

# Methoden für die Bewertung der Mensch- Maschine-Interaktion beim teilautomatisierten Fahren

Berichte der  
Bundesanstalt für Straßenwesen

Fahrzeugtechnik Heft F 141

**bast**

# Methoden für die Bewertung der Mensch- Maschine-Interaktion beim teilautomatisierten Fahren

von

Nadja Schömig  
Katharina Wiedemann  
Ruth Julier  
Alexandra Neukum

Würzburger Institut für Verkehrswissenschaften  
WIVW GmbH

André Wiggerich  
Heike Hoffmann

Bundesanstalt für Straßenwesen

**Berichte der  
Bundesanstalt für Straßenwesen**

Fahrzeugtechnik Heft F 141

**bast**

Die Bundesanstalt für Straßenwesen veröffentlicht ihre Arbeits- und Forschungsergebnisse in der Schriftenreihe **Berichte der Bundesanstalt für Straßenwesen**. Die Reihe besteht aus folgenden Unterreihen:

- A - Allgemeines
- B - Brücken- und Ingenieurbau
- F - Fahrzeugtechnik
- M - Mensch und Sicherheit
- S - Straßenbau
- V - Verkehrstechnik

Es wird darauf hingewiesen, dass die unter dem Namen der Verfasser veröffentlichten Berichte nicht in jedem Fall die Ansicht des Herausgebers wiedergeben.

Nachdruck und photomechanische Wiedergabe, auch auszugsweise, nur mit Genehmigung der Bundesanstalt für Straßenwesen, Stabsstelle Presse und Kommunikation.

Die Hefte der Schriftenreihe **Berichte der Bundesanstalt für Straßenwesen** können direkt bei der Carl Ed. Schünemann KG, Zweite Schlachtpforte 7, D-28195 Bremen, Telefon: (04 21) 3 69 03 - 53, bezogen werden.

Über die Forschungsergebnisse und ihre Veröffentlichungen wird in der Regel in Kurzform im Informationsdienst **Forschung kompakt** berichtet. Dieser Dienst wird kostenlos angeboten; Interessenten wenden sich bitte an die Bundesanstalt für Straßenwesen, Stabsstelle Presse und Kommunikation.

Die **Berichte der Bundesanstalt für Straßenwesen (BASt)** stehen zum Teil als kostenfreier Download im elektronischen BASt-Archiv ELBA zur Verfügung.  
<https://bast.opus.hbz-nrw.de>

## **Impressum**

**Bericht zum Forschungsprojekt 82.0708+82.0709**  
Methoden für die Bewertung der  
Mensch-Maschine-Interaktion  
beim teilautomatisierten Fahren

### **Fachbetreuung**

André Wiggerich  
Heike Hoffmann

### **Referat**

Automatisiertes Fahren

### **Herausgeber**

Bundesanstalt für Straßenwesen  
Brüderstraße 53, D-51427 Bergisch Gladbach  
Telefon: (0 22 04) 43 - 0

### **Redaktion**

Stabsstelle Presse und Kommunikation

### **Druck und Verlag**

Fachverlag NW in der  
Carl Ed. Schünemann KG  
Zweite Schlachtpforte 7, D-28195 Bremen  
Telefon: (04 21) 3 69 03 - 53  
Telefax: (04 21) 3 69 03 - 48  
[www.schuenemann-verlag.de](http://www.schuenemann-verlag.de)

ISSN 0943-9307

ISBN 978-3-95606-591-0

Bergisch Gladbach, Juni 2021

## Kurzfassung – Abstract

### **Methoden für die Bewertung der Mensch-Maschine-Interaktion beim teilautomatisierten Fahren**

Der Bericht stellt zwei standardisierte methodische Ansätze zur Bewertung der Effizienz und Sicherheit der Mensch-Maschine-Interaktion (MMI) bei der Nutzung von teilautomatisierten Fahrfunktionen vor. Dazu wurden Prüfkriterien definiert, welche die Erfüllung der dazu notwendigen Anforderungen an die Vermittlung eines angemessenen Systemwissens, eines adäquaten System- und Situationsbewusstseins, einer wenig beanspruchenden Systembedienung und einer sicheren Fahrerreaktion an Systemgrenzen erfassen sollen.

Mittels einer Expertenbewertung als erste Methode wird geprüft, ob die Mensch-Maschine-Schnittstelle (MMS, d. h. die Anzeigen, Systemausgaben und die Bedienelemente) die Mindestanforderungen bezüglich Wahrnehmbarkeit und Verständlichkeit von Systemzuständen erfüllt. Der Fokus liegt dabei auf visuellen und akustischen Systemausgaben. Zusätzlich werden Aspekte der Vermittlung eines angemessenen Systemverständnisses über das Benutzerhandbuch sowie der Systembedienung bewertet. Zur Überprüfung dieser Aspekte wurde eine Checkliste mit insgesamt 14 Anforderungen entwickelt. Das Verfahren kann auf Teststrecken oder im Realverkehr angewendet werden. Dazu wird ein Set an Use Cases durchfahren, anhand derer die Erfüllung der einzelnen Items abgeprüft wird. Die Checkliste wurde im Rahmen von zwei Realfahrtstudien und einer Fahrsimulatorstudie erprobt und weiterentwickelt. Dabei stand die Überprüfung der Validität (d. h. kann die Checkliste unterschiedlich gute Mensch-Maschine-Interaktionskonzepte voneinander unterscheiden) sowie der Reliabilität (d. h. inwieweit kommen unterschiedliche Rater bei der Bewertung eines Systems auf vergleichbare Ergebnisse) im Vordergrund.

Als zweite Methode wurde ein standardisiertes, szenarienbasiertes Beobachtungs- und Befragungsinstrument zur Bewertung der Systeminteraktion entwickelt. Als Bewertungskriterien werden Probleme bei der Systembedienung, im Fahrverhalten sowie Einschränkungen des Überwachungsverhaltens erfasst, die in ein übergeordnetes Versuchsleiterring pro Use Case münden. Darüber hinaus werden Aspekte des Systemverständnisses über

eine Befragung des Fahrers zu einzelnen MMS-Anzeigen und Systemreaktionen erfasst. Diese Bewertungskategorien sind in anschaulicher Form in einer Tablet-Anwendung, der sog. S.A.D.E.-App (Standardized Application for Automated Driving Evaluation), umgesetzt. Die Methode wurde in zwei Simulatorstudien und während einer Fahrt im Realverkehr mit einem Seriensystem erprobt und validiert.

Für eine umfassende, standardisierte Bewertung der Mensch-Maschine-Interaktion beim teilautomatisierten Fahren wird empfohlen, beide Methoden in Kombination zu verwenden. Das Verfahren kann um zusätzliche Methodenbausteine zur Erfassung der langfristigen Entwicklung von Systemvertrauen und -akzeptanz sowie Verhaltensanpassungen erweitert werden.

### **User studies for the evaluation of human-machine-interaction with partial automated driving functions**

The report presents two standardized methodological approaches for evaluating the efficiency and safety of human-machine interaction in the use of partially automated driving functions. For this purpose, test criteria were defined to assess the fulfilment of the necessary requirements for conveying adequate system knowledge, adequate system and situation awareness, easy system operation and safe driver reaction at system boundaries.

As a first method, an expert evaluation is used to check whether the human-machine interface (HMI, i.e. displays, system outputs and operating elements) meets the minimum requirements regarding the perceptibility and comprehensibility of system states. The focus is on visual and acoustic system outputs. In addition, aspects of conveying an appropriate understanding of the system via the user manual and system operation are evaluated. A checklist with a total of 14 requirements was developed to check these aspects. The procedure can be applied in a driving simulator, on test tracks or in real traffic. For this purpose, a set of use cases is experienced, on the basis of which the fulfilment of the individual items is checked. The checklist was tested and further developed in two real-life driving



studies and one driving simulator study. The focus was on evaluating the validity (i.e. can the checklist distinguish between human-machine interaction concepts of different quality?) and the reliability (i.e. to what extent do different experts arrive at comparable results when evaluating a system?)

As a second method, a standardized, scenario-based observation and survey tool for the evaluation of system interaction in the context of user studies was developed. As evaluation criteria, problems with system operation, driving behaviour and inadequacy of the monitoring behaviour are recorded, which lead to a superordinate examiner rating per use case. In addition, aspects of system understanding are recorded by questioning the driver about HMI-indicators and system reactions. These evaluation categories are implemented in a tablet application, the so-called S.A.D.E.-App (Standardized Application for Automated Driving Evaluation). The method was tested and validated in two simulator studies and under real traffic conditions with a series production vehicle.

For a comprehensive, standardized evaluation of the human-machine interaction during partially automated driving, it is recommended to use both methods in combination. The procedure can be extended by additional methods to evaluate the long-term development of system trust and acceptance as well as behavioural adaptations.

## Summary

### User studies for the evaluation of human-machine-interaction with partial automated driving functions

### Objective

By now many vehicle manufacturers offer vehicles with partial automation driving functions (so called Level 2-systems) with continuously automated longitudinal and lateral vehicle control. In such systems, the driver remains responsible for supervising traffic and the automated system.

Empirical research in this field shows that there is a relatively high level of acceptance and trust in such systems. The driver perceives a relief at least in some driving situations, such as driving in traffic jams. In general, however, the monitoring task is perceived as exhaustive. During the usage of these systems drivers tend to direct their attention to other non-driving-related tasks. Contradictory statements exist regarding the ability of drivers to react to system limits in Level-2-systems. It seems less problematic in cases where system limits are expectable, full attention is at the road and a warning is given. However, for most of the situations in L2-systems this cannot be guaranteed. Drivers must also be able to immediately intervene in sudden, critical events. There are initial indications that an appropriate design of the Human-Machine-Interface (HMI) might support drivers in this task.

### Overall evaluation concept

To date, there are no standardized tools for the assessment of the usability and safety of partial automated driving functions. The goal of the project was to develop approaches in order to evaluate whether drivers are able to adequately deal with such vehicles.

On the one hand, it must be evaluated whether the human-machine-interface (HMI) is able to support driver's understanding of system functionality and a safe management of system limits. This requires the definition of minimal requirements on the design of HMI for automated systems. Furthermore, in the

direct interaction with the system it must be checked to what extent the driver is able to understand the presented information and use them safely and effectively. It must be assured that the driver reaches a sufficient level of system, mode and situation awareness and is not overloaded due to the interaction with the system. For these issues, it is necessary to directly observe the user's interaction with the system and the HMI. For the assessment of the HMI of partial automated driving functions two methods are proposed that should be used in combination. The subordinate goal is to develop a comprehensive, standardized evaluation concept for HMI in automated driving which contains all relevant evaluation tools and criteria for a safe and efficient human-machine-interaction.

The first tool is the heuristic evaluation of the HMI performed by experts. By means of a checklist, experts rate whether the minimal requirements on the HMI regarding perceptibility and understandability of system states are fulfilled. This is mainly focused on the design of the outputs of the HMI. In addition, elements of system usage and the ability of the user manual to create an adequate system understanding are included as evaluation items. The checklist method was developed in a project running in parallel to this project and is documented in a separate report.

The second method proposed here are user studies in which the interaction-related aspects of a safe human-machine-interaction of partial automated driving functions are investigated. Beside the evaluation of HMI-aspects also effects of the system design can be assessed.

With this background a scenario-based, standardized assessment tool was developed in this project. By means of a tablet-application it is possible to assess observation and survey variables that can be used to derive aspects such as system understanding and system awareness, situation awareness, workload during the use of the system and the ability for a safe and efficient behavior in case of system limits and system errors. The method is applied in relevant use cases during user studies in the driving simulation or in real vehicles.

### Demand on user studies

As precondition of the usage of the tablet-based evaluation tool some basic, technical demands for

user studies in the driving simulation were defined. They refer to the relevant characteristics of a prototypic L2-system and its HMI-elements that must be implemented in the simulation. These characteristics were derived from a benchmark-analysis of existing serial L2-systems on the market. They include the implementation of the following features: an automated speed-and distance keeping function, an automated lateral control function, a lane-change assistance function, a hands-on detection in the steering wheel, a system logic on the basis of a system state diagram and defined transitions between the system states (by which actions or events and under which conditions does a transition from one to another state happen?), a HMI-logic including the respective visual display elements and a system operation logic including the respective control elements. Some parameters should be freely configurable in order to vary them in the context of an empirical study such as the way of operating the system, the hands-off strategy, the presence of warnings in case of system limits, the quality of lateral control and the frequency of failures in lateral control (based on “simulated” sensor problems).

The test scenarios for the evaluation of the systems were derived from a use case catalogue. This defines use cases as all possible steady system states (driving in a certain automation level or specific system state, such as change of system settings or degraded system systems, e.g. due to temporary failures in lateral control) and all transitions between these system states. The transitions are further categorized according to whether they are initiated by the system or the driver and whether they are directed upwards (towards higher automation levels) or downwards (towards lower automation levels). Based on the system logic of the evaluated system and the respective research question test scenarios are derived which contain a specification of the concrete human-machine-interaction in a certain environmental condition. A prototypic test layout created from the defined requirements was implemented both in the driving simulator of the Würzburg Institute for Traffic Sciences (WIVW GmbH) and in the simulator of the Federal Highway Research Institute (BAST).

## Development of the standardized evaluation tool

For the evaluation of interaction-related aspects of the safe human-machine-interaction of continuous L2 automated driving functions the following relevant evaluation criteria were formulated: First, aspects of the driver behavior which are assessed by observation of the driver in the interaction with the system via the human-machine-interface (HMI). Several errors/problems in three areas of system operation, driving behavior and monitoring behavior are defined which can separately be coded by the test administrator. Based on predefined rules, a generic rating about the ability of handling the use case is then derived from the observed problems in each scenario. Furthermore aspects of system understanding should be assessed by a specific survey to single HMI notifications and system reactions (e.g. how clear was the system operation? How understandable were HMI outputs, how comprehensible was the system's behavior?).

Those evaluation criteria were implemented in the so called tablet-application S.A.D.E. The acronym stands for Standardized Application for Automated Driving Evaluations and is an application of the driving simulation software SILAB® for the conduction of driver observations and driver surveys in partial automated driving on Level 2 via a tablet-PC. The test administrator can use the application for recording predefined error categories in system operation (e.g. delayed, inadequate, operating error, support necessary), driving behavior (e.g. delayed reaction, inadequate securing behavior, endangerment, collision, lane exceedance, unprecise lane keeping) and monitoring behavior (insecure hands-on behavior, inadequate inattention, occurrence of hands-off warnings) via the touch interface, can rate the managing of the scenario as a whole and can conduct a standardized questioning of the subject during a test drive. By the clearly arranged and vivid graphical presentation of the error categories and the questions including the rating scales as well as by the partially automated counting of driving errors for certain categories, the test administrator is supported in the observation and objectivity of the procedure is increased.

The application provides a data output file in a text-format, which contains the data for each variable per scenario. The usage and the graphical user interface as well as the procedure of the test is

described in detail in the project report. For an optimal training in the application of the tool training materials are available. The tablet-app and its evaluation criteria were explored and validated within two simulator studies with different research questions.

## **Application of the evaluation tool in a driving simulator study at the WIVW facility**

The first user study was conducted in the dynamic driving simulator of the WIVW GmbH. The objective was to evaluate whether HMI design issues defined from a previous expert evaluation using a checklist are also reflected in the behavior and experience of real users. For this purpose, two HMI variants were designed that differed in major design aspects of the HMI, e.g. regarding the system operation logic (way of activating ACC, location and labeling of control elements), the visual display elements (way of displaying the active lateral control and the lane change, contrast between foreground and background), and the warning concept (presence of a visual and acoustic warning in situations with detected system limits). These differences were deliberately designed that they result in different ratings on the developed checklist items. One "high compliant" HMI-variant A and one "low-compliant" HMI-variant B resulted from that procedure. The test course in the simulation had a length of 36 km and consisted of the following use case set: system activation, longer driving with active L2-system (several times during the route), driver-initiated lane change, temporary failures in lateral control due to missing lane markings, system failure at a broken-down vehicle (a stationary vehicle is not detected by the sensors, system stays active, driver has to intervene) and a system limit (system is not able to manage a sharp bend, lateral control is deactivated), system deactivation by driver. The test course was on a four lane highway (two lanes in each direction). The target speed was 120 km/h. Drivers should stay on the right lane with exceptions for the lane change and the avoidance maneuver at the stationary vehicle. On the left lane there was moderate traffic. In the scenarios lane change and stationary object there was no other traffic present to create comparable situations for all subjects. The complete duration of the drive was 20 minutes.

The two designed HMI-variants were varied as a between-subjects factor in the study. A second between-subject factor was the sequence of the two critical situations "sharp bend" and "stationary vehicle". In the two drive orders, these scenarios were changed while the sequence of the other scenarios remained constant.

N = 24 drivers participated in the study. They were recruited from the WIVW test driver panel. 11 from them were female. Mean age was 41.2 years (SD = 14.3 years). N = 12 drivers each experienced the HMI-variant A and B. Both groups were comparable with regard to characteristics such as age, driving performance, and prior experience with assistance systems and automated systems (from other simulator studies).

The study procedure was the following: after a short manual introduction drive the subjects read a short user manual which was adapted to the two HMI-variants. Then, they received general instructions for the simulator drive and the usage of the rating scales. This was followed by the test drive in the driving simulator where subjects experienced all relevant use cases and where immediately asked after each use case. The test administrator recorded observed problems during system operation, driving behavior and monitoring behavior by using the tablet-app for which they had received training before. After the test drive subjects had to fill out a questionnaire regarding different aspects about the HMI derived from the items of the checklist. The total duration of the study was 1 hour.

Beside the tablet-data also objective driving data were analyzed for the two critical scenarios.

The study showed that the designed differences in the two HMI-variants did also affect driver's behavior and subjective experiences of the system. The warning strategy in case of the system limit at the sharp bend had an influence. The absence of a visual and acoustic warning resulted in a worse lane keeping behavior and a worse system understanding. The differences in system operation resulted in a higher frequency of problems, such as a more frequent insecure or delayed activation behavior or incomplete activation which required a support by the test administrator. The variation in the distinctness of system states which were influenced by the visual contrast and the used indicators for the active lateral control function were also perceived as different by the two HMI-

variants although they had no effect on the driving performance (at least in the non-critical temporary failures in lateral control). Altogether however, due to the small sample size most effects only reached a tendency in significance.

The global rating of the test administrator per scenario mainly reflected the subordinate categories which makes it very suitable to reach a fast and efficient global impression of the system. The fact that the identified differences in the observed driving behavior on the dichotomous categories correspond with the analysis of time-based objective data (e.g. more frequent events of lane exceedances → higher maximum lateral deviation values, more frequent delayed braking reactions and endangerment → smaller minimum time-to-collision values) make the latter which are more time and resource-intensive obsolete. This makes the evaluation tool especially useful for the application in the real vehicle which does not provide data about driving performance.

In contrast to expectations the results showed a difference in the two experimental conditions with a worse driver behavior for the HMI-variant A in the scenario stationary vehicle although the system behavior and the HMI outputs were identical for the two HMI-variants. This effect was independent from scenario sequence or sample characteristics such as age, gender, driving performance or prior experience with assistance and automated systems. At this point, it can only be speculated what the reasons for this difference are. Probably the experience of a “good” HMI led to a higher system trust and to the impression of a higher system performance. As a result subjects might have waited longer before they realized that the system would not manage the situation. However, the data of the study do not provide any more evidence for that hypothesis.

Beside the described content-related results also some methodological issues arose that can be used for further improvements of the tool. Some recommendations were made by the test administrator regarding the integration of single observation categories into one. In general, there was a low need for corrections of the error recordings and the test administrator had no problems in creating the global ratings based on the single error categories and the predefined rules. This makes the method and the unidimensional rating scale a very effective tool.

## Application of the evaluation tool in a driving simulator study at the BAST facility

A second user study had the objective to investigate the effects of different system behavior characteristics, especially the quality of lane keeping performance, on the behavior and experience of a user. In the present report only the methodology of the study is described. The conduction and results of the study will be documented in the final public BAST report by the authors of the study.

The user study was conducted in the static driving simulator of BAST. In advance a suitable control element, a hands-on detection and a second camera in order to observe the driver’s inputs at the steering wheel had been implemented. The standardized evaluation tool S.A.D.E. was implemented on a tablet-PC and integrated in the network. As partial automated driving system the same one was used as in the WIVW study. It combines automated longitudinal and lateral vehicle control. The system is operated via buttons at the steering wheel.

For answering the research question the quality of lane keeping assistance was varied on two dimensions: the extent of lane centering of the lateral control (perfect vs. within a range) and the frequency of “simulated” sensor problems (no vs. frequent). These two factors were combined to four system variants and varied as between-subject-factor.

The driving course consisted of two 15-minutes sections in which the subjects could drive with the automated system on Level 2, each condition with its system variant of lateral control. After these longer sections a critical scenario followed. One test scenario was the system limit at the sharp bend where the lateral control was deactivated and the driver had to steer to avoid leaving the lane. The other test scenario was the stationary object which was announced by a warning sign at the road shortly before. The driver had to notice that the system will not detect the stationary vehicle and will stay active so that they had to intervene by steering and/or braking (both scenarios were identical to those in the WIVW study). Online driver behavior and survey were conducted by the S.A.D.E.-App.

The results of the study will show whether aspects of system behavior, such as the design of the lateral

control function will affect driver monitoring behavior and the resulting take-over ability at system limits.

## Application of the method in real traffic

For the application of the method in real traffic the app was reconfigured so that it can also be used flexible in a less controlled environment as it is the case in the real world. Some observation categories were summarized, some were deleted. The graphical user interface was adapted so that the name of the current use case was placed in a prominent location and that it included the instruction for the subjects how to behave in respective use cases. With regard to the operation of the app the test administrator can now actively switch to a new use case and can also deviate from the given use case sequence by using a forward and backward function.

The app was tested in a real drive with the autopilot function integrated in a Tesla Model 3. It helped to give useful recommendations for the usage of the tool and the test procedure in real traffic. Furthermore, suggestions for the choice of a suitable on-road test course are given.

