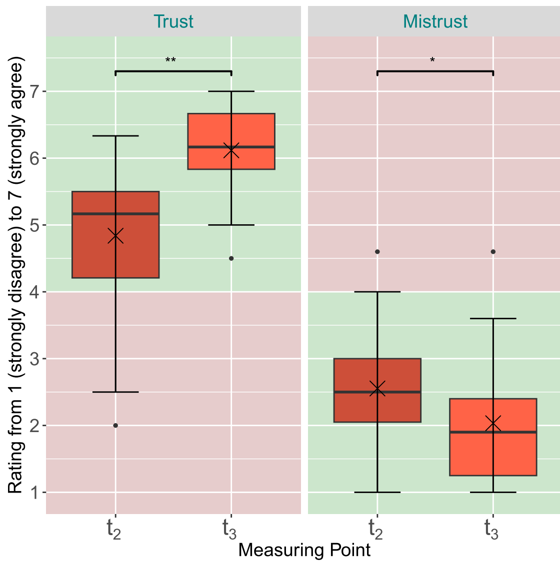


Fig. 7. Scales of **Trust** measured at t_2 and t_3 (* significance level < 0.0036, ** significance level < 0.001). Green background signifies positive and red background negative responses towards AD.

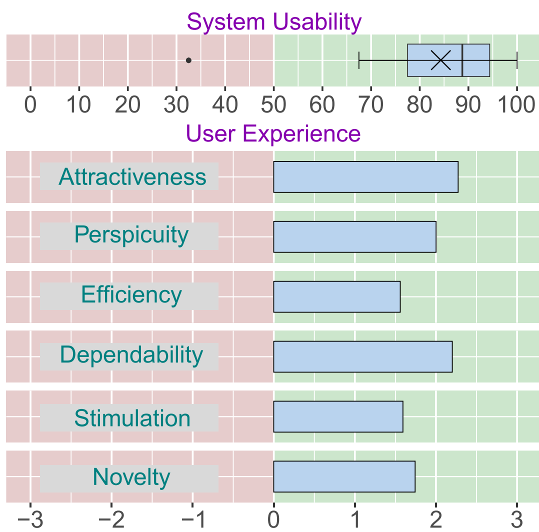


Fig. 8. (Top) **System Usability** measured at t_3 ($M = 84, 33$, $SD = 13, 52$). (Bottom) **User Experience** measured at t_3 (Attractiveness: $M = 2.27$, $SD = 0.56$; Perspicuity: $M = 2.0$, $SD = 0.9$; Efficiency: $M = 1.56$, $SD = 0.76$; Dependability: $M = 2.2$, $SD = 0.63$; Stimulation: $M = 1.59$, $SD = 0.73$; Novelty: $M = 1.74$, $SD = 0.77$). Green background signifies positive and red background negative responses towards AD.

away from the driving task more often. As a result, the entire driving experience would be more relaxed and restful. Overall, study subjects believed that ADS would enhance road safety due to their anticipatory driving style and ability to provide a better overview of traffic conditions compared to human drivers. However, five participants expressed concerns about the confusion between the SAE Level 2 and the SAE Level 3 ADS and the hazards that could be associated with such mode confusion when the SAE Level 3 ADS is deactivated. The ADS feature was deemed practical or even comfortable by 28 participants, although limitations such as low speed and early deactivation before road works were considered to be too restricting for regular use. On the other hand, 4 participants reported that using the ADS was just as strenuous as driving, as they had to keep an eye on the ADS and their surroundings.

IV. Discussion

In this research, the influence of the first contact with the approved SAE Level 3 ADS in the form of the Drive Pilot in the Mercedes-Benz EQS in real traffic situations on **Trust**, **Acceptance**, **Usability**, and **User Experience** was investigated. The study found a substantial increase in both **Acceptance** and **Trust** as well as a positive rating for **Usability** and **User Experience** after first contact with the SAE Level 3 ADS. This outcome aligns with the findings of the simulator studies presented in [21] and [23], indicating an overall positive initial situation towards ADS.