## Final evaluation of the method

In summary, the developed instrument is considered as a valuable tool for the evaluation of the human-machine-interaction in partial automated driving within the scope of user studies. It can be flexibly applied in various test environments and provides an efficient data entry and analysis. The testing procedure, the test scenarios and the evaluation criteria are standardized. The ratings of the observed behavior and experiences of the drivers are based on defined guidelines and therefore objective and normative.

In general, user studies have the advantage that not only hypothetical problems can be identified but also the real problems occurring in the direct interaction of a real user with the system. It allows the assessment of both the intuitive behavior in the first contact with the system and learning or sequence effects. By using realistic test scenarios the effects on driving safety can be estimated.

Potential constraints of the method mainly refer to the component of questioning the subjects. It can be expected that users are not very good in rating the relevant single HMI elements objectively and in the expected accuracy. For such design aspects it seems more appropriate to use the expert-based heuristic evaluation method.

Further constraints refer to the application of the method in real traffic. Due to safety reasons it is not possible to test potentially dangerous situations such as unpredictable system limits or system failures and their effects on driving behavior (in the sense of controllability testing). In addition, it is problematic that it is not possible to guarantee controlled and comparable conditions for all subjects experiencing the system. Due the named reason it is recommended to better use the method on test tracks.

Altogether, the user study and the checklist tool are recommended as two methods that should be used in combination. Those can be further expanded by other method-elements, e.g. a structured interview for the evaluation of the mental model developed during the usage of the system and naturalistic driving studies for the assessment of longterm-effects such as the developments of system trust, acceptance, intentions to use and behavioral adaptations.



## Inhalt

<b>1</b>	<b>Herleitung der Methoden zur Bewertung von teilautomatisierten Fahrfunktionen</b> . . . . .	13	2.5.2	Erprobung im Rahmen einer Simulatorstudie . . . . .	39
1.1	Definition und Beschreibung teilautomatisierter Fahrfunktionen . . . . .	13	2.5.3	Finale Erprobung im Realverkehr. . . . .	44
1.2	Potenzielle Problembereiche des teilautomatisierten Fahrens. . . . .	15	2.6	Gesamtbewertung und Limitationen des Verfahrens . . . . .	45
1.3	Bewertungsmethoden: Expertenbewertung und Nutzerstudie . . . . .	18	<b>3</b>	<b>Szenarienbasiertes Beobachtungsinstrument für die Bewertung der Mensch-Maschine-Interaktion in Nutzerstudien</b> . . . . .	46
1.4	Definition von Anforderungen und Prüfkriterien . . . . .	19	3.1	Anforderungen an das Bewertungsinstrument. . . . .	46
1.5	Definition der zu testenden Use Cases . . . . .	22	3.2	Definition der Bewertungskategorien . . . . .	48
1.6	Übersicht Methodenbeschreibung . . . . .	24	3.2.1	Beobachtungskategorien . . . . .	48
<b>2</b>	<b>Checkliste für die Expertenbewertung der Mensch-Maschine-Schnittstelle</b> . . . . .	25	3.2.2	Befragungskategorien . . . . .	49
2.1	Aufbau der Checkliste . . . . .	25	3.2.3	Gesamturteil . . . . .	50
2.2	Definition der Bewertungskategorien und -prinzipien . . . . .	27	3.3	Umsetzung der Methode in eine Tablet-Applikation. . . . .	50
2.3	Itemherleitung. . . . .	27	3.3.1	App-Oberfläche . . . . .	51
2.3.1	Anforderung an das Benutzerhandbuch . . . . .	28	3.3.2	Anwendung der App in der Fahrsimulation . . . . .	52
2.3.2	Anforderungen an die Ausgaben der Mensch-Maschine-Schnittstelle . . . . .	31	3.3.3	Anwendung der App im Realverkehr. . . . .	54
2.3.3	Anforderungen an die Systembedienung. . . . .	34	3.4	Prüfszenarien . . . . .	55
2.3.4	Anforderungen an das System. . . . .	35	3.4.1	Definition von Prüfszenarien für Simulatorstudien. . . . .	55
2.4	Anwendung der Checkliste. . . . .	36	3.4.2	Auswahl einer geeigneten Prüfstrecke im Realverkehr. . . . .	56
2.4.1	Anforderungen an die Rater. . . . .	36	3.5	Erprobung der Bewertungsmethode . . . . .	57
2.4.2	Zu prüfende Assistenzfunktion und MMS-Elemente. . . . .	36	3.5.1	Nutzerstudie zum Vergleich verschiedener MMS-Konzepte im WIVW-Simulator . . . . .	57
2.4.3	Testumgebung . . . . .	37	3.5.2	Nutzerstudie zur Auswirkung der Spurhaltegüte im BAST-Simulator. . . . .	65
2.4.4	Benötigte Ausrüstung . . . . .	37	3.5.3	Erprobung im Realverkehr . . . . .	71
2.4.5	Testablauf. . . . .	37	3.6	Gesamtbewertung und Limitationen des Verfahrens . . . . .	72
2.5	Erprobung und Anpassung der Checkliste. . . . .	38	<b>4</b>	<b>Fazit Gesamtbewertungskonzept</b> . . . . .	74
2.5.1	Erste Erprobung im Realverkehr . . . . .	38			



<b>Literatur</b> .....	77
<b>Bilder</b> .....	82
<b>Tabellen</b> .....	83
<b>Anhang</b> .....	84

# 1 Herleitung der Methoden zur Bewertung von teilautomatisierten Fahrfunktionen

## 1.1 Definition und Beschreibung teilautomatisierter Fahrfunktionen

Viele Hersteller bieten mittlerweile Fahrzeuge mit teilautomatisierten Fahrfunktionen (sog. Level-2-Systeme oder kurz L2-Systeme, gemäß SAE-Definition, 2013, 2016, 2018) an, bei denen die Fahrzeuge die kontinuierliche Längs- und Querführung übernehmen. Gemäß der Klassifikation von Automatisierungsstufen nach SAE ist der Fahrer<sup>1</sup> auf dieser Automationsstufe weiterhin für die Überwachung der Verkehrsumgebung verantwortlich und

muss bei Systemgrenzen oder -fehlern jederzeit in der Lage sein, korrigierend in die Fahrzeugführung einzugreifen (siehe Bild 1-1 mit Klassifikation der Automatisierungsstufen der SAE, 2018; für ähnliche Kategorisierungen siehe GASSER et al., 2012; NHTSA, 2017).

Die Verantwortlichkeiten des Fahrers beim teilautomatisierten Fahren können demnach wie in Tabelle 1-1 zusammengefasst werden. Die Aufgabe des Fahrers ist in der SAE-Definition als OEDR (Object and event detection and response) bezeichnet und meint die Objekt- und Eventerkennung, deren Klassifikation, die Verhaltensplanung und schließlich die Verhaltensreaktion selbst.

In Abgrenzung zum sog. teilautomatisierten Fahren, das Gegenstand dieses Berichts ist (im vorliegenden Bericht auch Fahren mit Level-2-System oder L2-System/-Fahrzeug genannt), das aus der Kombination einer kontinuierlichen Querführungsassistenz (z. B. einem Spurhalteassistenten) mit einer kontinuierlichen Längsführungsassistenz (z. B. Ad-

<sup>1</sup> Aus Gründen der einfachen Lesbarkeit wird im Text das generische Maskulinum verwendet, die Angaben beziehen sich jedoch auf alle Geschlechter.

		SAE J3016™ LEVELS OF DRIVING AUTOMATION					
		SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?		You are driving whenever these driver support features are engaged - even if your feet are off the pedals and you are not steering			You are <b>not</b> driving when these automated driving features are engaged - even if you are seated in "the driver's seat"		
		You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
What do these features do?		These are driver support features			These are automated driving features		
		These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met	This feature can drive the vehicle under all conditions	
Example Features		<ul style="list-style-type: none"> <li>• automatic emergency braking</li> <li>• blind spot warning</li> <li>• lane departure warning</li> </ul>	<ul style="list-style-type: none"> <li>• lane centering OR</li> <li>• adaptive cruise control</li> </ul>	<ul style="list-style-type: none"> <li>• lane centering AND</li> <li>• adaptive cruise control at the same time</li> </ul>	<ul style="list-style-type: none"> <li>• traffic jam chauffeur</li> </ul>	<ul style="list-style-type: none"> <li>• local driverless taxi</li> <li>• pedals/steering wheel may or may not be installed</li> </ul>	<ul style="list-style-type: none"> <li>• same as level 4, but feature can drive everywhere in all conditions</li> </ul>
		For a more complete description, please download a free copy of SAE J3016: <a href="https://www.sae.org/standards/content/J3016_201806/">https://www.sae.org/standards/content/J3016_201806/</a>					

Bild 1-1: Automatisierungsstufen nach SAE (2016) – Rot markiert ist Level 2, die sog. Teilautomation (partial driving automation)

L2 Verantwortlichkeiten	
Fahrer (kontinuierlich)	Automatisiertes System (wenn aktiv)
<ul style="list-style-type: none"> <li>• Führt die Teilaufgaben der dynamischen Fahraufgabe aus, die nicht vom automatisierten System ausgeführt werden.</li> </ul>	<ul style="list-style-type: none"> <li>• Führt Teile der Fahraufgabe aus, indem es die laterale (Querführung) und longitudinale (Längsführung) Fahrzeugkontrolle übernimmt.</li> </ul>
<ul style="list-style-type: none"> <li>• Überwacht das automatisierte System und greift notfalls ein, um die sichere Fahrzeugführung zu gewährleisten.</li> </ul>	<ul style="list-style-type: none"> <li>• Deaktiviert sich sofort nach Eingriff durch den Fahrer.</li> </ul>
<ul style="list-style-type: none"> <li>• Entscheidet selbstständig, ob/wann das System aktiviert vs. deaktiviert werden sollte.</li> </ul>	
<ul style="list-style-type: none"> <li>• Übernimmt sofort die komplette Fahraufgabe, wenn erforderlich oder gewünscht.</li> </ul>	

Tab. 1-1: Verantwortlichkeiten des Fahrers und der Automation in L2-Systemen gemäß SAE (übersetzt aus SAE, 2018, Seite 21)

aptive Cruise Control ACC) besteht und vom Fahrer weiterhin die Überwachungstätigkeit erfordert, sind weitere sog. Automationsstufen nach SAE zu unterscheiden, die sich im Grad der Unterstützung in der kontinuierlichen Fahraufgabe sowie in der Notwendigkeit zur Überwachung der Verkehrsumgebung unterscheiden (siehe Bild 1-1).

- Das sog. manuelle Fahren (ohne kontinuierliche Assistenzsysteme, aber evtl. mit temporär eingreifenden Notfallsystemen, wie dem Notbremsassistenten; im vorliegenden Bericht nach SAE auch Level 0 oder L0 genannt)
- Das sog. assistierte Fahren (d. h. entweder nur kontinuierliche Querführungsassistenz oder nur kontinuierliche Längsführungsassistenz; im vorliegenden Bericht nach SAE auch Fahren mit Level-1-System oder L1-System genannt)
- Das sog. hochautomatisierte Fahren, bei dem der Fahrer von der kontinuierlichen Überwachungsaufgabe befreit ist, aber weiterhin als Rückfallebene im Fall einer Systemgrenze zur Verfügung stehen muss (im vorliegenden Bericht nach SAE auch Fahren mit Level-3-System oder L3-System genannt)
- Das sog. vollautomatisierte bzw. autonome Fahren, bei dem der Fahrer nicht mehr als Rückfallebene benötigt wird (im vorliegenden Bericht nach SAE auch Fahren mit L4-/L5-System oder L4 bzw. L5 genannt). L4- und L5-Systeme unterscheiden sich dahingehend, dass für L5-Systeme keinerlei Beschränkungen vorliegen, während die Nutzung von L4-Systemen bestimmten einschränkenden Bedingungen unterliegt (z. B. bestimmte Straßentypen).

Bei einer Analyse der wesentlichen Systemcharakteristika von L2-Serienfahrzeugen (Stand: Frühjahr

2018; analysiert wurden Volvo XC60 PILOT ASSIST, Tesla Model S Lenkassistent, BMW Driving Assistant Plus, Cadillac CT6 Supercruise, Mercedes E-Klasse Lenk-Pilot, Audi Q7 Stauassistent, VW Passat Stauassistent) zeigt sich, dass in den meisten Systemen die Längs- und die Querführung als zwei voneinander getrennte Systeme betrachtet werden, die gemeinsam genutzt werden können, aber nicht müssen. Häufig ist die Aktivierung des ACC-Systems (Adaptive Cruise Control) eine Voraussetzung, um die Querführung aktivieren zu können. Entsprechend können die Systeme auch über verschiedene Bedienelemente – in der Regel über Tasten am Lenkrad – aktiviert und deaktiviert werden. In einigen Systemen ist die Querführung standardmäßig aktiv und muss im Menü explizit deaktiviert werden, wenn sie nicht genutzt werden soll. Bei allen Systemen auf dem europäischen Markt handelt es sich um sog. Hands-on-Systeme, d. h. der Fahrer ist aufgefordert, seine Hände während der Systemnutzung am Lenkrad zu lassen. Tut er dies nicht, wird nach einer bestimmten Zeit eine sog. Hands-Off-Warnung ausgegeben (siehe UN/ECE R 79, 2010 für eine genaue Spezifikation der visuellen und akustischen Ausgaben). An Systemgrenzen ist häufig ein sofortiger Fahrereingriff erforderlich, wenn sich das System entweder unmittelbar abschaltet (z. B. in sehr scharfen Kurven) oder das System in Situationen, in denen beispielweise aufgrund von Sensorproblemen Hindernisse nicht erkannt werden, trotzdem aktiv bleibt. Bis auf wenige Ausnahmen geben die meisten Systeme an Systemgrenzen keine explizite Warnung aus. In Fällen, in denen die Sensorik keine Fahrstreifenmarkierungen erkennt, wechselt die Querführungsassistenz häufig still, d. h. ohne eine Rückmeldung an den Fahrer in den Stand-by-Zustand, und aktiviert sich selbstständig wieder, sobald die Voraussetzungen wieder erfüllt sind.

Die Mensch-Maschine-Schnittstelle (MMS; d. h. die Anzeigen, Systemausgaben und Bedienelemente), unterscheidet sich ebenfalls zwischen den Systemen. Die Anzeige eines Lenkrads als Symbol für die Querführungsassistenten (Visualisierung der verschiedenen Systemzustände, d. h. aktiv vs. Standby, durch unterschiedliche Farbigkeit) wird weitgehend in allen Systemen verwendet. Zum Teil findet eine Ergänzung durch die Visualisierung der erkannten Fahrstreifenmarkierung statt. Die Aktivität der Längsregelung wird häufig durch Querbalken angezeigt, die gleichzeitig den eingestellten Folgeabstand zu Vorderfahrzeugen kodieren. Ein vorausfahrendes Fahrzeug wird eingeblendet, wenn dieses erkannt, oder wenn auf dieses geregelt wird. Neben der automatischen Quer- und Längsführung bieten einige Hersteller auch eine integrierte Spurwechselassistenten an, die allerdings immer vom Fahrer initiiert und überwacht werden muss. In der Regel wird diese Assistenten durch Betätigung des Blinkers (z. B. Blinkerhebel muss länger gedrückt gehalten werden) aktiviert. Daraufhin unterstützt das System den Fahrer bei der Lenkung des Fahrzeugs auf den benachbarten Zielfahrstreifen.

## 1.2 Potenzielle Problembereiche des teilautomatisierten Fahrens

Teilautomatisierte Fahrfunktionen sollen den Fahrkomfort erhöhen, bringen aus Sicht der Human-Factors-Forschung jedoch spezielle Herausforderungen mit sich, da dem Fahrer auf der einen Seite Teile der Fahraufgabe abgenommen werden, er jedoch weiterhin in der Verantwortung der Fahrzeugführung bleibt und jederzeit bereit sein muss, in zeitkritischen Situationen einzugreifen.

In den oft zitierten „Ironies of Automation“ beschrieb LISANNE BAINBRIDGE (1983) schon in den 80er Jahren, dass der Mensch kein guter Überwacher von Automationssystemen sei. Dies gilt besonders dann, wenn ihm in einem Mensch-Maschine-System die Aufgabe zufällt, plötzlich eingreifen zu müssen, wenn die Automation an ihre Grenzen stößt. GREENLEE, DELUCIA & NEWTON (2018) fanden heraus, dass das Fahren mit einem teilautomatisierten System zu einer Vigilanzminderung beiträgt, die zu einer Verschlechterung der Entdeckungsrate von potenziellen Gefahren mit zunehmender Fahrdauer und zunehmender Anforderung an die Überwachungsleistung führt. Zudem waren die subjektiv wahrgenommene Beanspruchung und das Stress-

level erhöht. BIONDI, LOHANI, HOPMAN, MILLS, COOPER & STRAYER (2018) konnten während einer Realfahrt mit einem teilautomatisierten Fahrzeug eine verringerte physiologische Aktivierung der Fahrer sowie eine verschlechterte Reaktion auf eine periphere visuelle Entdeckungsaufgabe (PDT) messen, was sie als Out-of-the-Loop Phänomen interpretierten.

Hinsichtlich der Reaktion auf Systemgrenzen gibt es unterschiedliche Befunde dahingehend, wie gut sie vom Fahrer gelöst werden können. In Notbremssituationen beispielsweise besteht im Vergleich zum manuellen Fahren und assistierten Fahren mit ACC die Gefahr, dass Fahrer mit teilautomatisierten Systemen verzögert reagieren (STANTON, YOUNG, WALKER, TURNER & RANDLE, 2001). Auch bei SCHALLER, SCHIEHLEN & GRADENEGGER (2008) zeigten sich verzögerte Reaktionen in kritischen Situationen im Vergleich zum manuellen Fahren, jedoch kam es mit Systemwarnung nie zu Kollisionen. LARGE, BANKS, BURNETT, BAVERSTOCK & SKRYPCHUK (2017) fanden beim teilautomatisierten Fahren ein geringes Maß an subjektiv wahrgenommenem Situationsbewusstsein sowie eine Beeinträchtigung der Leistung in der primären Fahraufgabe (Spurhaltung) und der sekundären Nebenaufgabe unmittelbar nach einer akustischen Übernahmesituation. De WINTER, STANTON, PRICE & MISTRY (2016) konfrontierten Fahrer auf verschiedenen Automationsstufen (von manuell bis hochautomatisiert) mit Situationen, die eine Fahrerreaktion erforderten. Auf besonders auffällige Reize (Stoppschilder, Fußgänger oder bremsendes Vorderfahrzeug) reagierten die Fahrer rechtzeitig, gerade bei geringen Geschwindigkeiten hatten sie aber häufiger Probleme, auf einen unvorhersehbaren Systemausfall zu reagieren. Generell führte eine fehlerbehaftete Automation zu hoher Frustration, während der physische Aufwand und die Beanspruchung sich verringerten. KIRCHER, LARSSON & HULTGREN (2014) berichten hingegen, dass Fahrer in Situationen, die das System erwartungsgemäß nicht selbst bewältigen konnte, rechtzeitig die Kontrolle übernahmen. Auch NAUJOKS, PURUCKER & NEUKUM (2016) fanden in Realfahrtstudien ebenfalls keine Beeinträchtigung der subjektiven und objektiven Sicherheit beim assistierten Fahren mit ACC vs. L2-System (im Vergleich zum manuellen Fahren). KLEEN (2014) konnte in einer Realfahrtstudie zeigen, dass die Fahrer beim Fahren mit einem L2-System sehr gut und verhältnismäßig zügig auf eine nicht zurückge-

meldete Deaktivierung der Querverführung, zum Beispiel an einer Baustelleneinfahrt oder in Kurvenszenarien, mit einem Lenkeingriff reagierten. Auch subjektiv bewerteten die Fahrer die Situationen als gut beherrschbar. Allerdings geben die Autoren zu bedenken, dass die Fahrer sich bewusst waren, an einer Untersuchung teilzunehmen, was gegenüber Alltagsfahrten eine höhere Aufmerksamkeit erwarten lässt.

Studien zeigen ebenfalls, dass Fahrer dazu neigen, sich vermehrt Nebentätigkeiten zuzuwenden und die Überwachung der Fahraufgabe zu vernachlässigen, wenn sie durch die entstehende Monotonie aufgrund der Wegnahme der Fahrtätigkeit unterfordert sind (LLANERAS, SALINGER & GREEN, 2013). GASPARG & CARNEY (2019) untersuchten in einer Naturalistic Driving Studie das Blickverhalten von Fahrern beim teilautomatisierten Fahren. Dabei beobachteten sie, dass Fahrer eher bereit waren, einzelne, besonders lange Blickabwendungen von der Straße zuzulassen und ihre Aufmerksamkeit insgesamt länger auf Bereiche außerhalb der Fahrbahn richteten als ohne Teilautomation. LOUW, KUO, ROMANO, RADHAKRISHNAN, LENNÉ & MERAT (2019) konfrontierten Fahrer insgesamt sechsmal in einer Simulationsstudie mit sog. „stillen“ Systemausfällen (d. h. ohne Rückmeldung an den Fahrer) auf Geraden und in Kurven und stellten dabei fest, dass die Fahrer, wenn sie durch eine visuelle Nebenaufgabe zusätzlich abgelenkt waren, zwar alle Ausfälle bemerkten und auf diese reagierten, dabei aber verzögert übernahmen und häufiger den Fahrstreifen verließen.

Darüber hinaus scheint die Gestaltung der Systemanzeigen wesentlich dazu beizutragen, inwieweit diese als zusätzliche Ablenkung wirken. KRAFT, NAUJOKS, WÖRLE & NEUKUM (2018) konnten zeigen, dass ein kritischer Grad der Ablenkung durch die Systemanzeigen bei einer angemessenen, d. h. reduzierten Gestaltung der MMS verhindert werden kann. WULF, ZEEB, RIMINI-DÖRING, ARNON & GAUTERIN (2013) untersuchten, ob spezifische Strategien der Mensch-Maschine-Interaktion das Situationsbewusstsein des Fahrers und dessen Fahrsicherheit verbessern können. Vorteilhaft in Bezug auf ein erhöhtes Situationsbewusstsein erwies sich ein Videobild, über das der Fahrer während der Beschäftigung mit einer visuellen Nebenaufgabe die Fahrszenarie peripher beobachten konnte. Allerdings hatte diese Variante vernachlässigbare Auswirkungen auf die Fahrsicherheit und wurde von den Autoren durch die längeren Blick-

dauern auf das Display gar als potenziell sicherheitsgefährdend in unerwarteten kritischen Situationen beurteilt. Ein sog. „Driver Safety Device“ (eine Art Totmann-Schalter, der nach Ablauf einer vorgegebenen Zeit regelmäßig gedrückt werden musste) führte im Vergleich zu einer schnelleren Reaktionsfähigkeit (weil bereits eine Hand am Lenkrad gehalten wurde), hatte aber keine positiven Auswirkungen auf das Situationsbewusstsein (gemessen über das SAGAT-Verfahren, siehe ENDSLEY, 1988).

Eine weitere Gefahr des teilautomatisierten Fahrens besteht in der potenziellen Verwechslung mit höheren Automationsstufen. Endkonsumenten sind sich der immer noch währenden Verantwortung für die Fahraufgabe bei der Nutzung dieser neuen Technologien möglicherweise nicht bewusst, falls diese nicht eindeutig als Fahrerunterstützung gekennzeichnet sind. Hersteller wie Tesla, die ihre teilautomatisierte Fahrzeugfunktion unter dem Namen „Autopilot“ vermarkten, könnten dem Benutzer – etwa aufgrund eines subjektiven Vergleichs mit Systemen aus der Luftfahrt – ein falsches Verständnis über das tatsächliche Unterstützungsniveau des Systems und die eigene Verantwortlichkeit suggerieren (IIHS, 2019) und ein übermäßiges Vertrauen in dessen Leistungsfähigkeit bewirken (siehe beispielsweise der Unfall eines Tesla mit einem Truck im Mai 2016 in Florida; NTSB-Bericht, 2017). In den Medien wurde in der Vergangenheit bereits von kritischen Fahrsituationen und Unfällen mit teilautomatisierten Funktionen berichtet, die mit einem Fehlverständnis der Fähigkeiten der Automation in Zusammenhang zu stehen scheinen (BANKS, PLANT & STANTON, 2018). Wissenschaftliche Untersuchungen stützen diese Annahme (LLANERAS, SALINGER & GREEN, 2013; ABRAHAM, SEPPELT, MEHLER & REIMER, 2017; NAUJOKS, PURUCKER, WIEDEMANN, NEUKUM, WOLTER & STEIGER, 2017). Aus Sicht potenzieller Nutzer teilautomatisierter Fahrfunktionen kommt für das Systemverständnis erschwerend hinzu, dass unter anderem in den Medien verschiedene Bezeichnungen für solche L2-Funktionen kommuniziert werden. So werden die Begriffe „autonom“, „selbstfahrend“ und „teilautomatisiert“ mitunter für dieselben teilautomatisierten Funktionen verwendet. Auch in der Wissenschaft gab es lange keine einheitliche Definition, obwohl die Forschung zur Mensch-Maschine-Interaktion (MMI) im Kontext des automatisierten Fahrens bis in die 90er Jahre des letzten Jahrhunderts zurückgeht (z. B. DESMOND, HANCOCK & MONETTE, 1998). Zudem ist davon auszugehen, dass

die wissenschaftliche Definition und Abgrenzung von Automationsstufen, wie z. B. die Klassifikation nach SAE (siehe Bild 1-1) den Verbrauchern nicht geläufig ist. Eine Einweisung oder Erklärung der in Fahrzeugen vorhandenen Assistenzsysteme findet außerdem nicht systematisch statt (ABRAHAM, MCANULTY, MEHLER & REIMER, 2018; CASNER & HUTCHINS, 2019), obwohl sich in der Literatur gezeigt hat, dass die Vorerfahrung und das Vorwissen über die automatisierte Fahrfunktion das Systemverständnis und die Art und Weise, wie Fahrer mit der Automation interagieren, beeinflussen können (FORSTER, HERGETH, NAUJOKS, KREMS & KEINATH, 2019b; HERGETH, LORENZ & KREMS, 2017; GOLD, KÖRBER, HOHENBERGER; LECHNER & BENGLER, 2015; KÖRBER; BASELER & BENGLER, 2018; BEGGIATO & KREMS, 2013; ZHOU, ITOH & KITAZAKI, 2019).

Dabei kann die selbst erworbene Vorerfahrung mit ähnlichen Systemen und Assistenzfunktionen eine Rolle spielen (WRIGHT, SAMUEL, BOROWSKI, ZILBERSTEIN & FISHER, 2016). Dies ist auch deshalb relevant, da vielen Fahrern oft nicht bewusst ist, über welche Funktionen sie in ihrem oder einem geliehenen Fahrzeug verfügen, und sie deren Funktionsweise durch Erfahrung im Umgang mit den Funktionen erlernen (VIKTOROVÁ & SUCHA, 2018). Die Erfahrung von Systemgrenzen während der Nutzung und die vorhandene Systemauslegung (bspw. maximal möglicher Lenkwinkel oder Querbesehleunigung) des Systems kann das mentale Modell entscheidend beeinflussen (PRICE, VENKATRAMAN, GIBSON, LEE & MUTLU, 2016). Häufigere Systemausfälle führen beispielsweise zu stärkerem Überwachungsverhalten und früherem Eingreifen im Falle einer Systemgrenze (SCHWARZ, GASPAR & BROWN, 2019). Das Vertrauen in die Funktion und damit die adäquate Verwendung kann sich ebenfalls auf das Eingreifen im Falle einer Systemgrenze auswirken (ZEEB, BUCHNER & SCHRAUF, 2015).

Im Kontrast zu den genannten Einschränkungen zeigt die empirische Forschung mit L2-Systemen jedoch ein hohes Vertrauen (KIRCHER et al., 2014) sowie eine hohe Akzeptanz solcher Systeme, zumindest bei geringen Geschwindigkeiten, wie beispielsweise im Stau, weil der Fahrer eine deutliche Entlastung erlebt (nachgewiesen sowohl über Befragung als auch über die Bereitschaft, die Querführung im Stau dem System zu überlassen; SCHALLER et al., 2008). Eine Verringerung der wahrgenommenen kognitiven Beanspruchung durch die

Fahraufgabe zeigte sich auch bei STANTON et al. (2001). BEGGIATO, HARTWICH, SCHLEINITZ, KREMS, OTHERSEN & PETERMANN-STOCK (2015) berichten allerdings davon, dass im direkten Vergleich zwischen manuellem, teilautomatisiertem und hochautomatisiertem Fahren das teilautomatisierte Fahren am anstrengendsten bewertet wurde, weil dort permanent überwacht werden muss (höhere Anzahl an Kontrollblicken in die Spiegel und auf das Kombi-Instrument-Display; ähnliche Befunde auch von LARGE et al., 2017). Fahrer wünschen sich hier insbesondere Informationen, die die Überwachung unterstützen (BEGGIATO et al., 2015). STAPEL, MULLAKKAL-BABU & HAPPEE (2017) konnten zeigen, dass die wahrgenommene Beanspruchung von der Erfahrung mit derartigen Systemen abhängt: Während Fahrer, die bereits Erfahrung mit automatisierten Fahrzeugen gemacht haben, eine Reduktion des Workloads wahrnehmen, ist dieser Aspekt bei unerfahrenen Fahrern sogar erhöht. Die Anforderungen an die visuelle Aufmerksamkeit reduzierten sich während des Fahrens mit Teilautomation im Vergleich zum manuellen Fahren nicht.

Zusammenfassend zeigt sich für die Nutzung von L2-Systemen eine relativ hohe Akzeptanz und Vertrauen. Der Fahrer nimmt zumindest in manchen Fahrsituationen, wie dem Fahren im Stau, eine Entlastung wahr. Generell wird die Überwachungsaufgabe aber als erschöpfend bewertet. Während der Nutzung der Systeme besteht die Gefahr von erhöhter Ablenkung durch die Zuwendung zu anderen Tätigkeiten, aber auch durch die Systemanzeigen selbst. Aktuell noch widersprüchliche Befunde gibt es dazu, wie gut Fahrer an Systemgrenzen reagieren. Es scheint, dass Fahrer eher dann keine Probleme erfahren, wenn Situationen antizipierbar sind, eine hohe Aufmerksamkeit auf der Fahrsituation liegt und das System zusätzlich eine Warnung ausgibt. Davon ist aber in L2-Systemen nicht immer auszugehen, da die Fahrer in der Lage sein müssen, bei unerwarteten Situationen unmittelbar einzugreifen. Es gibt erste Anhaltspunkte dafür, dass eine angemessene MMS-Gestaltung dies unterstützen kann. Die potenzielle Verwechslungsgefahr mit höheren Automationsstufen und damit einhergehendem übermäßigem Vertrauen in teilautomatisiertes Fahren muss durch die Vermittlung eines ausreichenden Systemwissens und eines angemessenen mentalen Modells erfolgen.

### 1.3 Bewertungsmethoden: Expertenbewertung und Nutzerstudie

Vor dem Hintergrund der dargestellten zu erwartenden Probleme im Umgang mit teilautomatisierten Systemen ist zu fordern, dass abgesichert werden sollte, ob Nutzern von teilautomatisierten Systemen ein adäquates Verständnis der Automation vermittelt wird. Ziel ist die Entwicklung einer ganzheitlichen Methode, die unabhängig von verschiedenen Systemausgestaltungen Aussagen über die Gebrauchstauglichkeit (engl. Usability) und Gebrauchssicherheit machen kann. Während erste Ansätze zur Überprüfung der Gebrauchstauglichkeit für hochautomatisiertes Fahren (SAE L3; NHTSA, 2017; NAUJOKS, HERGETH, WIEDEMANN, SCHÖMIG, FORSTER & KEINATH, 2019a; NAUJOKS, WIEDEMANN, SCHÖMIG, HERGETH & KEINATH, 2019c) sowie der Kontrollierbarkeit (NAUJOKS, WIEDEMANN, SCHÖMIG, JAROSCH & GOLD, 2018a, sog. TOC-Rating) bereits existieren, gibt es bisher noch keine standardisierte und ganzheitliche Bewertungsmethodik dieser Aspekte von Fahrzeugen mit teilautomatisierten Fahrfunktionen. Diesbezüglich müssen Ansätze entwickelt werden, mit deren Hilfe geprüft werden kann, ob die MMS die Voraussetzungen für eine sichere und dem Automationslevel angemessene Interaktion erfüllt, sodass Fahrer dementsprechend mit den Fahrzeugen umgehen können.

Zum einen ist hierbei zu bewerten, inwieweit die MMS geeignet ist, den Fahrer beim Verständnis der Systemfunktionalität und einem sicheren Umgang mit Systemgrenzen zu unterstützen. Hierzu müssen entsprechende Mindestanforderungen an die Gestaltung der MMS von automatisierten Systemen definiert werden (NAUJOKS, 2019c). Darüber hinaus ist zu prüfen, inwieweit der Fahrer in der direkten Interaktion mit dem System die ihm dargebotenen Informationen verstehen und effektiv und sicher nutzen kann (NAUJOKS, HERGETH, WIEDEMANN, SCHÖMIG & KEINATH, 2019b). Es sollte sichergestellt sein, dass der Fahrer über ein ausreichendes Maß an Systemverständnis, System- und Situationsbewusstsein verfügt und dabei nicht in übermäßigem Grad beansprucht wird. Für diese Fragen ist die Beobachtung der Interaktion des Nutzers mit dem System und der MMS notwendig.

Das übergeordnete Ziel ist die Entwicklung eines umfassenden, standardisierten Bewertungskonzepts für MMS beim automatisierten Fahren, in dem

verschiedene Prüfkriterien und Prüfverfahren für eine effiziente und sichere MMI definiert sind.

In der Usability-Forschung werden im Rahmen des Produktentwicklungszyklus zur Bewertung von Aspekten der MMS zwei grundsätzliche Ansätze vorgeschlagen (NIELSEN, 1994):

- Die expertenbasierte Bewertung durch Human-Factors Spezialisten (z. B. im Rahmen einer heuristischen Evaluation, d. h. der Sammlung von potenziell auftretenden Problemen durch selbstständiges Erproben und/oder anhand von bestehenden Guidelines und Ableitung eines Expertenurteils) und
- die empirische Bewertung, d. h. die Durchführung einer Nutzerstudie mit einer repräsentativen Stichprobe potenzieller Endnutzer des Systems (NIELSEN & MOLICH, 1990).

Die Methoden ergänzen sich in optimaler Weise darin, Usability- und Verständnisprobleme aufzudecken, die in der Interaktion mit dem Produkt über die MMS bestehen (JEFFRIES & DESURVIRE 1992). Während die Bewertung durch Experten in der Lage ist, potenzielle Probleme in einem frühen Entwicklungsstadium zu identifizieren, und einfach anzuwenden ist, ist die Nutzerstudie in der Regel besser dafür geeignet, reale Probleme in der tatsächlichen Nutzung des Produkts aufzudecken, wobei sie allerdings wesentlich aufwendiger in der Planung und Durchführung ist. Nachdem die beiden Methoden unterschiedliche Arten von Usability-Aspekten aufdecken, ist es empfehlenswert, beide Methoden im Rahmen eines umfassenden Bewertungskonzepts einzusetzen (JEFFRIES, MILLER WHARTON & UYEDA, 1991; ISO 9241 2018).

Als Methode für die Bewertung der MMI von L3-Fahrfunktionen haben NAUJOKS et al. (2019a) sich bereits auf dieses Vorgehen verständigt und einen zweistufigen, konsekutiven Bewertungsansatz vorgeschlagen. Als heuristische Evaluation empfehlen die Autoren die Anwendung einer Checkliste durch Human Factors Experten, anhand derer die Erfüllung von bestehenden Standards und Guidelines einer MMS überprüft werden kann. Zusätzlich wird ein Testprotokoll für Nutzertest im Fahrsimulator vorgestellt. In Anlehnung an diese Vorgehensweise werden die folgenden zwei Bausteine für eine Gesamtbewertung von teilautomatisierten Fahrfunktionen vorgeschlagen.

Zum einen wird eine heuristische Evaluation zur Bewertung der MMS durch Experten vorgeschlagen. Diese Bewertung erfolgt mithilfe einer Checkliste, bei der Experten der MMI beurteilen, ob Mindestanforderungen der MMS bezüglich Wahrnehmbarkeit und Verständlichkeit von Systemzuständen erfüllt sind. Die Checkliste setzt vor allem an der Gestaltung der MMS an, da diese während der Nutzung der primäre Kommunikationskanal ist. Dazu zählt beispielsweise die Gestaltung von visuellen und auditiven Systemausgaben. Teile der Checkliste sind angelehnt an die Checkliste von NAJJOKS et al., 2019c für das hochautomatisierte Fahren. Darüber hinaus wird sie um Aspekte der Systembedienung (d. h. Interaktion mit dem System über die MMS), gezielte Aspekte für das teilautomatisierte Fahren (z. B. bezüglich einer Fahrerüberwachung) sowie die Vermittlung eines angemessenen Systemverständnisses über das Benutzerhandbuch erweitert.

Für die Überprüfung von interaktionsbezogenen Aspekten der sicheren MMI von kontinuierlichen L2-Automatisierungsfunktionen werden Nutzerstudien empfohlen. Nutzerstudien fokussieren auf die Evaluation des natürlichen Nutzerverhaltens während der Verwendung des Systems und bieten die Möglichkeit, verschiedene Prozesse des Nutzerverhaltens in ihrer Wechselwirkung zu untersuchen. Dabei kann neben Aspekten der MMS auch die Systemauslegung (z. B. Güte der Querführung, Häufigkeit von Systemgrenzen) auf das Fahrerverhalten mitberücksichtigt werden. In diesem Rahmen wurde ein szenarienbasiertes, standardisiertes Bewertungsinstrument zur Erfassung von Beobachtungs- und Befragungsvariablen mittels einer Tablet-Anwendung entwickelt. Es ermöglicht die Erfassung eines angemessenen Systemverständnisses und -bewusstseins, eines ausreichenden Situationsbewusstseins, die Beanspruchung durch die Nutzung des Systems sowie die Fähigkeit zu einem sicheren und effizienten Verhalten im Fall von Systemgrenzen und bei Systemausfällen. Die Anwendung erfolgt in relevanten Use Cases im Rahmen von Nutzerstudien in der Fahrsimulation oder im Realverkehr.

Beide Methoden können im Simulator für ein in Entwicklung befindliches System oder zur Testung von Prototypen oder Serienfahrzeugen verwendet werden.

## 1.4 Definition von Anforderungen und Prüfkriterien

Die MMI muss bestimmte Anforderungen erfüllen, die in den folgenden Abschnitten beschrieben werden. Der Erfüllungsgrad dieser Anforderungen muss innerhalb der Expertenbewertung sowie der MMI von Nutzern im Umgang mit der teilautomatisierten Fahrfunktion geprüft werden. Je nach Methode sind dabei unterschiedliche Prüfkriterien anzunehmen und zu untersuchen: Das Checklistenverfahren formuliert Gestaltungsempfehlungen, um zu untersuchen, ob die grundlegenden Voraussetzungen gegeben sind, um die Anforderungen zu erfüllen (z. B. das reine Vorhandensein entsprechender Informationen und deren wahrnehmbare und verständliche Gestaltung). Dagegen versucht die Methode der Nutzerstudie aus objektiv beobachtbarem Verhalten der Fahrer in der Interaktion mit dem System die Usability und die Fahrsicherheit direkt zu erfassen und auf die darunter liegenden impliziten Anforderungen zu schließen.

Als relevante Anforderungen an die MMI beim teilautomatisierten Fahren werden dabei folgende definiert:

### Erstens

Der Fahrer muss über ein angemessenes mentales Modell, d. h. ein korrektes und verständliches Systemwissen verfügen. Nutzer können auf unterschiedliche Weise Wissen über das genutzte automatisierte Fahrzeugsystem erwerben. Darunter fallen neben Informationen, die über die Medien (bspw. Werbung durch Automobilhersteller) oder auf der Herstellerseite kommuniziert werden, auch der Markenname der Funktion (bspw. Driving Assistant Professional (BMW), Autopilot (Tesla) oder Supercruise (GM)) sowie Informationen, die im Benutzerhandbuch zu der jeweiligen Systemfunktion zu finden sind. Zu bewerten gilt deshalb, inwieweit die Nutzer mittels der genannten Quellen in adäquater Weise über die Fähigkeiten der Automation und ihre eigene Verantwortlichkeit im Vorfeld der Nutzung informiert werden. Neben der Vermittlung durch deklaratives Wissen (d. h. explizite Informationen), kann auch implizites Wissen eine Rolle für die Ausbildung eines adäquaten mentalen Modells spielen. Dabei spielt sowohl die erworbene Vorerfahrung mit ähnlichen Systemen und Assistenzfunktionen eine Rolle, als auch die Erfahrungen während der Fahrt über Systemeigenschaften und Systemgrenzen.



Zur Untersuchung der Güte der Vermittlung des Systemwissens eignet sich die Methode der heuristischen Evaluation am besten. Dazu findet in der Expertenbewertung durch die Checkliste eine Überprüfung des Benutzerhandbuchs dahingehend statt, ob relevante Informationen, die zum Systemverständnis nötig sind, enthalten sind. Im Rahmen der Nutzerstudie kann der Fahrer nach einer Fahrt zu entsprechenden Aspekten des Systems gezielt befragt werden, um zu prüfen, welches mentale Modell er durch das unmittelbare Erleben des Systems während der Fahrt aufgebaut hat.

### **Zweitens**

Das System- und Situationsbewusstsein (engl. Mode Awareness bzw. Situation Awareness) des Fahrers in der konkreten Bedien- und Fahrsituation muss durch die Gestaltung der MMS unterstützt werden. Fahrer sollten sich während der Systemnutzung jederzeit über ihre Verantwortung für die Fahraufgabe sowie über den Modus der Automation bewusst sein. Den Ausgaben der MMS kommt die Aufgabe zu, dem Fahrer während der Systemnutzung den Status der Automation (bspw. Querführung aktiv, passiv oder deaktiviert) und seine Rolle zu vermitteln, damit die Automation in angemessener Weise genutzt werden kann. Dadurch soll ein angemessenes System- und Situationsbewusstsein vermittelt werden, mit dem Ziel, ein der Fahrsituation jederzeit angemessenes Fahrerverhalten zu erreichen.

Die Prüfung einer angemessenen Mode- und Situation Awareness bzw. deren Voraussetzungen kann sowohl über das Checklistenverfahren als auch durch die Beobachtung und Befragung des Fahrers in einer Nutzerstudie erfolgen.

Das Checklistenverfahren zielt dabei vor allem darauf ab, die Gestaltung der MMS dahingehend zu begutachten, ob die Voraussetzungen für die Erfüllung dieser Anforderung erfüllt sind. Die Bewertung der MMS umfasst zum einen die Anzeigen und Ausgaben des Systems. Dies beinhaltet das Vorhandensein von Feedback über den Systemzustand, um den Nutzer über Zustände oder Transitionen zu informieren. Weiterhin ist die konkrete Ausgestaltung von Feedback relevant. Dieses sollte leicht wahrnehmbar und verständlich sein, damit der Fahrer im richtigen Moment die korrekten Schlüsse ziehen kann.

In der Nutzerstudie wird die Verständlichkeit von Systemausgaben sowie die Mode und Situation

Awareness indirekt über Befragung und Beobachtung erfasst. Während des Fahrens in einem bestimmten Systemzustand (vor allem während des Fahrens mit aktivem L2-System) wird geprüft, ob der Fahrer sich seiner Verantwortlichkeit bewusst ist und wie gut die MMS ihn dabei unterstützt. Dies wird im Verhalten daran deutlich, ob der Fahrer in angemessenem Grad seine Überwachungsaufgabe erfüllt, z. B. seine Hände am Lenkrad belässt bzw. im Fall einer ausgegebenen Hands-Off-Warnung seine Hände wieder zurück ans Lenkrad nimmt und einen ausreichenden Anteil seiner Aufmerksamkeit auf die Verkehrsumgebung richtet. Daneben soll der Fahrer auch subjektive Angaben über den wahrgenommenen Grad des Gefühls der Sicherheit im Umgang mit dem System treffen („ich weiß, was ich tun muss“; „ich verstehe, was das System mir anzeigt“; „ich kann nachvollziehen, warum sich das System derartig verhält“). Im Gegensatz zu einem generischen Systemverständnis, das auch in einer Nachbefragung erfasst werden könnte, soll der Fahrer hier die für einen Use Case spezifischen Systemanzeigen und Systemverhaltensweisen bewerten.