In particular, **Acceptance** of and **Trust** in the ADS in terms of fast, cost-effective, and safe destination reach, as well as simplicity of learning and ease of use, was rated significantly higher after the driving experience than expected before. The study subjects also rated their understanding and control of the vehicle, their knowledge of how to use the vehicle, and the feasibility of implementing the necessary technical and infrastructural requirements significantly higher after the driving experience than before. After the ride, they also felt significantly safer using the ADS and would consider its use to be much less dangerous. Similar the study by [6], no negative impact on **Trust** or **Acceptance** as a consequence of encountering the constraints or incomprehensible takeover requests could be identified. Since the study of [10] revealed that safety concerns before experiencing automated driving are a primary factor contributing to low acceptance, it can be postulated that the initial encounter with SAE Level 3 ADS could potentially address these concerns. Furthermore, similar to previous studies [7], [21], tendencies for careless behaviour or misuse were observed during SAE Level 3 ADS driving. For example, two participants attempted to change their seating position during automated driving, which may have inhibited or prevented them from quickly and adequately retaking the driving task. It was also observed that the study subjects neglected the driving task after using the SAE Level 3 ADS, even when using the SAE Level 2 ADS, and accordingly let go of the steering wheel, causing the SAE Level 2 ADS to issue a warning to retake it. Whether this was due to overestimation [7], [21] or rather due to mode confusion could not be conclusively determined.

Contrary to prior findings, this study could not establish a significant effect of age [11], [25], gender [1] or general attitude towards technology [4] on **Acceptance** and **Trust**. Nevertheless, it is worth mentioning that most participants were relatively young and tended to have a high Affinity for Technology as well as General Trust in Technology. It is hypothesized that individuals who have an aversion to technological innovation might be less inclined to take part in studies on ADS. Thus, it is necessary to consider whether a refusal to participate is due to a general mistrust of technological innovation.

The overall experiment was designed to allow the participants to experience the functionalities and limitations of the ADS in real-life situations mirroring German highway traffic conditions. Accordingly, a traffic jam was intentionally encountered in order to fulfill the activation requirements of the ADS, as well as a construction site in order to experience a further limitation in addition to velocity. However, due to the inherent variability of a real-world traffic study, it was not feasible to maintain consistent conditions for each test subject. The fluctuations in weather, light, and traffic density resulted in varying activation duration and frequencies, which represents a limitation of the study. Nevertheless, it was possible to complete a generalisable journey with an automated vehicle for each participant. Moreover, the study only surveyed **Acceptance** and **Trust** subjectively using questionnaires. Objective survey methods, such as eye tracking or the investigation of engagement with predefined NDRTs, for instance via comprehension tests, were intentionally avoided in order to prevent the occurrence of potentially dangerous situations.

For future studies, it is recommended that the sample of participants be diversified to include older individuals as well as those with a lower affinity for or trust in technology. Furthermore, research could examine how driving behavior adapts over longer or multiple trips in real traffic. This would be beneficial in understanding the nuances of driving behavior in varying contexts. Participants in the study mentioned that they would adapt their behavior after a longer period of familiarization, which is an important aspect to consider. In this context, the use of objective methods such as eye-tracking and the measurement of engagement in NDRT (cf. [24]) would be a valuable addition to the study, as it would provide insight that is not reliant on the participants' subjective perception. Furthermore, the investigation of the determination of driver readiness would be a beneficial avenue for exploration.

V. Conclusion

In summary, the findings of this study indicate that the initial experience of driving with a SAE Level 3 ADS vehicle was predominantly positive, leading to an increase in acceptance and trust of ADS as well as a positive Usability and User Experience. While participants acknowledge that the realization of automated functions represents a major achievement and has great potential, they also point out the need for further expansion of the limitations of the ADS to maximize the benefits for the driver, as well as to improve communication between the vehicle and the driver, thereby avoiding mode confusion and misuse. With regard to the introduction of SAE Level 3 vehicles, the results suggest that an iterative approach, whereby the ODD of the ADS is gradually expanded may have benefits over approaches, where the goal is to initially release the ADS with a large ODD. Early first contacts, enabled by incremental releases, have been shown to positively influence acceptance and trust.

Acknowledgment

The authors would like to thank all participants and especially Yannick Schwarz, Jonas Lehmann, and Meho Mahalbasic for their support in executing this study.