### **Drittens**

Der Fahrer soll in der Lage sein, das System effizient zu nutzen, d. h. die Bedienung des Systems sollte mit einem geringen Aufwand erfolgen können. Dazu darf die Bedienlogik nicht zu kompliziert und mental zu beanspruchend sein, um ein gewünschtes Handlungsziel zu erreichen.

In der Expertenbewertung findet dazu eine Überprüfung der Bedienelemente statt, z. B. ob diese leicht erreichbar sind und entsprechend benannt sind.

In der Nutzerstudie werden der Aufwand zur Systembedienung und das daraus resultierende Fahrverhalten als Maße für die Beanspruchung des Fahrers durch die MMS erfasst. Dies umfasst zum einen vom Fahrer initiierte Transitionen zwischen Systemzuständen, d. h. Aktivierung des Systems bzw. Teilfunktionen sowie deren Deaktivierung. Fehler, Unsicherheiten oder unnötige Verhaltensweisen in der Systembedienung weisen auf eine hohe Komplexität der System- und Bedienlogik und u. U. Schwierigkeiten in der Verständlichkeit der notwendigen Bedienhandlungen für eine erfolgreiche Zielerreichung hin (z. B. bei der Aktivierung: ist das System schon aktiv oder noch im Stand-by-Zustand?).

## Viertens

Die Fahrsicherheit darf durch die Interaktion mit dem System zu keinem Zeitpunkt beeinträchtigt sein. Dies betrifft zum einen die Systembedienung, die ohne Einbußen in der Fahrleistung stattfinden können muss. Insofern ist eine wenig beanspruchende Systembedienung eine wichtige Voraussetzung für die Fahrsicherheit. Besonders relevant darüber hinaus ist aber die angemessene Reaktion im Fall von Systemgrenzen, an denen der Fahrer oft ohne Vorwarnung durch das System die Fahraufgabe wieder übernehmen muss.

Dieser Aspekt der Fahrsicherheit oder Kontrollierbarkeit des Systems kann am besten in Nutzerstudien anhand des Fahrerverhaltens bei systeminitiierten Transitionen im Fall von Systemgrenzen oder Systemausfällen erfasst werden. Dies geschieht über die Beobachtung der Interaktion mit dem System, d. h. ob der Fahrer das System oder Teilfunktionen im Fall von Systemgrenzen oder Systemaus-

fällen rechtzeitig und angemessen abschaltet bzw. ob er angemessen, d. h. ohne Einbußen in der Fahrsicherheit für sich und andere Verkehrsteilnehmer mit einer Übernahme der Fahraufgabe oder einer Teilaufgabe reagiert. Fahrfehler oder gar Gefährdungen von sich und anderen weisen wiederum indirekt darauf hin, dass der Fahrer über ein unzureichendes System- und/oder Situationsbewusstsein verfügt und daher von einem Systemausfall „überrascht“ und in einem übermäßigen Grad beansprucht wird.

Das übergeordnete Bewertungskonzept für die Bewertung der Effizienz und Sicherheit der MMI mit den zugrunde liegenden Anforderungen und deren jeweiligen Prüfkriterien über die beiden Methoden „Heuristische Evaluation“ und „Nutzerstudie“ ist in Bild 1-2 grafisch dargestellt. Die Pfeile deuten dabei an, welche der Anforderungen über welche Methode mittels welchen Kriteriums geprüft werden kann. Wenn beide Methoden einen Beitrag leisten können, sind zwei Pfeile dargestellt. Es wird darauf hin-

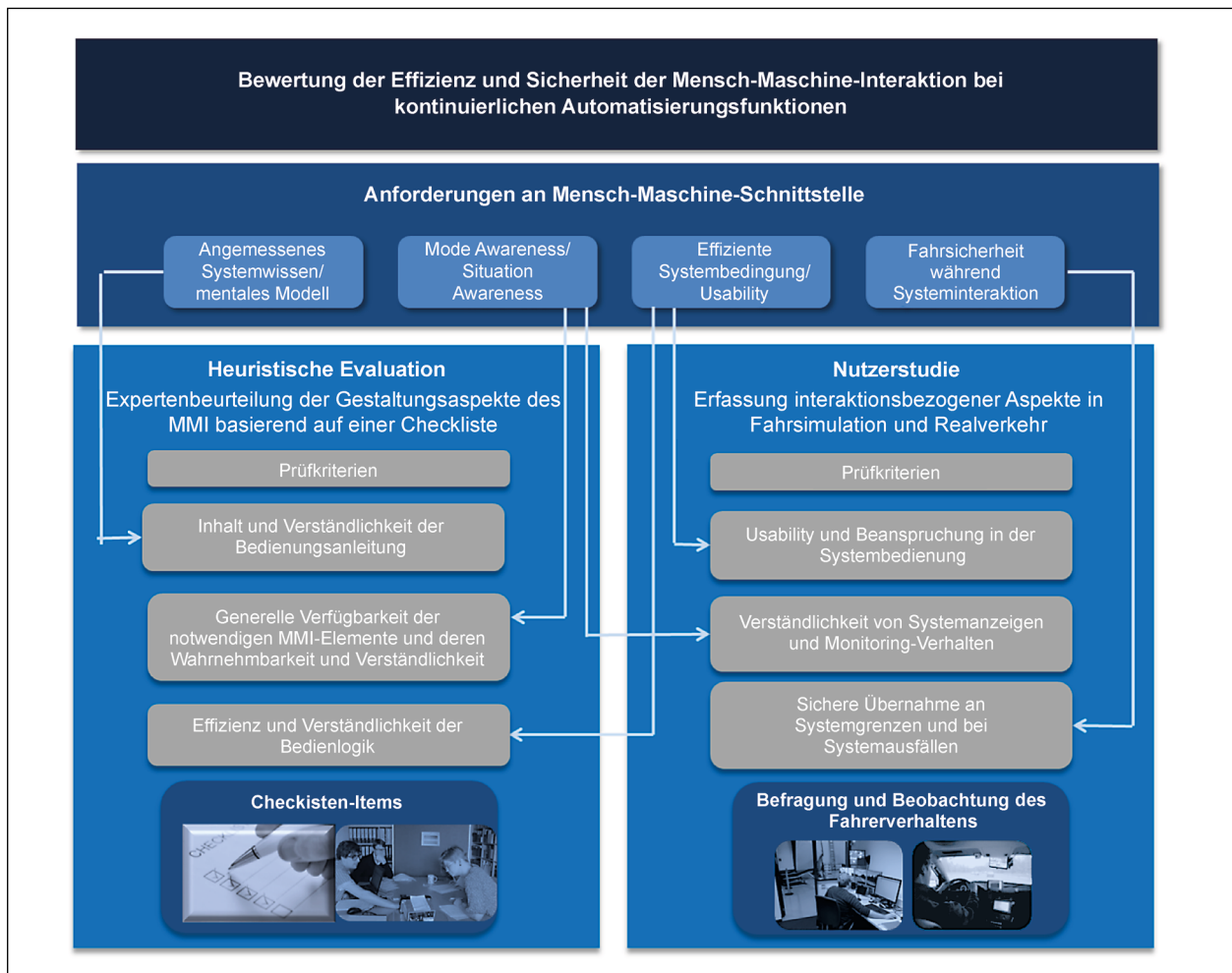


Bild 1-2: Definition von Prüfkriterien, die zur Bewertung einer erfolgreichen und sicheren Interaktion mit einem teilautomatisierten System herangezogen werden.

gewiesen, dass im Bild nicht alle Abhängigkeiten zwischen den einzelnen Konstrukten abgebildet werden können.

## 1.5 Definition der zu testenden Use Cases

Zur Bewertung einer effizienten und sicheren MMI von kontinuierlichen Automatisierungsfunktionen ist zu definieren, in welchen Anwendungsfällen, sog. Use Cases, diese Interaktion zu prüfen ist. Die Auswahl dieser Use Cases stellt eine zentrale Anforderung für die Durchführung beider Methoden dar.

Unter dem Begriff „Use Case“ soll in diesem Zusammenhang eine Sequenz von Interaktionen zwischen zwei Aktoren (Mensch und Maschine) in einem Mensch-Maschine-System mit dem Ziel, ein beobachtbares Ergebnis oder Ziel zu erreichen, verstanden werden (NAUJOKS, HERGETH, WIEDEMANN, SCHÖMIG & KEINATH, 2018b). So stellt eine Bedienhandlung, die das System in einen bestimmten Zustand versetzt, wie beispielsweise die Aktivierung, einen Use Case dar.

Aus dieser zunächst eher generisch formulierten Beschreibung von Zustandsübergängen oder Systemzuständen können anschließend konkrete Testfälle abgeleitet werden, die mit dem zu evaluierenden L2-System für die Erfüllung des Use Cases verwendet werden können (NAUJOKS et al., 2018b). So kann die Aktivierung eines L2-Systems durch verschiedene Bedienpfade möglich sein, welche allesamt Instanzen (d. h. Testfälle) des generischen Use Cases „Systemaktivierung“ darstellen. Diese Testfälle zeichnen sich folglich dadurch aus, dass sie systemspezifische, kontextabhängige Interaktionssequenzen beschreiben. So kann ein und derselbe Use Case innerhalb verschiedener Szenarien geprüft werden (z. B. eine Übernahme der Kontrolle an verschiedenen Systemgrenzen wie einer Baustelle, einer Autobahnabfahrt oder einem Pannenfahrzeug). Demnach beinhaltet ein Prüfscenario (oder Testfall; die Begriffe werden hier synonym verwendet) eine genauere Spezifikation der Umgebungs- und Kontextfaktoren, in denen der generische Use Case stattfindet.

Die Herleitung der Use Cases orientiert sich an dem von NAUJOKS et al. (2018b) vorgeschlagenen Ansatz, bei dem ein schrittweises Vorgehen angewandt wird: In einem ersten Schritt werden alle prinzipiell möglichen Use Cases definiert. Darunter fal-

len sämtliche Zustandsübergänge zwischen verschiedenen Systemzuständen unter Berücksichtigung von im Fahrzeug ebenfalls vorhandenen geringeren Automationsstufen. Dazu gehört beispielsweise die Aktivierung des Systems aus dem manuellen Fahren heraus oder eine Transition von L2 auf eine niedrigere Automationsstufe (z. B. Fahren mit ACC; d. h. L1). Je nach Fahrzeugausstattung sind auch Übergänge zu höheren Automationsstufen zu betrachten. Zu beachten ist, dass die Übergänge entweder vom Fahrer initiiert werden können (z. B. Aktivierung) oder vom System angestoßen werden (z. B. Deaktivierung aufgrund einer Systemgrenze). Weiterhin sind dauerhafte Systemzustände, wie etwa das Fahren mit aktivem L2-System zu betrachten, welches eine permanente Systemüberwachung erfordert.

Weiterhin werden besondere Systemzustände innerhalb einer Automationsstufe berücksichtigt, wie den Vorgang einer Veränderung der Systemfunktionalität durch den Fahrer (z. B. Verstellung der Set Speed oder des Wunschabstands im ACC) oder verringerte Systemumfänge. Diese können auftreten, wenn das System (kurzzeitig) den Systemumfang verringert z. B. aufgrund mangelnder Sensorik, ohne dass eine explizite Transition zu einer geringeren Automatisierungsstufe stattfinden muss. Hiermit ist für L2-Systeme der häufige Fall von kurzzeitigen Ausfällen der Querführung und dadurch bedingten Übergängen in den Stand-by-Modus und die anschließende Reaktivierung der Funktion gemeint. Ebenso fallen darunter Zustände, in denen höhere Automationsstufen nicht aktiviert werden können (z. B. Querführungsassistenten nicht aktivierbar, weil aktuell keine Fahrstreifenmarkierungen erkannt werden). Besondere Use Cases, die aus weiteren integrierten Assistenten und damit Systemzuständen resultieren, sind beispielsweise ein fahrerinitiiertes Fahrstreifenwechsel.

Entsprechend werden die in Tabelle 1-2 dargestellten Arten von Use Cases definiert. Die Instanzen der Use Cases (Zielerreichung auf unterschiedliche Arten) werden ebenfalls aufgeführt. Die Herleitung berücksichtigt Systeme mit der Möglichkeit, geringere Automationsstufen (siehe Stufen nach SAE) zu aktivieren (z. B. nur ACC; siehe aktueller Benchmark). Insgesamt sind 21 Use Case bzw. unter Berücksichtigung der verschiedenen Instanzen der Use Cases (Testfälle) 27 Testfälle möglich.

Nicht berücksichtigt sind in dieser Auflistung Transitionen zu höheren Automationsstufen. Wenn ein

System auch höhere Automationsstufen enthält (z. B. L3), sollten auch Transitionen von höheren Automationsstufen betrachtet und die Use Cases entsprechend erweitert werden (siehe NAUJOKS et al., 2019a).

Anschließend erfolgt eine Einschränkung des Use Case Sets auf eine Auswahl, die in den beiden Methoden getestet wird. Die Auswahl erfolgt dabei basierend auf der Systemlogik des jeweiligen zu testenden L2-Systems und den beinhalteten System-

		Use Case
Systemzustände	Fahren in bestimmter Automationsstufe	<b>Regulär</b>
		(1) L0 (d. h. manuelles Fahren), (= L0)
		(2) L1, wobei nur die Querführung aktiv ist (= L1 <sub>quer</sub> )
		(3) L1, wobei nur die Längsführung aktiv ist (= L1 <sub>längs</sub> )
	(4) L2 (d. h. Quer- und Längsführung sind aktiv; = L2)	
	<b>Verringerter Systemumfang</b>	
	(5) in L0, d. h. höhere Automationsstufen (z. B. Querführung) können nicht aktiviert werden (L0 verringert)	
	(6) in L1 <sub>quer</sub> , d. h. höhere Automationsstufen (z. B. durch zusätzliche Längsführung) können nicht aktiviert werden (L1 <sub>quer</sub> verringert)	
	(7) in L1 <sub>längs</sub> , d. h. höhere Automationsstufen (z. B. durch zusätzliche Querführung) können nicht aktiviert werden (L1 <sub>längs</sub> verringert)	
	(8) in L2, z. B. kurzzeitiger Übergang der Querführung in Stand-by-Modus mit anschließender automatischer Reaktivierung (L2 verringert)	
	Verstellen von Settings in Automationsstufe	(9) während des Fahrens in L1 <sub>quer</sub>
(10) während des Fahrens in L1 <sub>längs</sub>		
(11) während des Fahrens in L2		
Transitionen	Aktivierung fahrerinitiiert	(12) Von L1 <sub>quer</sub> (der Querführung) <ul style="list-style-type: none"> <li>• Instanz (12.1): Aus L0 heraus</li> </ul>
		(13) Von L1 <sub>längs</sub> (der Längsführung) <ul style="list-style-type: none"> <li>• Instanz (13.1): Aus L0 heraus</li> </ul>
		(14) Von L2 <ul style="list-style-type: none"> <li>• Instanz (14.1): von L0 direkt</li> <li>• Instanz (14.2): aus aktivem L1<sub>quer</sub> heraus (d. h. Hinzuschalten der Längsführung)</li> <li>• Instanz (14.3): von L2 aus aktivem L1<sub>längs</sub> heraus (d. h. Hinzuschalten der Querführung)</li> </ul>
		(15) Eines unterstützten Fahrstreifenwechsels in L2
		(16) Von L1 <sub>quer</sub> (der Querführung)
	Deaktivierung fahrerinitiiert	(17) Von L1 <sub>längs</sub> (der Längsführung)
		(18) Von L2 (z. B. im Fall eines Systemausfalls nötig) <ul style="list-style-type: none"> <li>• Instanz (18.1): direkt</li> <li>• Instanz (18.2): Runterschalten von L2 in L1<sub>quer</sub> (d. h. Wegschalten der Längsführung)</li> <li>• Instanz (18.3): Runterschalten von L2 in L1<sub>längs</sub> (d. h. Wegschalten der Querführung)</li> </ul>
		(19) Von L1 <sub>quer</sub> (z. B. im Fall eines Sensorausfalls oder einer Systemgrenze)
	Deaktivierung systeminitiiert	(20) von L1 <sub>längs</sub> (z. B. im Fall eines Sensorausfalls oder einer Systemgrenze)
		(21) von L2 <ul style="list-style-type: none"> <li>• Instanz (21.1): direkt (z. B. aufgrund einer Systemgrenze)</li> <li>• Instanz (21.2): Runterschalten von L2 in L1<sub>quer</sub> (d. h. Abschalten der Längsführung)</li> <li>• Instanz (21.3): Runterschalten von L2 in L1<sub>längs</sub> (d. h. Abschalten der Querführung)</li> </ul>

Tab. 1-2: Übersicht aller theoretisch möglichen Use Cases in L2-Systemen

zuständen, etwa ob ein System Stand-by-Zustände vorsieht oder nicht bzw. ob die separate Nutzung einzelner Automationsstufen vorgesehen ist (z. B. alleinige Nutzung der Querführung möglich vs. nur in Kombination mit bereits aktivem ACC nutzbar). Dies führt zu einer ersten Reduktion der Anzahl der relevanten Use Cases, die zu prüfen ist.

In einem weiteren Schritt findet eine Priorisierung der verbleibenden Use Cases statt, um die Anzahl der Prüffälle mit einem akzeptablen Aufwand untersuchen zu können. Wie bei den Prüfkriterien erwähnt, sollte das Ziel eines guten Systems mit guter MMS sein, dass Fahrer sicher und effizient damit interagieren können. Die Expertenbewertung prüft dafür die Voraussetzungen, die Nutzerstudie das tatsächliche Fahrerverhalten. Deswegen ist die Auswahl der Use Cases eng mit der Definition der Bewertungskategorien verknüpft, und es ist zielführend, in beiden Methoden die gleichen Use Cases zu untersuchen.

Die Auswahl erfolgt basierend auf der Definition potenziell zu erwartender Probleme und daraus abgeleiteter Bedingungen für die erfolgreiche und effiziente Bewältigung eines Use Case. Probleme bei der Systeminteraktion, die aufgrund der Gestaltung der MMS oder der Systembedienung auftreten können, sind vor allem an Systemzustandsübergängen zu erwarten. Relevante Use Cases, bei denen die Interaktion geprüft werden sollte, sind daher die fahrerinitiierte Aktivierung (sowohl aus dem manuellen Fahren als auch aus anderen durch die jeweilige Funktionsausstattung möglichen Systemzuständen heraus; z. B. ACC oder automatisierte Querführung) sowie die fahrerinitiierte Deaktivierung. Dabei sollte das System zum einen vollständig deaktiviert werden. Zum anderen sollte – wenn mit dem zu evaluierenden System möglich – ein Übergang zu niedrigeren Automationsstufe erfolgen (z. B. ACC oder automatisierte Querführung).

Der Use Case des „Fahrens in L2“ ist notwendig, um zu überprüfen, ob sich der Fahrer während der Nutzung des Systems seiner Verantwortlichkeit zur System- und Situationsüberwachung bewusst ist. Innerhalb des aktiven L2-Systems sollte ein durch das System unterstützter, fahrerinitiiertes Fahrstreifenwechsel erlebt werden, sofern dieser im Funktionsumfang vorhanden ist und eine Hands-Off-Warnung ausgelöst werden. Die bewusste Auslösung der Hands-Off-Warnung wird nur in der Checklistenmethode angewandt, um deren Gestaltung bewerten zu können, nicht jedoch in der Nutzerstudie, da

dort die Frage, ob den Fahrern bewusst ist, dass sie beim teilautomatisierten Fahren die Hände am Lenkrad lassen müssen, eine zentrale Fragestellung ist.

Zudem müssen die Verständlichkeit von systeminitiierten Zustandsänderungen und das angemessene Verhalten auf diese zentraler Bestandteil der Prüfung sein. Dies betrifft zum einen Use Cases, in denen das System in L2 nur einen verringerten Systemumfang bieten kann (z. B. im Fall eines kurzzeitigen Ausfalls einer Teilfunktion, wie der Querführung). Außerdem sollten Transitionen von höheren auf niedrigere Automatisierungsstufen (typischerweise auf L1<sub>längs</sub> oder L0), wie es bei einem (Teil-)systemausfall oder an Systemgrenzen der Fall ist, erlebt werden, da diese eine besondere Herausforderung darstellen, nachdem der Fahrer sie bemerken und mit einer entsprechenden Fahrerreaktion reagieren muss. Dies ist in einer Nutzerstudie in der Fahrsimulation entsprechend leichter umzusetzen als bei der heuristischen Evaluation, die typischerweise im Realverkehr stattfindet.

Basierend auf dieser Argumentation sollten in den beiden Methoden folgende Use Cases bei der Bewertung der MMS geprüft werden, die sich auf prototypische L2-Systeme ähnlich dem in Kapitel 1.1 beschriebenen konzentrieren (Tabelle 1-3). Sehr ähnliche Use Cases auf anderen Automationsstufen sowie Use Cases, die gemäß dem spezifischen Systemzustandsdiagramm ungewöhnlich bzw. nicht vorgesehen sind, können in der Prüfung entfallen.

## 1.6 Übersicht Methodenbeschreibung

Der vorliegende Bericht beschreibt die beiden methodischen Ansätze. Zuerst werden die Methode der heuristischen Evaluation sowie die dazugehörige Checkliste und deren Anwendung erläutert. Anschließend wird die Methode der Nutzerstudie und die dafür entwickelte Tablet-Anwendung S.A.D.E. als szenarienbasiertes, standardisiertes Bewertungsinstrument zur Erfassung interaktionsbezogener Aspekte der sicheren MMI von kontinuierlichen L2-Automatisierungsfunktionen dargestellt. Nach jedem dieser Berichtsteile erfolgt eine Beschreibung der empirischen Überprüfung des jeweiligen Konzeptes und eine Diskussion der jeweiligen Methode. Abschließend werden nochmals Vor- und Nachteile der einzelnen Methoden sowie die Integration der beiden zu einer Gesamtbewertung diskutiert.

Use Cases
Aktivierung <ul style="list-style-type: none"> <li>• Von Manuell auf L2</li> <li>• Von nur Querführung/Längsführung auf L2 (systemspezifisch)</li> </ul>
Deaktivierung <ul style="list-style-type: none"> <li>• Von L2 zu Manuell</li> <li>• Von L2 zu nur Querführung oder nur Längsführung (systemspezifisch)</li> </ul>
Fahren in L2, inklusive <ul style="list-style-type: none"> <li>• Erleben einer Hands-Off-Warnung (nur Expertenbewertung)</li> <li>• Erleben eines temporären Passiv-Werdens der aktiven Querführung aufgrund Nicht-Erkennung von Fahrstreifenmarkierungen o. Ä. (Querführung wechselt in Stand-by-Zustand)</li> </ul>
Fahrerinitiierte Ausführung eines vom System unterstützten Fahrstreifenwechsels (falls Funktion vorhanden)
Systemgrenze des L2-Systems, welche zur vollständigen Deaktivierung der Querführung führt
Optional: Verstellen von Settings, z. B. der Set Speed in L2

Tab. 1-3: Übersicht über zu prüfenden Use Cases in der MMS-Bewertung mittels Expertenbewertung und Nutzerstudie

## 2 Checkliste für die Expertenbewertung der Mensch-Maschine-Schnittstelle

Aus einer Literaturanalyse relevanter Quellen werden Anforderungen an die Gestaltung der Ausgaben und die Bedienung der Mensch-Maschine-Schnittstelle sowie das Benutzerhandbuch abgeleitet, deren Erfüllungsgrad mithilfe einer Checkliste abgeprüft werden soll. Die Erfüllung der Anforderungen stellt eine Voraussetzung für eine hinreichend positive Bewertung der MMS dar. In diesem Kapitel werden der Aufbau der Checkliste, die Itemherleitung, sowie die Anwendung des Verfahrens beschrieben. Im Anschluss werden die Ergebnisse mehrerer Erprobungsphasen der Checkliste in Real- und Fahrsimulatorstudien vorgestellt. Am Ende des Kapitels erfolgt eine kritische Betrachtung der Methode.

### 2.1 Aufbau der Checkliste

Die Checkliste enthält 14 Anforderungen an die MMS beim teilautomatisierten Fahren, deren Erfüllungsgrad von Experten bewertet werden soll. Basierend auf dem Erfüllungsgrad der Anforderungen wird ein finales Rating zur Systembewertung abgegeben. Die Erfüllung einer Anforderung wird anhand von mehreren Items bestimmt (siehe Kapitel 2.2 für Bewertungskategorien und -prinzipien und Kapitel 2.3 für Itemherleitung). Die Themen, denen die Anforderungen zugeordnet sind, entsprechen den in Kapitel 1.4 genannten Prüfkriterien. Die Checkliste beginnt mit Anforderungen zum Benut-

zerhandbuch, deren Erfüllung dazu beitragen soll, dass der Nutzer im Vorfeld der Nutzung ein korrektes mentales Modell aufbaut und ein korrektes Systemwissen erwirbt. Anschließend befasst sich die Checkliste mit Anforderungen zur Gestaltung der MMS, deren Erfüllung ein korrektes System- und Situationsbewusstsein sowie eine angemessene Systembedienung während der Nutzung begünstigen sollen. Im Kapitel zur Bewertung der Bedienungsanleitung werden die Darstellung der Inhalte „Systemfunktionalität“, „Systemmodi“, „Anzeigen“ und „Systembedienung“ sowie deren Verständlichkeit und das Vorhandensein einer Kurzanleitung anhand von Items geprüft. Die Bewertung der MMS gliedert sich in zwei übergeordnete Kapitel: Um die Checkliste einfach anwenden zu können, sind manche, die MMS betreffende Items, bestimmten Use Cases zugeordnet (Use Case spezifische Bewertung), die während der Überprüfung im Fahrzeug erlebt werden und deren Chronologie der Systeminteraktion entspricht (d. h. Aktivierung, Fahren mit aktivem System, Erlebens eines fahrerinitiierten Fahrstreifenwechsels, Erleben von Systemgrenzen, Deaktivierung etc.; siehe Kapitel 1.5 zur Herleitung der Use Cases). Weiterhin erfolgt im Kapitel „Generelle MMS-Anforderungen“ der Checkliste eine Use Case unabhängige Bewertung der Anzeige-Elemente und Systemausgaben hinsichtlich deren Anwesenheit, Lesbarkeit und Verständlichkeit sowie die Bewertung der Bedienelemente. Tabelle 2-1 gibt eine Übersicht über die Gliederung der Checkliste. Ein Auszug aus der Checkliste in englischer Sprache ist in Bild 2-1 zu sehen. Die Checkliste befindet sich im Anhang.



Category: General HMI requirements					
11	<b>Requirements for the display of system mode (e.g. ACC, lateral control, L2).</b>				
	11.1 System mode is indicated at all times.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	11.2 Indicators for different modes (e.g. ACC active, lateral control active; L2 active) are distinguishable from each other.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="width: 100%; text-align: center;"> <tr> <td style="width: 33%;">Fulfilled</td> <td style="width: 33%;">Satisfactory</td> <td style="width: 33%;">Not-acceptable</td> </tr> </table> <p>Comments:</p> <hr/> <hr/>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
12	<b>Requirements for legibility of visual information.</b>				
	12.1 Visual information is displayed close to driver's line of sight.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.2 Indicators relevant to the system are in spatial proximity of each other in each display.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.3 Colour of indicators contrasts adequately from background colour.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.4 Red/green and blue/yellow colour combinations are avoided.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.5 Text is easy to read (e.g. size, length).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.6 Flashing indicators are used only to attract attention when a warning message is issued.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="width: 100%; text-align: center;"> <tr> <td style="width: 33%;">Fulfilled</td> <td style="width: 33%;">Satisfactory</td> <td style="width: 33%;">Not-acceptable</td> </tr> </table> <p>Comments:</p> <hr/> <hr/>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
13	<b>Requirements for the comprehension of visual information.</b>				
	13.1 The colours confirm to conventions or stereotypes (for example green related to "system active"; red to "warning or danger").	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.2 The use of different colours is limited to as few as possible (not more than five).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.3 Text messages must be easy to understand (i.e. as short as possible, no technical language, language of user).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.4 Informative semantics is used for non-critical information, command style is used for safety critical information.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.5 Indicators are commonly accepted or standardized or are designed in a way that the underlying function can be assumed.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="width: 100%; text-align: center;"> <tr> <td style="width: 33%;">Fulfilled</td> <td style="width: 33%;">Satisfactory</td> <td style="width: 33%;">Not-acceptable</td> </tr> </table> <p>Comments:</p> <hr/> <hr/>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			

Bild 2-1: Auszug aus der Checkliste

Kapitel der Checkliste	Anforderungen bezüglich:
Bedienungsanleitung	<ul style="list-style-type: none"> <li>• Systemfunktionalität</li> <li>• Systemmodi und Anzeigen</li> <li>• Systembedienung</li> <li>• Verständlichkeit</li> <li>• Vorhandensein einer Kurzanleitung</li> </ul>
Use Cases	<ul style="list-style-type: none"> <li>• Aktivierung</li> <li>• Deaktivierung</li> <li>• Fahren mit aktivem System</li> <li>• Fahrerinitiiertes Fahrstreifenwechsel (falls diese Unterstützungsfunktion Bestandteil des Systems ist)</li> <li>• Systemgrenzen</li> </ul>
Generelle MMS-Anforderungen	<ul style="list-style-type: none"> <li>• Anwesenheit von Anzeige-Elementen</li> <li>• Lesbarkeit</li> <li>• Verständlichkeit</li> <li>• Bedienelemente</li> </ul>

Tab. 2-1: Übersicht über die Kapitel der Checkliste

## 2.2 Definition der Bewertungskategorien und -prinzipien

Die Items sind mit „ja oder „nein“ bzw. mit „nicht anwendbar“ zu beantworten. Die Option „nicht anwendbar“ wird dann vergeben, wenn das jeweilige Item nicht beurteilt werden kann, beispielsweise weil das System die entsprechende Funktionalität (z. B. Fahrstreifenwechsel) nicht enthält oder die zur Bewertung nötige Erfahrung des Iteminhalts nicht gemacht werden konnte (z. B. Systemgrenze). Die Bewertung orientiert sich somit am Antwortformat wie es z. B. im „Response Code of Practise“ (Response Consortium, 2006) oder der Checkliste von NAUJOKS et al. (2019c) genutzt wurde. Durch Urteile auf Itemebene wird die Entscheidung, ob eine übergeordnete Anforderung erfüllt ist oder nicht, getroffen. Die Bewertung erfolgt dabei nach folgender Regel:

- Fulfilled: Optimale Lösung; wenn alle Unteritems mit „ja“ beantwortet wurden.
- Satisfactory: Nicht perfekt, aber ausreichend; wenn maximal ein Unteritem mit „nein“ beantwortet wurde.
- Not acceptable: Nicht zufriedenstellend; wenn mehr als ein Unteritem mit „nein“ beantwortet wurde.

Die Checkliste enthält innerhalb jeder Anforderung ein Kommentarfeld (Comments), sodass eine Begründung für die Abgabe eines Urteils notiert werden kann. Nach Ausfüllen der Checkliste erfolgt die Abgabe des finalen Ratings zur Systembewertung.

Grundlage für das Urteil bildet der Erfüllungsgrad der 14 übergeordneten Anforderungen. Die Abgabe des Urteils erfolgt jedoch nach Ermessen der Rater, d. h. es wird bewusst auf die Vorgabe von Richtlinien für die Vergabe des Urteils verzichtet.

- Very good – no improvements necessary: Sehr gut – keine Verbesserungen notwendig,
- acceptable with minor options for improvement: Zufriedenstellend mit geringfügigen Möglichkeiten der Verbesserung,
- acceptable with major options for improvement: Zufriedenstellend mit großen Möglichkeiten der Verbesserung,
- not acceptable: Nicht zufriedenstellend.

## 2.3 Itemherleitung

Zur Auswahl der Checklistenitems wurde eine Literaturrecherche zu bereits bestehenden Standards, Normen, Guidelines und Gestaltungsrichtlinien sowie relevanten empirischen Studien durchgeführt. Zudem fanden Inhalte der Checkliste von NAUJOKS (2019c) zu Level 3 Systemen Eingang in die Checkliste. Der Schwerpunkt der L2-Checkliste liegt auf der Bewertung von visuell-auditiven Ausgaben einer MMS. Eine Neuerung ergibt sich im Hinblick auf die Berücksichtigung von Anforderungen zur Systeminteraktion sowie der Gestaltung des Benutzerhandbuchs. Außerdem werden Besonderheiten von teilautomatisierten Fahrzeugen wie beispielsweise die Notwendigkeit einer Fahrerüberwachung spezifiziert.



Im Folgenden werden die Items anhand von vier inhaltlichen Schwerpunktthemen vorgestellt. Die chronologische Item-Nummerierung der Checkliste, die sich aus der Zuordnung zu speziellen Use Cases zur einfacheren Anwendbarkeit während der Fahrzeugbewertung ergibt (siehe Kapitel 2.1 zum Aufbau der Checkliste), wird dadurch an manchen Stellen aufgebrochen. Die Checkliste ist mit dem Ziel der internationalen Anwendbarkeit in englischer Sprache formuliert. Zur besseren Verständlichkeit wurden die Items für diesen Bericht ins Deutsche übersetzt.

### 2.3.1 Anforderung an das Benutzerhandbuch

Da nicht davon ausgegangen werden kann, dass Fahrer beim Kauf oder der Miete eines Fahrzeugs umfassend über alle Fahrzeugfunktionen (inkl. deren Grenzen und der intendierten Nutzung) aufgeklärt werden (ABRAHAM et al., 2018; CASNER & HUTCHINS, 2019), muss das Benutzerhandbuch in der Lage sein, dem Fahrer die wichtigsten Punkte bezüglich der Systembedienung, des Systemverhaltens und der Systemgrenzen in einer leicht verständlichen Form zu vermitteln. Nutzungsinformationen sind laut DIN EN 82079-1 (S. 11) „für den sicheren Gebrauch eines Produkts unerlässlich“, „für den effizienten und effektiven Gebrauch eines Produkts nützlich“ und „zur Erfüllung marktrechtlicher und behördlicher Verpflichtungen erforderlich.“ Auch können dadurch eventuell bestehende Falschannahmen korrigiert werden, wodurch sich Fehlgebrauch verringern lässt. Dass die Güte der Beschreibung von automatisierten Fahrzeugfunktionen einen Einfluss auf die Qualität der Interaktion mit diesen Systemen hat, konnten FORSTER et al. (2019b) in einer Untersuchung zeigen, bei der sich das Verständnis und das Bedienverhalten der Probanden nach Lesen der Gebrauchsanweisung im Vergleich zu einer generischen Information verbesserten.

Grundsätze zu Inhalt und Gestaltung von Benutzerhandbüchern finden sich beispielsweise in der DIN EN 82079-1 zum „Erstellen von Gebrauchsanleitungen – Gliederung, Inhalt und Darstellung“, der Neufassung des Europäischen Grundsatzkatalogs zur Mensch-Maschine-Schnittstelle über sichere und effiziente bordeigene Informations- und Kommunikationssysteme der Europäischen Kommission (2008) sowie im „Response Code of Practice“ (Response Consortium, 2006). Angaben über die minimal notwendigen Informationen, die Fahrer von teil-

automatisierten Fahrzeugen bekommen sollten, geben auch CASNER & HUTCHINS (2019). Im Folgenden ist die Herleitung der Items dargestellt. Eingangs werden immer die übergeordneten Anforderungen, die das Benutzerhandbuch erfüllen sollte, dargestellt.

### Beschreibung der Systemfunktionalität

Zur sicheren Anwendung des Systems muss der Fahrer die Funktionsweise, den Anwendungsbereich sowie die Grenzen des Systems kennen. Die Beschreibung der korrekten Funktionsweise kann verhindern, dass Fahrer Fehler bei der Benutzung machen (BEGGIATO & KREMS, 2013). Um zu verhindern, dass Fahrer, die Automatisierungsfunktion in nicht geeigneten Situationen nutzen, müssen Fahrer den vorgesehenen Anwendungsbereich kennen. Andernfalls werden die Systeme beispielsweise auf nicht dafür vorgesehenen Strecken verwendet (DICKIE & BOYLE, 2009). Auch um im Fall von plötzlich auftretenden Systemgrenzen eine angemessene Reaktion zeigen zu können, sollten diese Aspekte vorher bekannt sein. In empirischen Studien konnte gezeigt werden, dass das Wissen über Systemgrenzen Reaktionszeiten verringern kann (HERGETH, LORENZ & KREMS, 2017). Neben unzureichenden Reaktionen kann sich die Unkenntnis über Systemgrenzen auch negativ auf die Akzeptanz und das Systemvertrauen auswirken. BEGGIATO & KREMS (2013) konnten zeigen, dass das Erleben von unbekanntem Systemgrenzen bei ACC über mehrere Termine zu einer konstanten Abnahme von Akzeptanz und Vertrauen der Funktion führte. Auch BLÖMACHER, NÖCKER & HUFF (2018) stellten fest, dass beim Vergleich verschiedener Instruktionen zu L3-Fahrzeugen eine inkorrekte Instruktion ein falsches mentales Modell zur Folge hatte, woraus Probleme beim Verständnis der Verkehrssituation resultierten. Die EU-Kommission (2008) fordert, dass die Systemfunktionalität in den Nutzungsinformationen genau beschrieben wird (Grundsatz V für Informationen zum System). Auch die Checkliste des „Response Code of Practice“ (Response Consortium, 2006; Kapitel A.2.9, Item 17, S. A33) fordert ebenso wie die DIN EN 82079-1 die Beschreibung der Systemfunktion und Systemgrenzen im Benutzerhandbuch. Die folgenden Items wurden aus den Ausführungen abgeleitet:

- „Das System wird als Assistenz- oder Komfortsystem beschrieben, das Teile der Fahraufgabe

ausführt, d. h. sowohl die Quer- als auch die Längsführung.“ (Item 1.1)

- „Das Handbuch beschreibt die Bedingungen, unter denen die Nutzung des Systems vorgesehen ist (z. B. Straßentyp, Verkehr, Beschränkungen durch Umweltbedingungen).“ (Item 1.2)
- „Das Handbuch beschreibt Systemgrenzen und/oder Gründe für eine eingeschränkte Systemfunktionalität und deren Folgen.“ (Item 1.3)

Während der Nutzung eines teilautomatisierten Systems ist der Fahrer weiterhin für die Fahraufgabe verantwortlich (SAE, 2018). Entsprechend des Wiener Übereinkommens der Vereinten Nationen von 1968 muss ein Fahrer jederzeit in der Lage sein, sein Fahrzeug zu beherrschen. Der Fahrer ist somit für die Überwachung des Systems zuständig, um im Falle von plötzlich auftretenden Systemgrenzen sofort eingreifen zu können. Aufgrund auftretender Probleme mit Mode Confusion (BANKS, ERIKSSON, O'DONOGHUE & STANTON, 2018a; LARSSON, 2012) ist es zu empfehlen, dem Fahrer Informationen zu seiner Rolle und Verantwortung im Benutzerhandbuch zu vermitteln.

In diesem Zusammenhang fordert die EU-Kommission (2008) in Grundsatz VII für Informationen zum System (S. 30): „Darstellungen zur Nutzung des Systems (z. B. Beschreibungen, Fotos und Skizzen) sollten weder unrealistische Erwartungen bei potenziellen Nutzern begründen, noch zu einer Nutzung Anlass geben, von der eine Gefährdung ausgehen kann.“ Dieser Grundsatz ist auch in der Checkliste des „Response Code of Practice“ (Response Consortium, 2006) mit der Forderung danach, dass die Systeminformationen realistische Erwartungen generieren sollen, abgebildet (Kapitel A2.14, Item 7). Auch die Alliance of Automobile Manufacturers (AAM; 2006) stellt diese Forderung. In einer Studie von NEES (2018) konnte gezeigt werden, dass Fahrer unter anderem bei der Funktionsbezeichnung „Autopilot“, wie sie beispielsweise vom Hersteller Tesla für das L2-System verwendet wird, die Verantwortung für die Fahrzeugkontrolle mehrheitlich beim Fahrzeug sehen. Dadurch besteht die Gefahr, dass Fahrer den Funktionsumfang überschätzen. Die folgenden Items bilden die Anforderungen an die Kommunikation der Verantwortlichkeit des Fahrers im Handbuch ab:

- „Das Handbuch informiert den Fahrer, dass er dafür verantwortlich ist, das System permanent

zu überwachen, um bei Systemgrenzen oder Fehlfunktionen eingreifen zu können.“ (Item 1.4)

- „Durch die Benennung der Funktion werden keine unangemessenen Erwartungen hervorgerufen.“ (Item 1.5)

### **Beschreibung der Systembedienung**

Damit der Fahrer in der Lage ist, das System sicher im Straßenverkehr zu bedienen, ist die Kenntnis der Systembedienung und der dazu benötigten Steuerelemente erforderlich. Das Benutzerhandbuch sollte deshalb die Benutzung des Systems in angemessener Weise beschreiben (DIN EN 82079-1). Die Erklärung der Systembedienung wird auch in der Checkliste des „Response Code of Practice“ (Response Consortium, 2006; Kapitel A.2.9, Item 17, S. A33) gefordert. Für ein L2-System fallen darunter die Interaktionen, die für die erfolgreiche Durchführung der unter Kapitel 1.5 dargestellten Use Cases wie „Aktivierung“ oder „Deaktivierung“ des Systems erforderlich sind. Sofern innerhalb eines Fahrzeugs verschiedene Stufen der Automation existieren (z. B. L1 und L2) und sich das L2-System auch aus verschiedenen Ausgangszuständen (z. B. manuell vs. ACC) aktivieren lässt bzw. den Wechsel zu anderen Automationsstufen (d. h. zu anderen SAE-Level, z. B. ACC) zulässt, soll die Systemaktivierung und -deaktivierung auch unter dem Aspekt des Wechsels der Automationsstufe erläutert werden. Sofern das System einen automatisierten Fahrstreifenwechsel beinhaltet (in der Checkliste als „Spurwechsel“ bezeichnet), sollen auch die dazu notwendigen Bedienhandlungen beschrieben werden. Da es bei L2-Funktionen je nach Auslegung zu Systemgrenzen, bei denen der Fahrer eingreifen muss, kommen kann (FAVARÒ, EURICH & NADER, 2018), ist es wichtig, dass der Fahrer entsprechendes Wissen darüber hat, wie die Kontrolle wiederzuerlangen ist, wie das System übersteuert werden kann und welche Steuerelemente dazu benötigt werden. Auch aus diesem Grund sollten die dazu benötigten Kontrollelemente (inkl. Deaktivierung und Übersteuerung) im Handbuch beschrieben werden (DIN EN 82079-1). Die folgenden Items fassen die genannten Anforderungen zusammen:

- „Das Handbuch beschreibt, wie das System aktiviert werden kann und welche Voraussetzungen für die Aktivierung erfüllt sein müssen.“ (Item 2.1)

- „Das Handbuch beschreibt, wie das System deaktiviert oder übersteuert werden kann.“ (Item 2.2)
- „Das Handbuch beschreibt, wie der Fahrer zwischen verschiedenen Systemmodi wechseln kann.“ (Item 2.3)
- „Wenn implementiert, beschreibt das Handbuch, wie ein Spurwechsel durchgeführt werden kann, wie das System sich währenddessen verhält und welche Verantwortlichkeiten der Fahrer hat.“ (Item 2.4)
- „Das Handbuch beschreibt die notwendigen Bedienelemente, um die oben genannten Aktionen auszuführen.“ (Item 2.5)