Declaration of Conflict of Interest

The author Prof. Peters was not involved in the development of the Drive Pilot during his prior employment with Mercedes-Benz Group AG. The EQS was ordinarily purchased by the Technical University of Darmstadt. Therefore the authors declare that there is no conflict of interest.

References

[1] K. Weigl, D. Eisele, and A. Riener, "Estimated years until the acceptance and adoption of automated vehicles and the willingness to pay for them in germany: Focus on age and gender," *Int. J. Transp. Sci. Technol.*, vol. 11, no. 2, pp. 216–228, Jun. 2022.

[2] E. Lehtonen, F. Malin, T. Louw, Y. M. Lee, T. Itkonen, and S. Innamaa, "Why would people want to travel more with automated cars?" *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 89, pp. 143–154, Aug. 2022.

[3] *Surface Vehicle Recommended Practice: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-road Motor Vehicles*, document J3016_202104, SAE Int., Apr. 2021.

[4] I. Nastjuk, B. Herrenkind, M. Marrone, A. B. Brendel, and L. M. Kolbe, "What drives the acceptance of autonomous driving? An investigation of acceptance factors from an end-user's perspective," *Technol. Forecasting Social Change*, vol. 161, Dec. 2020, Art. no. 120319.

[5] M. Beggiato and J. F. Krems, "The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 18, pp. 47–57, May 2013. https://www.sciencedirect.com/science/article/pii/S1369847813000028

[6] M. Körber, E. Baseler, and K. Bengler, "Introduction matters: Manipulating trust in automation and reliance in automated driving," *Appl. Ergonom.*, vol. 66, pp. 18–31, Jan. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003687017301606

[7] S. Nordhoff, J. D. Lee, S. Calvert, S. Berge, M. Hagenzieker, and R. Happee, "(Mis-)use of standard Autopilot and Full Self-Driving (FSD) Beta: Results from interviews with users of Tesla's FSD Beta," *Frontiers Psychol.*, vol. 14, 2024, doi: 10.3389/fpsyg.2023.1101520.

[8] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quart.*, vol. 13, no. 3, p. 319, Sep. 1989.

[9] M. Körber, "Theoretical considerations and development of a questionnaire to measure trust in automation," in *Proc. 20th Congr. Int. Ergonom. Assoc. (IEA)*, in Advances in Intelligent Systems and Computing, S. Bagnara, Ed. Cham, Switzerland: Springer, 2019, pp. 13–30. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-96074-6_2

[10] S. Cheng, H. Dong, Y. Yue, and H. Tan, "Automated driving: Acceptance and chances for young people," in *Cross-Cultural Design. Applications in Cultural Heritage, Tourism, Autonomous Vehicles, and Intelligent Agents*, vol. 12773, P.-L. P. Rau, Ed., Cham, Switzerland: Springer, 2021, pp. 182–194.

[11] S. Haghzare, J. L. Campos, K. Bak, and A. Mihailidis, "Older adults' acceptance of fully automated vehicles: Effects of exposure, driving style, age, and driving conditions," *Accident Anal. Prevention*, vol. 150, Feb. 2021, Art. no. 105919.

[12] J. Helgath, P. Braun, A. Pritschet, M. Schubert, P. Böhm, and D. Isemann, "Investigating the effect of different autonomy levels on user acceptance and user experience in self-driving cars with a VR driving simulator," in *Design, User Experience, and Usability: Users, Contexts and Case Studies*, vol. 10920, A. Marcus and W. Wang, Eds., Cham, Switzerland: Springer, 2018, pp. 247–256.

[13] M. Johansson, F. Ekman, H. Strömberg, M. Karlsson, and L.-O. Bligård, "Capable and considerate: Exploring the assigned attributes of an automated vehicle," *Transp. Res. Interdiscipl. Perspect.*, vol. 10, Jun. 2021, Art. no. 100383.

[14] C. Schmidt, F. Hartwich, and J. F. Krems, "Looking at driving automation from a passenger's perspective: Driving experience and gaze behavior during fully automated vs. human vehicle control," in *Human Interaction, Emerging Technologies and Future Applications III* (Advances in Intelligent Systems and Computing), vol. 1253. Cham, Switzerland: Springer, 2021, pp. 3–9.