### Beschreibung der Systemmodi und Anzeigen

Zur Förderung eines vollständigen Systemverständnisses ist eine Nennung und Beschreibung aller Systemmodi sowie der dazugehörigen Indikatoren und Anzeige-Elemente im Benutzerhandbuch enthalten (DIN EN 82079-1). Dazu gehören neben den Zuständen „aktiv“, „inaktiv“ und der „Standby-Funktion“ auch weitere Systemzustände wie beispielsweise „ACC aktiv“, sofern diese im Systemumfang enthalten sind. Die positive Wirkung von zusätzlichen Illustrationen gegenüber rein textlicher Darstellung wurde in der Vergangenheit vielfach empirisch untersucht (siehe HINDER; 1982; LEVIE & LENTZ, 1982; CARNEY & LEVIN, 2002). Die textliche Beschreibung von Systemzuständen sollte daher immer durch Abbildungen der vom System verwendeten Zustandsindikatoren ergänzt werden (EU-Kommission, 2008; DIN EN 82079-1). Gleichzeitig sollten Informationen über den Ort der Anzeigen und Zustandsindikatoren enthalten sein (z. B. Art des verwendeten Displays und Position der Anzeige). Die Anforderungen können mithilfe der folgenden Items abgeprüft werden:

- „Das Handbuch beschreibt alle verschiedenen Systemmodi und ihre relevanten Indikatoren.“ (Item 3.1)
- „Zur Veranschaulichung werden informative Illustrationen verwendet.“ (Item 3.2)
- „Das Handbuch nennt die relevanten Displays, die der Fahrer beachten muss (z. B. Clusteranzeige und HUD).“ (Item 3.3)

### Verständlichkeit des Handbuchs

Damit es dem Handbuch gelingt, beim Benutzer ein angemessenes Systemverständnis zu generieren, sollten die Informationen einfach und verständlich vermittelt werden (EU-Kommission, 2008; Grundsatz II für Informationen zum System, S. 27: „Systembezogene Anleitungen sollten zutreffend und einfach sein.“). Grundlegend beinhaltet eine gute Verständlichkeit, dass die Informationen in der jeweiligen Landessprache des Nutzers formuliert sind und auf die Verwendung einer technischen Sprache oder Fachausdrücken verzichtet wird (Grundsatz II bzw. III der EU-Kommission, 2008; AAM, 2006). Auch die Checkliste des „Response Code of Practice“ enthält ein Item (Kapitel A.2.9, Item 17, S. A33), welches die verständliche Formulierung von Informationen zu Systemfunktionen, -bedienung und -grenzen fordert. Neben der zusätzlichen Verwendung von Abbildungen oder Illustrationen (EU-Kommission, 2008; DIN EN 82079-1) wird das sprachliche Verständnis durch die Verwendung von kurzen, auf relevante Informationen beschränkte Sätze gefördert (DIN EN 82079-1). Auch auf eine übersichtliche Struktur, welche die relevanten Informationen in übersichtliche Abschnitte unterteilt und bei Bedarf Verweise auf andere relevante Kapitel enthält, sollte geachtet werden (DIN EN 82079-1). Daraus ergeben sich die folgenden Items:

- „Das Handbuch verwendet einfach Sätze moderater Länge.“ (Item 4.1)
- „Das Handbuch verwendet eine nicht-technische Sprache. Wenn technische Begriffe verwendet werden, werden sie erklärt.“ (Item 4.2)
- „Die Kapitel sind gut strukturiert und durch Überschriften getrennt.“ (Item 4.3)
- „Verweise zu anderen relevanten Kapiteln (z. B. ACC) werden angeführt.“ (Item 4.4)

### Vorhandensein einer Kurzanleitung

Häufig informieren sich Fahrer im Vorfeld der Nutzung von assistierten oder automatisierten Systemen nicht mithilfe des Benutzerhandbuchs über die Systembedienung (EU-Kommission, 2008; MEHLENBACHER, WOGALTER & LAUGHERY, 2002). Laut EU-Kommission (2008, S. 27) wird dieser Zustand „...durch die unzulängliche Gestaltung der Anleitungen noch verschärft.“ (EU-Kommission, 2008, S. 27, Kapitel 4.3.6.2). Durch die Bereitstellung einer Kurzanleitung, in der die wichtigsten In-

formationen zur Systemnutzung und -bedienung übersichtlich zusammengefasst werden, sollte dieser Tatsache entgegengewirkt werden. Aus diesem Grund stellt die Checkliste folgende Forderung:

- „Für den schnellen Zugriff steht eine Kurzanleitung zur Verfügung.“ (Item 5.1)

Wenn jedoch davon ausgegangen werden muss, dass ein Nutzer sich vor der erstmaligen Systemnutzung lediglich an einer Kurzanleitung orientiert, müssen an die Gestaltung der Kurzanleitung ähnliche Maßstäbe gelegt werden, wie an die Gestaltung des ausführlichen Benutzerhandbuches. Daraus folgt, dass auch in der Kurzanleitung alle relevanten Informationen zur Nutzung des Systems enthalten sein müssen, diese gut verständlich dargestellt sind und eindeutig kenntlich gemacht wird, an welcher Stelle des Benutzerhandbuches ausführlichere Informationen zu erhalten sind. Diese Forderungen finden sich in den folgenden Items wieder:

- „Wenn eine Kurzanleitung verfügbar ist, sind alle für die Systemnutzung erforderlichen Informationen enthalten.“ (Item 5.2)
- „Wenn eine Kurzanleitung verfügbar ist, werden die Informationen verständlich dargestellt.“ (Item 5.3)
- „Wenn eine Kurzanleitung verfügbar ist, wird auf die entsprechenden Kapitel des Benutzerhandbuchs verwiesen.“ (Item 5.4)

### 2.3.2 Anforderungen an die Ausgaben der Mensch-Maschine-Schnittstelle

Den Ausgaben der MMS kommt die Aufgabe zu, eine sichere Interaktion mit dem System zu gewährleisten und dem Fahrer während der Systemnutzung durch die Angabe des Systemzustands jederzeit seine Rolle zu vermitteln. Werden weder das Handbuch noch eine Kurzanleitung vor der erstmaligen Systembenutzung gelesen, stellt die MMS die einzige Informationsquelle dar. Ziel der Checkliste ist es, anhand von Items die Einhaltung generell akzeptierter Prinzipien der MMS-Gestaltung zu überprüfen, die eine sichere Systemnutzung dadurch begünstigen, indem Informationen über den aktuellen Systemstatus und Systemzustandsänderungen effektiv kommuniziert werden (NAUJOKS et al., 2019c). Für die Ausgaben der MMS existieren Mindestanforderungen, die in europäischen Normen und allgemein akzeptierten Gui-

delines zu In-Vehicle-Information Systemen (IVIS), festgehalten sind. Dabei handelt es sich um generische Anforderungen, die für MMS vielfältiger Art generelle Gültigkeit besitzen, und deswegen auch bereits in die L3-Checkliste von NAUJOKS et al. (2019c) Eingang gefunden haben. Neben Referenzen, die explizit im Automobilkontext entstanden sind, werden für die Itemableitung auch Referenzen herangezogen, die sich im weiteren Sinne auf die Gestaltung von MMS beziehen und auf das automatisierte Fahren anwendbar sind. Weiterhin werden Ergebnisse empirischer Studien ergänzend herangezogen. Die abgeleiteten Items beziehen sich vor allem auf die Darstellung visueller Anzeigen der MMS, da diese den aktuellen Standard (Stand: 2019) von auf dem Markt verfügbaren Fahrzeugen mit L2-Funktion darstellen. Viele Hersteller (z. B. Audi, BMW, Daimler, und Tesla) nutzen für ihre L2-Systeme eine visuell-auditive MMS, die den aktiven Systemstatus anzeigt.

Die nachfolgenden Kapitel sind nach den Prozessen der menschlichen Informationsverarbeitung strukturiert: Neben der Grundvoraussetzung, dass über den aktuellen Systemstatus informiert wird (Anwesenheit von Statusinformationen), müssen die Indikatoren wahrgenommen und erkannt (Lesbarkeit von Anzeigen) sowie korrekt interpretiert werden können (Verständlichkeit von Anzeigen; NAUJOKS et al., 2019c).

#### Anwesenheit von Statusinformationen

Die Anzeige des Systemmodus ist relevant für die Kommunikation der Verantwortung des Fahrers. Während es beim hochautomatisierten Fahren (L3-Automatisierung) wichtig sein wird, dem Fahrer zu übermitteln, wann er die Kontrolle wieder übernehmen muss, ist es beim teilautomatisierten Fahren (L2) bedeutsam, dem Fahrer zu vermitteln, dass er während der Systemnutzung weiterhin für die Fahraufgabe und die Überwachung des Systems verantwortlich ist, auch wenn kein kontinuierlicher Kontakt mit den Stellteilen des Fahrzeugs (z. B. Fahrpedal) besteht. Da die (visuell-auditive) MMS bei Fahrzeugen mit L2-Systemen aktuell den primären Informationskanal darstellt, ist die Anzeige des Systemstatus essenziell, um den Fahrer permanent über den Modus der Automation zu informieren (STEVENS, QUIMBY, KERSLOOT & BURNS, 2002; Response Consortium, 2006; AAM, 2006) und ihm seine Rolle zu kommunizieren (Adaptive Consortium, 2017).

Häufig existieren im selben Fahrzeug mehrere Assistenzsysteme. Gerade Systeme mit mehreren Automationsstufen (z. B. L1, L2 und L3) bergen das Problem der Mode Confusion (FELDHÜTTER, HÄRTWIG; KURPIERS, MEJIA HERNANDEZ & BENGLER, 2018). Deshalb ist es wichtig, dass jede Automationsstufe ihren eigenen Modus-Indikator hat, um den aktuellen Systemzustand jederzeit identifizieren zu können (Adaptive Consortium, 2017, NAUJOKS et al., 2019c). Um Zustandsverwechselungen, die zu potenziell kritischen Fahrsituationen führen können, zu vermeiden, müssen die Indikatoren voneinander unterscheidbar sein (Adaptive Consortium, 2017; NAUJOKS et al., 2019c; Response Consortium, 2006). Neben der reinen Anwesenheit von Mode-Indikatoren fordern NAUJOKS et al. (2019c), dass Zustandsänderungen effektiv kommuniziert werden. Dies ist von besonderer Bedeutung, wenn die Transition nicht vom Fahrer sondern durch das System initiiert wird. Unter der Annahme, dass eine fahrerinitiierte Transition meistens bewusst herbeigeführt wird, erwartet der Fahrer gewöhnlich die Änderung des Systemzustands. Dies ist bei systeminitiierten Änderungen nicht der Fall. Der temporäre Verlust der Querführung, ohne dass das System durch den Nutzer deaktiviert wurde (sog. Stand-by-Modus), und der anschließenden selbstständigen Reaktivierung, stellt einen Sonderfall beim teilautomatisierten Fahren dar. In diesem Fall ist die effektive Kommunikation der Zustandsänderung besonders wichtig, damit der Fahrer entnehmen kann, ob das System tatsächlich aktiv regelt. Anhand der folgenden Items werden die Anforderungen überprüft:

- „Der Systemmodus wird immer angezeigt.“ (Item 8.1 und 11.1).
- „Die Indikatoren der Systemmodi sind voneinander unterscheidbar.“ (Item 11.2)
- „Eingeschränkte Systemfunktionalität ist klar von der vollen Systemfunktionalität zu unterscheiden.“ (Item 8.3)

Die Änderungen der Modus-Indikatoren aufgrund einer Bedienung durch den Nutzer sollten ohne zeitliche Verzögerung stattfinden (EU-Kommission, 2008; Department of Defense, 2012; AAM, 2006). Die empfohlene Zeitspanne, die dabei nicht überschritten werden sollte, liegt zwischen 100 ms (Department of Defense, 2012), 200 ms (BUBB, BENGLER, GRÜNEN & VOLLRATH, 2015) oder 250 ms (STEVENS et al., 2002). Wenn eine längere Verzögerung unvermeidbar ist, sollte eine Rückmeldung

darüber erfolgen, dass das System die Eingabe des Fahrers registriert hat (NAUJOKS et al., 2019c; STEVENS et al., 2002). Bis eine Systemantwort verfügbar ist, sollte der Fahrer über den Fortschritt des beabsichtigten Moduswechsels informiert werden (Adaptive Consortium, 2017; BUBB et al., 2015; NAUJOKS et al., 2019c). Die folgenden Items überprüfen diese Anforderungen:

- „Die erfolgreiche Aktivierung oder Deaktivierung des Systems oder des Spurwechsels wird dem Fahrer gleichzeitig zur Fahrereingabe rückgemeldet.“ (Aktivierung: Item 6.3; Deaktivierung: Item 7.3; Spurwechsel: Item 9.2 und 9.4)
- „Eine vorhandene Verzögerung zwischen Fahrerinput und der Aktivierung des Systems oder des Spurwechsels oder eines Fehlschlagens der Aktivierung wird angezeigt.“ (Aktivierung: Item 6.4 und Spurwechsel: Item 9.3)

Wenn das System oder Teile davon nicht mehr fehlerfrei funktionieren (z. B. aufgrund eines Sensorfehlers) und daraufhin ein systemseitiges Abschalten des Systems erfolgt, muss dies durch einen Hinweis oder eine Warnung angezeigt werden. Um eine effektive Kommunikation sicherzustellen, müssen Anzeigen, die nicht permanent vorhanden sind, mindestens so lange angezeigt werden, dass sie vom Fahrer wahrgenommen werden können (NAUJOKS et al., 2019c). Meldungen, die den Fahrer über die Deaktivierung des Systems informieren, müssen von weniger dringlichen Informationen in der MMS klar unterscheidbar sein (Japan Automobile Manufacturers Association, 2004, NAUJOKS et al., 2019c). Wichtig ist hierbei, dass eine Warnung zwar die Aufmerksamkeit des Fahrers auf sich zieht, der Fahrer bei der Ausgabe einer Warnung jedoch nicht erschreckt wird (NAUJOKS et al., 2019c; ROSS et al., 1996). Die folgenden Items bilden diese Anforderungen ab:

- „Wenn das System oder Teile des Systems nicht mehr fehlerfrei funktionieren, wird ein Hinweis/Warnung ausgegeben.“ (Item 10.1)
- „Ein Hinweis/Warnung wird lange genug angezeigt, um vom Fahrer bemerkt zu werden.“ (Item 10.2)
- „Durch die Ausgabe einer Warnung wird der Fahrer nicht erschreckt.“ (Item 10.4)

## Lesbarkeit von Anzeigen

Die EU-Kommission (2008) legt in ihrem Grundsatz I für die Darstellung von visuellen Informationen folgendes fest (S. 13, Kapitel 4.3.3.1): „Die von einem System zu einem beliebigen Zeitpunkt optisch angezeigten Informationen sollten so gestaltet werden, dass der Fahrer die betreffenden Informationen mit wenigen Blicken erfassen kann, die kurz genug sein müssen, um das Fahrverhalten nicht zu beeinträchtigen.“ Um die Lesbarkeit von Anzeigen sicherzustellen, sollten mehrere Grundsätze verfolgt werden: Displays, die visuelle Informationen transportieren, sollten so nah wie möglich an der normalen Sichtlinie des Fahrers angebracht werden (AAM, 2006; NAUJOKS et al., 2019c; ROSS et al., 1996). Weiterhin sollten Anzeigen, die zu einer gemeinsamen Funktion gehören (in diesem Fall beispielsweise die Anzeige von ACC und Querverführung) zusammen gruppiert werden (NAUJOKS et al., 2019c; STEVENS & CYNK, 2011). Zudem ist der Kontrast zwischen der relevanten Information und dem Hintergrund ausschlaggebend für eine gute Lesbarkeit (NAUJOKS et al., 2019c). Gemäß DIN 15008:2003 soll der minimale Kontrast zwischen einer Anzeige und dem Hintergrund bei Tageslichtverhältnissen mindestens 3:1, bei Nachtbedingungen 5:1 ausmachen (BUBB et al., 2015). Bei der Verwendung verschiedener Farben ist darauf zu achten, dass sich diese gut voneinander unterscheiden lassen (GREEN, LEVISON, PAELKE & SERAFIN, 1994). Da gewisse Farbkombinationen einer guten Lesbarkeit entgegenstehen, sollten Farbkombinationen aus rot und grün, sowie blau und gelb vermieden werden (NAUJOKS et al., 2019c; STEVENS & CYNK, 2011). Für die Lesbarkeit von textlichen Anzeigen spielen vor allem Schriftgröße und Schriftart eine Rolle (NAUJOKS et al., 2019c). Für Buchstaben empfiehlt sich eine Mindesthöhe von 6.4 mm (GREEN et al., 1994). Dabei sollte eine Mischung aus Groß- und Kleinbuchstaben genutzt werden (ROSS et al., 1996). Um ein klares Schriftbild zu erreichen, sollte auf die Verwendung einer serifenlosen Schriftart (z. B. Arial, Helvetica) geachtet werden (STEVENS et al., 2002; Department of Defense, 2012; NAUJOKS et al., 2019c). Die Verwendung von blinkenden Anzeigen sollte auf kritische Ereignisse, in denen die Aufmerksamkeit des Fahrers geweckt werden soll, beschränkt sein (DIN 15008:2003; Department of Defense, 2012; STEVENS et al., 2002). Die folgenden Items sollen die Überprüfung der Lesbarkeit der Anzeigen ermöglichen:

- „Visuelle Informationen werden nahe der Sichtlinie des Fahrers angezeigt.“ (Item 12.1)
- „Systemrelevante Anzeigen befinden sich in räumlicher Nähe zueinander.“ (Item 12.2)
- „Die Farbe von Anzeigen weist einen angemessenen Kontrast zur Hintergrundfarbe auf.“ (Item 12.3)
- „Die Farbkombinationen rot/grün und blau/gelb werden vermieden.“ (Item 12.4)
- „Text ist gut lesbar (z. B. hinsichtlich Größe und Länge).“ (Item 12.5)
- „Blinkende Darstellungen werden nur verwendet, um auf eine Warnmeldung aufmerksam zu machen.“ (Item 12.6)

## Verständlichkeit von Anzeigen

Neben der reinen Lesbarkeit ist es wichtig, dass der Inhalt der durch die MMS vermittelten Informationen verständlich ist, damit der Fahrer schnell ableiten kann, ob bzw. welche Konsequenzen die jeweiligen Anzeigen für sein Handeln bedeuten (Adaptive Consortium, 2017). Erfordert die Dekodierung der relevanten Anzeigen zu viel Aufmerksamkeit, wird dadurch die eigentliche Aufgabe des Fahrers – die Überwachung der Fahrsituation – zu stark beeinträchtigt. Generell gilt, dass die Gestaltung von Informationen (z. B. Hinweise vs. Warnungen) an deren Dringlichkeit angepasst werden sollte (NAUJOKS et al., 2019c). Beispielsweise sollten Warnungen auffälliger sein als die reine Mitteilung einer Information (Adaptive Consortium, 2017; NAUJOKS et al., 2019c). Dies kann z. B. über die farbliche Gestaltung von Icons und Symbolen, Multimodalität oder den Inhalt einer Nachricht geschehen.

Zur schnelleren Verarbeitung sollten sicherheitskritische Informationen, die eine sofortige Fahrerreaktion erfordern, multimodal dargeboten werden (ROSS et al., 1996; GREEN et al., 1994; NAUJOKS et al., 2019c). Visuelle Warnungen sollten dementsprechend mit einem akustischen Warnton gekoppelt werden (CAMPBELL, CARNEY & KANTOWITZ, 1998; NAUJOKS, MAI & NEUKUM, 2014). Eine schnelle Informationsverarbeitung, die nicht das Lesen von textlichen Informationen erfordert, kann durch die Verwendung von Farben, die gängigen Konventionen und Stereotypen entsprechen, erzielt werden (NAUJOKS et al., 2019c). So wird häufig der aktive Systemstatus mit grün angezeigt,

während rot häufig auf eine Gefahr hindeutet (BUBB et al., 2015; ROSS et al., 1996; NAUJOKS et al., 2019c; STEVENS & CYNK, 2011). Bei der Farbwahl sollte jedoch berücksichtigt werden, dass nicht mehr als fünf verschiedenen Farben – ausgenommen Schwarz und Weiß – verwendet werden, um einer Verwechslung der Farben vorzubeugen (NAUJOKS et al., 2019c; STEVENS et al., 2002). Zu berücksichtigen ist dabei, dass Text generell nicht farbig angezeigt werden sollte, da farbiger Text die Leszeit erhöht (STEVENS et al., 2002). Bei der Verwendung von Textmeldungen sollten für eine kurze Blickabwendungszeit von der Straße weiterhin darauf geachtet werden, dass diese prägnant und verständlich formuliert sind (NAUJOKS et al., 2019c). Dazu sollten sie nicht zu lang sein (CAMPBELL, CARNEY & KANTOWITZ, 1998), keine technischen Begriffe enthalten (GREEN et al., 1994) sowie in der jeweiligen Landessprache des Nutzers präsentiert werden (ROSS et al., 1996). Bezüglich des Inhalts der Nachricht sollte bei textlichen Aussagen darauf geachtet werden, für unkritische Informationen auf informative Semantik zurückzugreifen, während bei sicherheitskritischen Informationen der Aufforderungsstil zu verwenden ist (STEVENS et al., 2002). Die Auswahl von Symbolen oder Zustandsindikatoren sollte erfolgen, indem standardisierte oder allgemein akzeptierte Symbole oder Anzeigen verwendet werden (CAMPBELL, CARNEY & KANTOWITZ, 1998; AAM, 2006; NAUJOKS et al., 2019c). Beim Gebrauch von neuartigen Symbolen sollten diese von einer textlichen Beschreibung begleitet sein (NAUJOKS et al., 2019c; STEVENS et al., 2002). Die folgenden Items berücksichtigen diese Anforderungen:

- „Hinweise/Warnungen sind an die Dringlichkeit der Situation angepasst (hinsichtlich der Verwendung von Farben, zusätzlichem Sound (d. h. Multimodalität) sowie Inhalt der Nachricht.“ (Item 10.3)
- „Die Auswahl der Farben entspricht gängigen Konventionen oder Stereotypen (z. B. Grün bezogen auf „System aktiv“; Rot auf „Warnung oder Gefahr“).“ (Item 13.1)
- „Die Verwendung unterschiedlicher Farben ist auf möglichst wenige beschränkt (nicht mehr als fünf).“ (Item 13.2)
- „Textmeldungen müssen verständlich formuliert sein (nicht zu lang, keine technischen Begriffe, in Landessprache).“ (Item 13.3)

- „Informative Semantik wird für unkritische Informationen verwendet, der Aufforderungsstil für sicherheitskritische Informationen.“ (Item 13.4)
- „Die verwendeten Indikatoren sind allgemein akzeptiert oder standardisiert oder sind so gestaltet, dass auf die zugrunde liegende Funktion geschlossen werden kann.“ (Item 13.5)