[15] K. Weigl, C. Schartmüller, A. Riener, and M. Steinhauser, "Development of the questionnaire on the acceptance of automated driving (QAAD): Data-driven models for level 3 and level 5 automated driving," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 83, pp. 42–59, Nov. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847821002151

[16] M. Wicki, "How do familiarity and fatal accidents affect acceptance of self-driving vehicles?" *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 83, pp. 401–423, Nov. 2021.

[17] J. de Winter and S. Nordhoff, "Acceptance of conditionally automated cars: Just one factor?" *Transp. Res. Interdiscipl. Perspect.*, vol. 15, Sep. 2022, Art. no. 100645.

[18] S. Nordhoff, J. Stapel, B. van Arem, and R. Happee, "Passenger opinions of the perceived safety and interaction with automated shuttles: A test ride study with 'hidden' safety steward," *Transp. Res. A, Policy Pract.*, vol. 138, pp. 508–524, Aug. 2020.

[19] J. Dillmann, R. J. R. den Hartigh, C. M. Kurpiers, F. K. Raisch, D. de Waard, and R. F. A. Cox, "Keeping the driver in the loop in conditionally automated driving: A perception-action theory approach," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 79, pp. 49–62, May 2021.

[20] F. Naujoks, Y. Forster, K. Wiedemann, and A. Neukum, "Improving usefulness of automated driving by lowering primary task interference through HMI design," *J. Adv. Transp.*, vol. 2017, pp. 1–12, Aug. 2017.

[21] B. Metz, J. Wörle, M. Hanig, M. Schmitt, A. Lutz, and A. Neukum, "Repeated usage of a motorway automated driving function: Automation level and behavioural adaption," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 81, pp. 82–100, Aug. 2021.

[22] N. Zangi, R. Srour-Zreik, D. Ridel, H. Chassidim, and A. Borowsky, "Driver distraction and its effects on partially automated driving performance: A driving simulator study among young-experienced drivers," *Accident Anal. Prevention*, vol. 166, Mar. 2022, Art. no. 106565.

[23] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, "Trust in automation—Before and after the experience of take-over scenarios in a highly automated vehicle," *Proc. Manuf.*, vol. 3, pp. 3025–3032, Jan. 2015.

[24] X. Liu, R. Madigan, E. Sadraei, Y. M. Lee, and N. Merat, "Drivers' engagement in NDRTs during automated driving linked to travelling speed and surrounding traffic," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 101, pp. 332–339, Feb. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847824000160

[25] Y. M. Lee, R. Madigan, T. Louw, E. Lehtonen, and N. Merat, "Does users' experience and evaluation of level 3 automated driving functions predict willingness to use: Results from an on-road study," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 99, pp. 473–484, Nov. 2023.

[26] *Un Regulation, no. 157—Uniform Provisions Concerning the Approval of Vehicles With Regards to Automated Lane Keeping Systems [2021/389]*, document R157, 2021.

[27] Honda. (2021). *Initiatives in the Area of Automated Driving Technologies*. [Online]. Available: https://global.honda/en/tech/Automated_drive_safety_and_driver_assistive_technologies/

[28] A.-K. Frison, Y. Forster, P. Wintersberger, V. Geisel, and A. Riener, "Where we come from and where we are going: A systematic review of human factors research in driving automation," *Appl. Sci.*, vol. 10, no. 24, p. 8914, Dec. 2020.

[29] G. Pöhler, T. Heine, and B. Deml, "Itemanalyse und faktorstruktur eines fragebogens zur messung von vertrauen im umgang mit automatischen systemen," *Zeitschrift für Arbeitswissenschaft*, vol. 70, no. 3, pp. 151–160, Nov. 2016.

[30] M. R. Lehto, T. House, and J. D. Papastavrou, "Foundations for an empirically determined scale of trust in automated systems," *Int. J. Cogn. Ergonom.*, vol. 4, no. 1, pp. 73–86, Mar. 2000.

[31] C. Hewitt, I. Politis, T. Amanatidis, and A. Sarkar, "Assessing public perception of self-driving cars: The autonomous vehicle acceptance model," in *Proc. 24th Int. Conf. Intell. User Interfaces*, Mar. 2019, pp. 518–527.

[32] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quart.*, vol. 27, no. 3, p. 425, 2003.