### 2.3.3 Anforderungen an die Systembedienung

Neben den Anforderungen an die visuellen Ausgaben des Systems gehören zur MMS auch die zur Systembedienung erforderlichen Interaktions- und Bedienhandlungen. Für eine sichere Nutzung sollte das Bediensystem der Teilautomation eine intuitive, einfache und fehlerrobuste Verwendung der Steuer- und Bedienelemente fördern. In diesem Zusammenhang wird in den „Human Factors Design Guidelines für Fahrerinformationssysteme“ von UMTRI (GREEN et al., 1994) folgender Grundsatz (S. 11) gefordert: „Operations that occur most often or have the greatest impact on driving safety should be the easiest to perform“. Zu den häufig durchgeführten Bedienhandlungen gehören beim teilautomatisierten Fahren die Aktivierung (sowohl aus dem manuellen als auch aus niedrigeren Automationsstufen wie beispielsweise ACC heraus, wenn diese im System getrennt vorhanden sind) sowie die Deaktivierung entweder des Gesamtsystems oder einzelner Teilfunktionen des Systems (z. B. nur Quерführung). Von einer einfachen Bedienung kann dann ausgegangen werden, wenn der Nutzer in der Lage ist, den Zielzustand des Systems mithilfe von wenigen und einfach auszuführenden Bediensritten zu erreichen (FORSTER et al., 2019a). Je weniger Bediensritte beispielsweise zur Erreichung des Zielzustands notwendig sind, umso besser ist die Systembedienung auch unter dem Aspekt der Beanspruchung zu bewerten (PURUCKER; NAUJOKS, PRILL & NEUKUM, 2017), da lange Bedienwege ablenkend wirken können. STEVENS & CYNK (2011) fordern in ihrer „Checklist for the assessment of in-vehicle information systems“, dass alle relevanten Bedienelemente aus der normalen Fahrposition erreicht werden können, ohne, dass umständliche Bewegungen notwendig sind. Zudem sollen sich zur intuitiven Nutzung Bedienelemente, die für die Systembedienung notwendig sind – wie die dazugehörigen Anzeigenelemente auch – in räumlicher Nähe zueinander befinden (NASA, 2015; GREEN et al., 1994). Eine einfache und intuitive Bedienung wird auch durch die Forderung

nach Beschriftung von Bedienelementen erzielt (NASA, 2015; Department of Defense, 2012). Die Beschriftungen sollen dabei aus der normalen Fahrposition gut lesbar, d. h. nicht durch die Handposition verdeckt werden (AAM, 2006; Department of Defense, 2012) und leicht verständlich sein. Die Verständlichkeit kann durch die Konsistenz zwischen Systemanzeigen und Beschriftungen auf den Bedienelementen gefördert werden (GREEN et al., 1994). Wie auch die Systemanzeigen müssen die Beschriftungen untereinander gut voneinander unterscheidbar sein (NASA, 2015). Dies gilt auch im Hinblick auf die Abgrenzung zu anderen Systemen. Die Anforderung der Unterscheidbarkeit ist auch auf die Bedienlogik zur Erreichung verschiedener Systemmodi oder -stufen, anzuwenden (NASA, 2015). Weiterhin sollen die Bedienlogik oder die Anordnung der Bedienelemente gängigen Konventionen entsprechen. Beispielsweise wird die Drehung eines Kontrollelements im Uhrzeigersinn mit „Hoch“, „Rechts“ oder „Steigern“ verbunden (STEVENS et al., 2002). Die folgenden Items dienen zur Überprüfung der Anforderung an die Systembedienung:

- „Die Aktivierung und Deaktivierung des Systems, Teilfunktionen des Systems oder des Spurwechsels ist einfach durchzuführen.“ (Aktivierung: Item 6.1 und 6.2; Deaktivierung: Item 7.1 und 7.2; Spurwechsel: Item 9.1)
- „Die Bedienelemente sind aus der normalen Fahrposition gut zu erreichen.“ (z. B. Multifunktionslenkrad; Item 14.1)
- „Systemzugehörige Bedienelemente (z. B. Tasten) sind in Gruppen zusammengefasst.“ (Item 14.2)
- „Die Bedienelemente sind beschriftet.“ (Item 14.3)
- „Die Beschriftungen sind aus der normalen Fahrposition sichtbar und leicht zu verstehen.“ (Item 14.4)
- „Die Bedienlogik oder die Anordnung der Bedienelemente (z. B. obere Taste  $\triangleq$  +; untere Taste  $\triangleq$  -) ist erwartungsgemäß und entspricht gängigen Konventionen.“ (Item 14.5)
- „Bedienelemente oder die erforderliche Bedienungshandlungen zur Erreichung verschiedener Systemmodi oder -stufen sind unterscheidbar.“ (Item 14.6)

### 2.3.4 Anforderungen an das System

Es gibt Hinweise darauf, dass Fahrer teilautomatisierter Fahrzeuge ihrer Überwachungsaufgabe nicht immer in geforderter Form nachkommen und sich stattdessen vermehrt mit fahrfremden Tätigkeiten beschäftigen (LLANERAS, SALINGER & GREEN, 2013; GASPAR & CARNEY, 2019). Aus diesem Grund sollen Strategien der Fahrerüberwachung eine unsachgemäße Systemnutzung verhindern. Die EU-Kommission (2008) fordert als Maßnahme beispielsweise in Grundsatz III für das Systemverhalten (S. 24, Kapitel 4.3.5.3): „Die Interaktion mit Systemfunktionen, die nicht zur Nutzung durch den Fahrer während der Fahrt vorgesehen sind, sollte unmöglich gemacht werden, solange sich das Fahrzeug in Bewegung befindet; wenn dies nicht möglich ist, sollte zumindest in unmissverständlicher Weise vor entsprechenden nicht erwünschten Interaktionen gewarnt werden.“ Auch die Alliance of Automobile Manufacturers (2006) stellt diese Anforderung. Um zu verhindern, dass Fahrer längere Zeit ihre Hände vom Lenkrad nehmen und dadurch riskieren, nicht mehr jederzeit in die Fahrzeugsteuerung eingreifen zu können, wird in der UN ECE R79 gefordert, dass das System detektieren können muss, ob der Fahrer die Hände am Lenkrad hat. Falls erkannt wird, dass der Fahrer dies für länger als 15 s nicht tut, soll eine visuelle Hands-Off-Warnung ausgegeben werden, die dem Fahrer die Notwendigkeit vermittelt, die Hände ans Lenkrad zu nehmen. Nach insgesamt 30 s ohne Handerkennung folgt eine visuell-auditive Warnung. Wenn nach weiteren 30 s noch keine Fahrerreaktion erfolgt ist, soll eine Systemdeaktivierung begleitet von einem akustischen Warnsignal erfolgen (UN ECE R79). Die genannten Maßnahmen können jedoch nicht garantieren, dass der Fahrer seine visuelle Überwachungstätigkeit hinreichend gut ausführt. Aus diesem Grund ist zu fordern, dass überwacht wird, ob die visuelle Aufmerksamkeit des Fahrers in ausreichendem Maß auf die Fahraufgabe gerichtet ist. Die folgenden Items dienen der Überprüfung der Maßnahmen zur ausreichenden Systemüberwachung:

- „Das Fahrzeug begrenzt ablenkende Optionen für den Fahrer durch die Einschränkung von Unterhaltungsfunktionen während der Systemnutzung.“ (z. B. Internetnutzung; Item 8.2)
- „Das System fordert den Fahrer auf, die visuelle Aufmerksamkeit wieder auf die Straße zu len-



ken, wenn diese längere Zeit nicht auf der Straße ist.“ (Item 8.4)

- „Es wird ein Hinweis/Warnung ausgegeben, wenn die Hände längere Zeit nicht am Lenkrad sind.“ (Hands-Off-Warnung; Item 8.5)
- „Die Warnung vermittelt die Notwendigkeit, die Hände ans Lenkrad zu nehmen oder das Lenkrad festzuhalten.“ (Item 8.6)
- „Das Eskalationsschema der Hands-Off-Warnung entspricht der Forderung der UN ECE R79: erste optische Warnung spätestens nach 15 s, erste akustische Warnung spätestens nach 30 s, Systemdeaktivierung inkl. akustischer Warnung spätestens nach 60 s.“ (Item 8.7)

## 2.4 Anwendung der Checkliste

### 2.4.1 Anforderungen an die Rater

Ein System soll von zwei Ratern bewertet werden, bei denen es sich um Human-Factors Experten mit mehrjähriger Erfahrung in der Bewertung der MMS von Fahrerassistenz- und Automationssystemen handeln soll. Sie verfügen außerdem über grundlegendes Wissen zu den relevanten Human Factors Konzepten (siehe Kapitel 1.4) und wissen, worum es sich bei einem teilautomatisierten System handelt, von welchen anderen Automationsstufen und -systemen es abgegrenzt werden muss und welche Rolle dem Fahrer bei der Nutzung zukommt. Ist dies nicht der Fall, sollte anhand einer Schulung ein grundlegendes Verständnis über den Bewertungsgegenstand geschaffen werden und vermittelt werden, warum bestimmte Aspekte überprüft werden und worauf bei der Bewertung besonderes Augenmerk gelegt werden muss. In diesem Zusammenhang sollte eine Sensibilisierung für speziell bei diesen Fragestellungen auftretende Probleme (z. B. Mode Confusion) und psychologische Konstrukte (z. B. Situationsbewusstsein, s. ENDSLEY, 1987) stattfinden. Außerdem ist das Verständnis wichtig, dass es sich nicht um eine technische Systembewertung, sondern um die Evaluation von Human Factors Aspekten handelt. Aspekte wie das Systemverhalten (z. B. Güte der Querführung, Häufigkeit von Ausfällen) oder dessen vermeintliche Auswirkungen auf das Fahrerverhalten (z. B. Überwachungsverhalten oder Vertrauen in die Automation) sind explizit nicht Gegenstand der Methode.

Neben der Kenntnis über die durchzuführenden Use Cases machen sich die Rater vor der Testdurchführung mit den Items und den Bewertungsprinzipien vertraut. Die wichtigsten Informationen zum Testablauf und der Benutzung der Checkliste sind auf der ersten Seite der Checkliste zusammengefasst (siehe Anhang 8.1). Darüber hinaus gibt es Schulungsmaterialien zu den oben genannten Aspekten, die für die Schulung der Experten benutzt werden können.

### 2.4.2 Zu prüfende Assistenzfunktion und MMS-Elemente

#### Zu prüfende Assistenzfunktionen

Wird in der Checkliste von „dem System“ gesprochen, ist damit die Kombination aus der kontinuierlichen Längsführungsassistenz (ACC) und der kontinuierlichen Querführungsassistenz (Spurhalteassistenz) gemeint. Einzelne Items erfordern auch die Betrachtung der Einzelsysteme, etwa bei der Aktivierung des Kombinationssystems, wenn ACC bereits aktiv ist. Explizit nicht Bestandteil des teilautomatisierten Systems sind Funktionen wie kurzfristig eingreifende oder warnende Systeme (z. B. Notbremsassistent, Spurverlassenswarnung, Totwinkelassistent). Auch andere Assistenzfunktionen, wie Speed Limiter oder Navigationssysteme sind nicht Gegenstand der Bewertung.

#### Zu prüfende Elemente der Mensch-Maschine-Schnittstelle

Die MMS kann verschiedene Kanäle nutzen, um Informationen zu vermitteln und Rückmeldungen zu geben. Dazu gehören:

- Visuelle Rückmeldungen: Anzeigen von Text, Symbolen, etc. auf einem Display oder durch Beleuchtung.
- Akustische Rückmeldungen: Darbietung von Tönen, akustischen Signalen, Sprachausgaben
- Haptische Rückmeldungen: Rückmeldungen über den taktilen Kanal, z. B. Bewegungsmuster, Vibrationen, pulsierende Signale am Sitz oder Lenkrad

Der Fokus der Checkliste liegt auf der Bewertung von visuellen Anzeigen. Wenn in der Checkliste jedoch allgemein von „Feedback“ gesprochen wird, werden alle Arten dieser Rückmeldungen berück-

sichtigt, außer eine spezifische Rückmeldung ist angesprochen (z. B. nur die visuelle). Von Multimodalität wird dann gesprochen, wenn die Information über mehr als eine Art von Feedback vermittelt wird (z. B. visuell und akustisch). Wenn ein System verschiedene Displays oder Anzeigeorte zur Anzeige von Informationen verwendet, z. B. ein Head-Up-Display, sollen die Anzeigen auf diesem Display ebenfalls mit berücksichtigt werden. Allerdings sollen explizit nur die MMS-Elemente bewertet werden, die Bestandteil des Systems oder dessen Einzelfunktionen sind, d. h. den Systemzustand des teilautomatisierten Systems wiedergeben. Andere Elemente, die zu anderen Assistenzfunktionen oder generell anderen Fahraufgaben gehören, sollen nicht mitbewertet werden (z. B. Darstellung des Tachos etc.). Die Bedienung des Systems kann auf verschiedene Arten erfolgen. Dazu zählen spezielle Tasten, Hebel, Schalter und Ähnliches. Diese können sich auf dem Lenkrad oder an anderen Stellen im Fahrzeug wie beispielsweise der Mittelkonsole befinden. Bei der Bewertung der Systembedienung in der Checkliste sollen nur die Elemente berücksichtigt werden, die für die Bedienung des teilautomatisierten Systems notwendig sind. Andere Elemente sollen nicht bewertet werden (wie z. B. Tasten für andere Assistenzfunktionen etc.).

### 2.4.3 Testumgebung

Die Expertenbewertung kann sowohl im Fahrsimulator als auch in der Realfahrt oder auf Teststrecken erfolgen. Je nach Testumgebung kann es sein, dass manche Use Cases nicht durchzuführen sind oder erlebt werden können. Insbesondere bei der Auslösung einer Hands-Off-Warnung und Erleben einer Systemgrenze handelt es sich um potenziell kritische Fahrsituationen, weswegen die Durchführung auf der Teststrecke oder abgesperrtem Gelände empfohlen wird. Hinzu kommt, dass das Erleben einer Systemgrenze in der Realität schwer zu erzeugen ist. Falls die Bewertung aus ökonomischen Gründen oder dem fehlendem Zugang zu Teststrecken im Realverkehr durchgeführt wird, sollte folgendes beachtet werden: Bei der Strecke sollte es sich um eine Autobahn oder eine Strecke mit baulich getrennten Fahrtrichtungen handeln. Die Strecke sollte eine Fahrtdauer von einer halben Stunde in eine Fahrtrichtung ermöglichen, um ausreichend Zeit für die Absolvierung der Use Cases zu haben. Das Verfahren sollte generell bei eher geringem Verkehrsaufkommen durchgeführt werden (außerhalb der Hauptverkehrszeit). Fahrer sollten vorwie-

gend auf dem rechten Fahrstreifen fahren mit einer Geschwindigkeit, die die Verwendung des L2-Systems erlaubt ohne jedoch den rückwärtigen Verkehr zu behindern. Für die Vergleichbarkeit der Testbedingungen wird eine Geschwindigkeit von 100 – 120 km/h empfohlen. Um Fahrstreifenwechsel sicher durchführen zu können, sollte vor diesen die Setzgeschwindigkeit des Systems erhöht werden. Generell sollten die Use Cases nur dann ausgeführt werden, wenn die Situation es zulässt.

### 2.4.4 Benötigte Ausrüstung

Für die Testdurchführung wird die Checkliste, die aktuell als Paper-Pencil-Variante verfügbar ist, in dreifacher Ausführung (pro Rater eine Checkliste + eine unbeschriebene Version für Abgabe der Urteile auf Item-, Anforderungs- und Systemebene nach Einigung) benötigt inkl. Stifte und Klippboards. Zu Dokumentationszwecken kann im Vorfeld eine Videokamera ins Fahrzeug mitgenommen oder fest installiert werden, um die Systemausgaben bei möglichen Unklarheiten im Nachhinein überprüfen zu können. Die Kameras sollten in der Lage sein, die Anzeigen der MMS und die Systembedienung abzubilden, ohne die Sicht des Fahrers auf die Elemente zu verdecken. Falls die Durchführung im öffentlichen Raum stattfindet, ist aus datenschutzrechtlichen Gründen darauf zu achten, dass sich keine umgebenden Fahrzeuge oder Personen auf den Aufzeichnungen befinden.

### 2.4.5 Testablauf

Im Folgenden werden der Testablauf und das Ausfüllen der Checkliste beschrieben, das von zwei Ratern durchgeführt wird (siehe Kapitel 2.4.1 für Anforderungen an die Rater). Die Gliederung der Checkliste in die Kapitel „Bedienungsanleitung“, „Use Cases“ und „Generelle MMS-Anforderungen“ bildet den zeitlichen Ablauf der Testdurchführung ab. Zuerst soll eine Bewertung der Bedienungsanleitung erfolgen. Die Bewertung kann mithilfe des Benutzerhandbuchs in Papierform oder der systemintegrierten Bedienungsanleitung erfolgen. Welche Art von Bedienungsanleitung verwendet wurde, wird auf dem dafür vorgesehenen Kommentarfeld auf der Checkliste vermerkt. Alle für die Systembedienung relevanten Kapitel des Benutzerhandbuchs müssen gelesen werden. Die Bewertung der Bedienungsanleitung wird erfahrungsgemäß eine Stunde in Anspruch nehmen.

Anschließend sollen die Use Cases während einer Testfahrt erlebt werden. Die Use Cases sind innerhalb der Checkliste so angeordnet, dass sie einer üblichen Systembedienung entsprechen. Die Use Cases dürfen beliebig oft wiederholt werden, bis dem Rater eine ausreichende Beurteilung der Items möglich ist. Während des Erlebens der Use Cases verzichten die Rater auf eine inhaltliche Kommunikation über die zu bewertenden Items, um eine unabhängige und unvoreingenommene Bewertung sicherstellen zu können. Nach Erleben aller Use Cases findet ein Fahrerwechsel statt. Generell gilt, dass jeder Use Case von beiden Testern für die Bewertung erlebt werden sollte. Pro Rater wird ungefähr mit einer Dauer von einer Stunde gerechnet, um das System zu erleben.

Die Bewertung der Use Cases auf Itemebene erfolgt, nachdem beide Fahrer das System ausreichend erlebt haben. Die Bewertung nehmen die Rater getrennt voneinander vor. Im Fall von Unsicherheit bei der Bewertung besteht jederzeit die Möglichkeit, die entsprechenden Use Cases erneut zu testen. Außerdem werden nach Fahrtende die Items bewertet, die sich auf die Anforderungen zur generellen MMS-Bewertung beziehen. Hierbei handelt es sich um eine allgemeine Beurteilung übergreifender Punkte, deren Grundlage eine ausreichende Kenntnis von Systemanzeigen erfordert, die durch das Erleben der Use Cases gewährleistet werden sollte. Auch hier kann bei Bedarf eine Wiederholung der Bedienhandlung erfolgen. Alternativ kann auch auf eine vorhandene Videoaufzeichnung zurückgegriffen werden.

Anschließend erfolgt der Abgleich der bewerteten Items unter den beiden Ratern, mit dem Ziel der Einigung auf Item-Ebene. Bei Nicht-Übereinstimmung werden die Gründe notiert, wobei jeder Rater den Grund seiner Entscheidung vertritt. Kommt es in der Diskussion um die Itembewertung zu Unklarheiten, die das nochmalige Erleben von Use Cases erfordern, ist dies, genauso wie der Rückgriff auf das Foto- oder Videomaterial, jederzeit möglich. Der Prozess der Diskussion wird erfahrungsgemäß einige Zeit in Anspruch nehmen (abhängig davon, ob Use Cases nochmals durchfahren werden müssen oder in vielen Punkten Uneinigkeit besteht). Das Ergebnis der Einigung auf Item-Ebene wird auf einer unbeschriebenen Version der Checkliste festgehalten. Die Bewertung einer Anforderung auf übergeordneter Ebene wird erst im Anschluss an die Diskussion vorgenommen, wenn die Rater in allen Untertiteln eine Einigung erzielen konnten. Diese Be-

wertung wird entsprechend den Bewertungskriterien aus Kapitel 2.2. durchgeführt.

Im Anschluss erfolgt die gemeinsame Abgabe der finalen Systembewertung. Insgesamt ist mit einer Testdauer von 4 – 6 Stunden pro Fahrzeug zu rechnen.

## 2.5 Erprobung und Anpassung der Checkliste

Während ihrer Entwicklungsphase durchlief die Checkliste mehrere Erprobungsphasen. Ziel dieser Iterationen war die Überprüfung der Checkliste auf ihre Anwendbarkeit sowie auf Reliabilität (d. h. kommen unterschiedliche Rater auf ähnliche Ergebnisse) und Validität (d. h. ist die Checkliste in der Lage, Unterschiede von unterschiedlich guten Systemen aufzudecken). Insgesamt wurde die Checkliste zweimal im Realverkehr anhand eines Serienfahrzeugs mit L2-Funktion sowie einmal in Bezug auf zwei verschiedene MMS-Varianten, die im Fahrsimulator implementiert wurden, von Experten getestet. Die Ergebnisse der Expertenbewertung im Fahrsimulator wurden mit den Ergebnissen einer Nutzerstudie dahingehend verglichen, ob sich Gestaltungsunterschiede in der MMS auch im Verhalten der Probanden niederschlagen. Diese werden im zweiten Teil des Berichts (Kapitel 3.5.1) dargestellt.

### 2.5.1 Erste Erprobung im Realverkehr

Zur ersten Erprobung der Checkliste (Juli 2018) wurde ein Tesla Model S mit Autopilot-Funktion verwendet. Im Fokus der Erprobung stand die Verständlichkeit der Items und die Anwendbarkeit des Verfahrens. Insgesamt zwei Raterpaare, bestehend aus je zwei Human Factors-Experten der WIVW GmbH, durchführten dabei die vorgegebenen Use Cases entsprechend des Testprotokolls, welches ihnen im Vorfeld ausführlich durch die Entwickler der Checkliste erläutert wurde. Als Prüfstrecke wurde ein Autobahnabschnitt auf der A 70 zwischen Schweinfurt und Bamberg genutzt. Die Strecke ist zweistreifig pro Fahrtrichtung und ohne generelle Geschwindigkeitsbegrenzung. Sie enthält einen Tunnel, der sich zur Prüfung des Displaykontrasts eignet (Geschwindigkeitsbegrenzung: 80 km/h). Im Anschluss füllten die Experten die Checkliste aus. In der anschließenden gemeinsamen Diskussion konnten Unklarheiten bei der Itemformulierung, die

zu Verständnisproblemen und dadurch bedingten Abweichungen in der Bewertung führten, identifiziert und behoben werden. Die Checkliste wurde daraufhin angepasst.