[33] S. Osswald, D. Wurhofer, S. Trösterer, E. Beck, and M. Tscheligi, "Predicting information technology usage in the car," in *Proc. 4th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, Oct. 2012, pp. 51–58.

[34] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds., London, U.K.: Taylor & Francis, 1996, pp. 189–194.

[35] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *HCI and Usability for Education and Work* (Lecture Notes in Computer Science). Berlin, Germany: Springer, Nov. 2008, pp. 63–76. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-540-89350-9_6

[36] T. Franke, C. Attig, and D. Wessel, "A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale," *Int. J. Human–Computer Interact.*, vol. 35, no. 6, pp. 456–467, Apr. 2019.

[37] Mercedes-Benz Group. (2023). *Mercedes-Benz Drive Pilot*. [Online]. Available: https://www.mercedes-benz.de/passengercars/technology/drive-pilot.html

[38] Bundesanstalt für Straßenwesen. (2021). *Automatische Zählstellen 2021*. [Online]. Available: https://www.bast.de/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/Daten/2021_1/Jawe2021.html

[39] N. Döring, *Forschungsmethoden Und Evaluation in Den Sozial- und Humanwissenschaften*, 6th ed., Berlin, Germany: Springer, 2023. [Online]. Available: http://www.springer.com/

[40] D. Vorberg and S. Blankenberger, "Die auswahl statistischer tests und Maße," *Psychologische Rundschau*, vol. 50, no. 3, pp. 157–164, Jul. 1999.

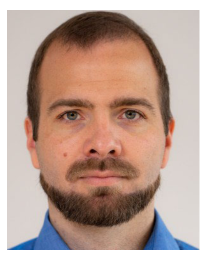
[41] J. Cohen, "Statistical power analysis," *Current Directions Psychol. Sci.*, vol. 1, no. 3, pp. 98–101, 1992.

[42] J. Bortz and C. Schuster, *Statistik Für Human- und Sozialwissenschaftler*, 7th ed., Berlin, Germany: Springer, 2010. [Online]. Available: http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10448295

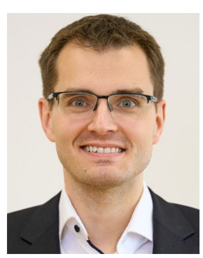
[43] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," *Int. J. Human-Computer Interact.*, vol. 24, no. 6, pp. 574–594, Jul. 2008.

[44] P. Mayring, *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, 13th ed., Weinheim, Germany: Beltz, 2022. [Online]. Available: https://www.content-select.com/index.php?id=bib_view&ean=9783407258991

Sarah Schwindt-Drews received the M.Sc. degree in mechanical and process engineering from the Technical University of Darmstadt in 2020. She is currently pursuing the Ph.D. degree. Since 2020, she has been a Research Associate with the Institute for Ergonomics and Human Factors, Technical University of Darmstadt. Her research interests center around the mental models of drivers in automated vehicles.



Kai Storms received the M.Sc. degree in mechanical and process engineering from the Technical University of Darmstadt, Germany, in 2020, and the Ph.D. (Dr.-Ing.) degree in February 2024. He is currently a Research Associate with the Technical University of Darmstadt. His topic was context aware data reduction for highly automated driving. His research interests include verification and validation of automated vehicles, with a focus on data reduction.



Steven Peters was born in 1987. He received the Ph.D. (Dr.-Ing.) from Karlsruhe Institute of Technology, Karlsruhe, Baden-Württemberg, Germany, in 2013. From 2016 to 2022, he was the Manager of Artificial Intelligence Research with Mercedes-Benz AG, Germany. He has been a Full Professor with the Technical University of Darmstadt, Darmstadt, Germany, and heads the Institute of Automotive Engineering, Department of Mechanical Engineering, since 2022.

Bettina Abendroth received the Ph.D. degree in industrial engineering, majoring in mechanical engineering from the TU Darmstadt University of Technology in July 2001. Her topic was driver types of vehicles and driving support at longitudinal guiding. She is currently the Deputy Director of the Institute of Ergonomics and Human Factors and heads the Human-Machine Interaction & Mobility Research Group. She was appointed as a Professor (apl.) in 2024.