### 2.5.2 Erprobung im Rahmen einer Simulatorstudie

In einer weiteren Erprobungsphase (März 2019) wurde die optimierte Version der Checkliste im Rahmen einer Expertenbewertung erneut getestet. Diese Erprobung hatte das Ziel zu überprüfen, ob die Checkliste in der Lage ist, unterschiedlich gute Varianten der MMS identifizieren zu können. Dazu wurden zwei verschiedene MMS-Varianten in den Fahr-Simulator implementiert, welche die Anforderungen der Checkliste in unterschiedlichem Ausmaß erfüllten. Zudem sollte geprüft werden, inwieweit die Checkliste sowie die zugehörigen bereitgestellten Schulungsmaterialien (im Unterschied zur ersten Erprobung wurden Erläuterung und Definition der zu bewertenden Funktionen und MMS-Anzeigen sowie Erläuterungen zum Bewertungsprozess und zum Ratingprinzip entworfen) verständlich sind und der Bewertungsprozess praktikabel angewendet werden kann.

#### Teilautomatisierte Fahrfunktion

Basierend auf Benchmark-Analysen wurde eine Automatisierungsfunktion entwickelt, die sich dadurch auszeichnet, dass die kontinuierlich assistierende Quer- und Längsführung separat voneinander genutzt werden können. Bevor die Längsführung aktiviert werden kann, muss sie zunächst in einen Stand-by-Modus geschaltet werden. Die Querführung wird dagegen automatisch aktiviert, sobald das Fahrzeug in einem definierten Korridor innerhalb des Fahrstreifens fährt und das System Fahrstreifen erkennt, nachdem sie einmalig vom Fahrer in den Stand-by-Modus geschaltet wurde. Die Längsführung kann durch Betätigung des Fahrpedals kurzzeitig übersteuert werden. Wird die Bremse betätigt, fällt die Längsführung in den Stand-by-Modus zurück und muss über die entsprechende Lenkradtaste wieder aktiviert werden. Durch einen Lenkeingriff über einer definierten Schwelle kann die Querführung übersteuert werden. Sobald sich das System wieder innerhalb des Fahrstreifens befindet, reaktiviert sich die Querführung. Durch Betätigung der jeweiligen Tasten für Quer- und Längsführung am Lenkrad können die Systeme vollstän-

dig deaktiviert werden. Das System passt seine Geschwindigkeit automatisch an vorgegebene Geschwindigkeitslimits an. Darüber hinaus kann die gefahrene Geschwindigkeit manuell in 10-km/h-Schritten angepasst werden. Der Abstand zu einem vorausfahrenden Fahrzeug kann prinzipiell verstellt werden, war aber in der vorliegenden Studie standardmäßig auf der mittleren Stufe eingestellt.

Weiterhin verfügt das System über eine Hands-Off-Warnung in zwei Stufen, die den Fahrer nach Wegnahmen der Hände vom Lenkrad für 15 Sekunden in einer ersten Stufe durch ein gelbes Lenkrad-Symbol warnt. In einer zweiten Stufe wird nach weiteren 15 Sekunden durch ein rotes Lenkrad und einen zusätzlichen akustischen Ton gewarnt. Nach weiteren 30 Sekunden schaltet sich die Querführung automatisch ab.

#### Varianten der Mensch-Maschine-Schnittstelle

Zur Validierung wurden zwei unterschiedlich „gute“, d. h. im Hinblick auf die in der Checkliste definierten Anforderungen in unterschiedlichem Maße erfüllende Versionen der MMS erstellt. Die MMS-Versionen sollten sich bezüglich möglichst vieler Items unterscheiden, dabei aber in einem möglichst realistischen Rahmen bleiben. Beispielsweise ist die Bedienlogik in der „schlechten“ MMS-Variante insofern komplexer, als dass das ACC-System vor einer Aktivierung zunächst in den Stand-by-Zustand gebracht werden muss und die dazu benötigten Tasten zur Bedienung an unterschiedlichen Positionen am Lenkrad angebracht sind. Zudem korrespondiert deren Beschriftung nicht mit den in der visuellen Anzeige dargestellten Symbolen. Weiterhin aktiviert sich das System erst nach einer leichten Verzögerung ohne explizite Rückmeldung dazu. Außerdem gibt es Unterschiede bei der letzten Stufe der Hands-Off-Warnung dahingehend, dass in einer Variante mit der Abschaltung der Querführung zusätzlich eine ACC-Verzögerung einsetzte.

Weitere Unterschiede ergeben sich in den visuellen Anzeigen, die die Unterscheidbarkeit der einzelnen Systemzustände erschweren können (siehe Bild 2-2). So wird in der „schlechten“ MMS-Variante die aktive Spurhalteassistentz nur durch die farbige Darstellung von vertikalen Linien angezeigt, während in der „guten“ MMS zusätzlich ein blaues Lenkrad mit Händen angezeigt wird. Insgesamt ist die Lesbarkeit der Anzeigen durch einen schlechteren Farbkontrast zwischen Vorder- und Hintergrund in

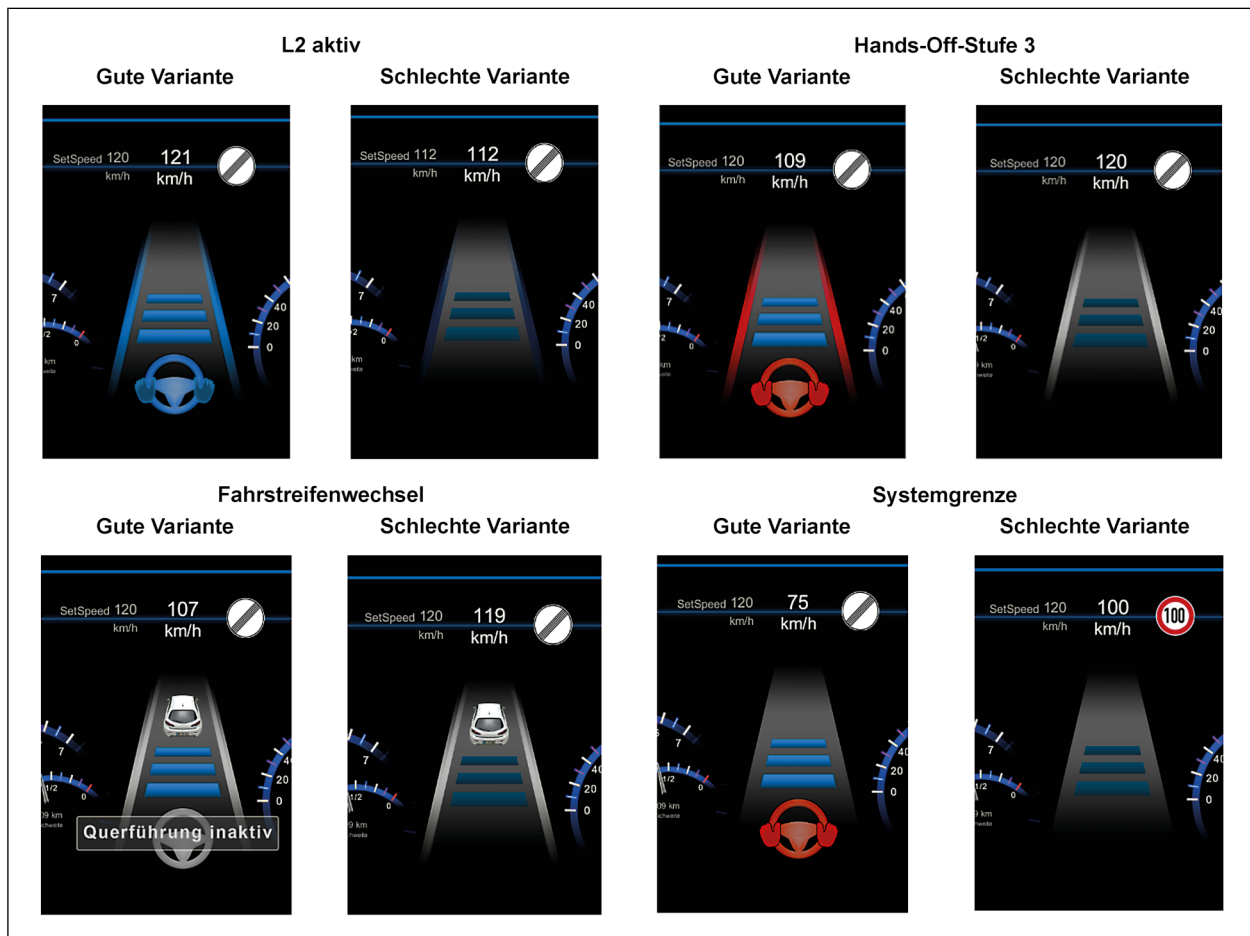


Bild 2-2: Vergleich der visuellen MMS-Varianten in verschiedenen Systemzuständen (Quelle: WIVW GmbH)

der „schlechten“ MMS erschwert. Unterschiede ergeben sich außerdem durch eine zusätzliche visuelle und akustische Warnung im Fall einer vorhersehbaren Systemgrenze (scharfe Kurve) in der guten Variante. Während eines fahrerinitiierten Fahrstreifenwechsels wird in der „guten“ MMS-Variante eine zusätzliche Textanzeige über den Status des Systems gelegt (Querführung inaktiv). Tabelle 2-2 im Ergebnisteil stellt die theoretisch erzeugten Unterschiede nochmals auf der Itemebene der Checkliste dar.

Im Folgenden sind die Unterschiede überblicksartig zusammengefasst:

- Unterschiedliche Bedienlogik
  - Aktivierung des ACC: direkt vs. Zwischenschritt über „Stand-by-Zustand“,
  - Positionierung der Tasten: am Lenkrad nah beieinander vs. weit entfernt,
  - Beschriftung der Lenkradtasten: korrespondierend mit Symbolen der visuellen Anzeigen vs. nicht korrespondierend,
- Aktivierung des Systems nach Tastendruck: unmittelbar vs. mit zeitlicher Verzögerung von 1 s.
- Unterschiede in den visuellen Anzeigen
  - Anzeige der aktiven Querführung: über Lenkrad + Fahrstreifenmarkierungen vs. nur Fahrstreifenmarkierungen,
  - Lesbarkeit: guter vs. schlechter Kontrast zwischen Symbolen und Hintergrund,
  - Fahrerinitiiertes Fahrstreifenwechsel: zusätzliche Textanzeige über den Status des Systems während des Fahrstreifenwechsels (Querführung inaktiv) vs. keine Textanzeige.
- Unterschiede im Warnkonzept
  - Warnung: visuelle + akustische Warnung vor einer vom System erkannten Systemgrenze (scharfe Kurve) vs. keine Warnung,
  - Hands-Off Strategie: Implementierung einer Verzögerung des Fahrzeugs in der letzten Stufe der Hands-Off-Warnung vs. keine Verzögerung.

## Prüfstrecke

Zur Bewertung der Varianten wurde eine 20-minütige Autobahnstrecke im Simulator durchfahren, welche die verschiedenen Use Cases enthielt, die als relevant definiert wurden und die Prüfung der Checklistenitems ermöglichen. Diese Use Cases sind

- Systemaktivierung (instruiert; 2 mal),
- Fahren mit aktivem L2-System,
- Fahrstreifenwechsel bei langsamem Vorderfahrzeug (instruiert),
- Systemausfall: Hindernis,
- Ausfall der Querführung bei schlechten Fahrstreifenmarkierungen,
- Systemgrenze: Scharfe Kurve mit Ausfall der Querführung,
- Systemdeaktivierung (instruiert).

## Methodik

Die Expertenbewertung fand im dynamischen Fahr-Simulator der WIVW GmbH statt und dauerte insgesamt 3.5 h. Bei den Experten handelte es sich um zwei wissenschaftliche Mitarbeiter des WIVW, die Erfahrung in der Bewertung von MMS hatten. Bewertet werden sollten die zwei entwickelten MMS-Varianten:

- Die „gute“-Variante A, die viele der Checklisten-Items per Definition erfüllen sollte,
- Die „schlechte“-Variante B, die wenige der Checklistenitems erfüllen sollte

Als Bewertungsgrundlage diente die WIVW-Checkliste in der Version vom Februar 2019, welche bereits die Anpassungen der ersten Erprobungsrunde enthielt. Als Prüfstrecke wurde die in Kapitel 3.5.1 beschriebene Autobahnstrecke verwendet. Zusätzlich zu den dort beschriebenen Use Cases sollten die Experten während der Fahrt mit aktivem System einmal den vollständigen Ablauf einer Hands-Off-Warnung erleben. Als Vorbereitung für die Bewertung erhielten die Rater folgende Materialien:

- Informationen über die Aufgaben der Rater
- Erläuterungen und Definitionen der zu bewertenden Systeme und MMS-Elemente (welche Elemente der MMS sind Gegenstand der Bewertung, was soll nicht bewertet werden)

- Erläuterungen zum Aufbau der Checkliste und des Ratingprinzips
- Checkliste in der Version vom Februar 2019 (enthält 18 Anforderungen; Achtung: die Reihenfolge und Inhalte der Items weichen leicht von der finalen Version nach dem letzten Erprobungsschritt ab, siehe Kapitel 2.5.3)

Der Ablauf der Bewertung erfolgte in folgenden Schritten: Zunächst sollte die als Ersatz einer Bedienungsanleitung erstellte Kurzanleitung für die im Anschluss erlebte Systemvariante (Variante A oder B) durchgelesen werden. Die Kurzanleitungen der beiden Varianten unterschieden sich lediglich hinsichtlich der Beschreibung der Systembedienung, die sich zwischen den Varianten unterschied. Danach sollte die Erfüllung der Unteritems (nach ja-/nein-Prinzip) zu Anforderungen für das „Quick Manual“ aus der Kategorie „User Manual“ der Checkliste beurteilt werden. Die ersten vier Items zur Bedienungsanleitung, die sich auf ein ausführliches Benutzerhandbuch beziehen, wurden nicht bewertet, sodass von 18 möglichen, d. h. in der Checkliste enthaltenen Anforderungen, 14 beurteilt wurden.

Anschließend wurde das System während der Simulatorfahrt in den verschiedenen Use Cases erlebt. Beide Rater fuhren dafür abwechselnd die Prüfstrecke mit den beiden Systemvarianten in umgekehrter Reihenfolge. Nach den spezifischen Use Cases (Aktivierung, Deaktivierung, Fahrstreifenwechsel, L2-Fahrt inkl. Hands-Off-Warnung und kurzzeitiger Ausfall der Querführung, Systemgrenze) wurde die Simulation jeweils pausiert, sodass die Experten die Bewertung der jeweiligen Unteritems der entsprechenden Checklisten-Anforderungen unmittelbar vornehmen konnten. Nachdem die Rater jeweils die Prüfstrecke mit einer Systemvariante durchfahren hatten, wurden die Anforderungen der generischen Kategorien („General HMI-requirements“ und „Control elements“) bewertet. Daraufhin sollte pro Anforderung die dreistufige Bewertung in „fulfilled“, „satisfactory“ und „not-acceptable“ erfolgen<sup>2</sup>. In einer abschließenden Diskussionsrunde nach Erleben beider MMS-Varianten wurde ein gemeinsames Urteil auf Item- und Anforderungsebene erzielt bevor ein einstimmiges Gesamturteil ge-

<sup>2</sup> Dieser Bewertungsschritt wird in der finalen Methode (siehe Kapitel 5.2.2) erst im Anschluss an die gemeinsame Diskussion und Einigung auf Item-Ebene durchgeführt.

ben wurde. Dieses Gesamturteil wurde in drei Stufen anhand der Kategorien „very good – no improvements necessary“, „acceptable with options for improvement“, „not acceptable“ abgegeben.

**Ergebnisse**

Die Ergebnisse des Expertenratings zeigen, dass die beiden MMS-Varianten unterschiedlich gut auf den Items der Checkliste bewertet werden. Tabelle 2-2 stellt die theoretisch erzeugten Unterschiede auf der Itemebene sowie die Bewertung durch die Rater dar (Grün = Item erfüllt, Gelb = Item akzeptabel, aber verbesserungsfähig, Rot = nicht akzeptabel).

bel, aber verbesserungsfähig, Rot = nicht akzeptabel).

In der gemeinsamen Bewertung der 14 Anforderungen (vier Anforderungen zur Bedienungsanleitung wurden nicht bewertet) ergab sich für Variante A das Urteil „fulfilled“ in 79 % der Items und das Urteil „satisfactory“ in 21 % der Items (siehe Tabelle 2-3). Für keine der Anforderungen wurde das Urteil „not-acceptable“ vergeben. Im Gegensatz dazu wurde in Variante B keine Anforderung als „fulfilled“ bewertet. 57 % der Anforderungen wurden als „satisfactory“ und 43 % als „not-acceptable“ bewertet (siehe Ta-

Nr.	Item	„gute“ MMS (Var A)	„schlechte“ MMS (Var B)	Empirische Unterschiede in Expertenbewertung				
5	A quick way to retrieve the most important information is provided	Kurzbedienungsanleitung identisch für Var A und B, mit Ausnahme der Beschreibung der Systembedienung		Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Rot</td></tr> </table>	Var A	Var B	Grün	Rot
Var A	Var B							
Grün	Rot							
6	The activation of the system is easy <ul style="list-style-type: none"> <li>Activation of the system out of manual driving is easy to perform</li> <li>Activation of the system out of lower automation levels (e.g. ACC active) is easy to perform</li> </ul>	ACC und Querführung jeweils direkt über eine Taste aktivierbar	ACC muss zunächst in Stand-by gebracht, kann dann erst aktiviert werden (2 Stufen)	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Gelb</td><td>Rot</td></tr> </table>	Var A	Var B	Gelb	Rot
Var A	Var B							
Gelb	Rot							
7	Feedback about the activation process is given <ul style="list-style-type: none"> <li>Clearly indicated/simultaneously to the driver input</li> <li>In case of a delay, process itself is indicated</li> </ul>	Keine Verzögerung zwischen Tastendruck und Aktivierung	Verzögerte Aktivierung (1s) ohne besondere Rückmeldung	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Gelb</td></tr> </table>	Var A	Var B	Grün	Gelb
Var A	Var B							
Grün	Gelb							
8	The deactivation of the system is easy <ul style="list-style-type: none"> <li>Deactivation of the full system is easy to perform</li> <li>Deactivation of single parts of the system is easy to perform</li> </ul>	ACC und Querführung jeweils direkt über eine Taste deaktivierbar	ACC kann zunächst in Stand-by gebracht, dann deaktiviert oder über separate Taste direkt komplett deaktiviert werden	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Gelb</td></tr> </table>	Var A	Var B	Grün	Gelb
Var A	Var B							
Grün	Gelb							
9	Status information is permanently displayed in L2 <ul style="list-style-type: none"> <li>Active system status is displayed</li> <li>Information about driver's responsibility can be derived</li> </ul>	Anzeige des aktiven Systems permanent über ACC-Balken, Fahrstreifenmarkierung + Lenkrad mit Händen dran zeigt Verantwortlichkeit	Anzeige des aktiven Systems über ACC-Balken, Fahrstreifenmarkierung; kein zusätzliches Lenkrad	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Gelb</td></tr> </table>	Var A	Var B	Grün	Gelb
Var A	Var B							
Grün	Gelb							
10	Information about the function working properly can be derived <ul style="list-style-type: none"> <li>...Degraded system functionality is clearly indicated (e.g. short-term deactivation of lateral control)</li> </ul>	Anzeige der Sensorgüte über graue/blauere Fahrstreifenmarkierung + Lenkrad	Anzeige der Sensorgüte über graue/blauere Fahrstreifenmarkierung (kein Lenkrad – daher weniger salient)	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Gelb</td></tr> </table>	Var A	Var B	Grün	Gelb
Var A	Var B							
Grün	Gelb							
11	The system provides a strategy to deal with inappropriate system use (driving hands-off for too long; driving without monitoring the system and the road)	Hands-Off-Warnung nach UNR79-Fahrzeug verzögert nach Ablauf der 2. Warnstufe + Querführung geht nach Warnung in Stand-by	Hands-Off-Warnung nach UNR79 – Querführung geht in Stand-by	Ja <table border="1"> <tr><td>Var A</td><td>Var B</td></tr> <tr><td>Grün</td><td>Rot</td></tr> </table>	Var A	Var B	Grün	Rot
Var A	Var B							
Grün	Rot							

Tab. 2-2: Unterschiedliche Gestaltung der MMS-Varianten basierend auf den Checklisten-Items in der für die Studie verwendeten Version und die tatsächliche Bewertung durch 2 Rater in der vorangegangenen Expertenbewertung Grün: fulfilled, Gelb: satisfactory, Rot: not-acceptable)

Nr.	Item	„gute“ MMS (Var A)	„schlechte“ MMS (Var B)	Empirische Unterschiede in Expertenbewertung				
12	The way of using the Lane Change is easy to execute • ...The assistance by the system is clearly perceivable during the lane change process itself	Zusätzliche Textanzeige „Spurwechsel wird ausgeführt“	Keine Rückmeldung durch System	Nein <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Yellow</td><td>Yellow</td></tr></table>	Var A	Var B	Yellow	Yellow
Var A	Var B							
Yellow	Yellow							
13	Predictable limitations of the functionality are clearly communicated • A warning/take-over request is issued	Bei Systemgrenzen visuelle und akustische Warnung	Keine Warnung an Systemgrenzen	Ja <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Green</td><td>Red</td></tr></table>	Var A	Var B	Green	Red
Var A	Var B							
Green	Red							
14	The system mode is indicated at all times • ...All potential system modes have at least one indicator (e.g. available, ACC active, L2 active, degraded, deactivated)	Aktive Querführung durch Fahrstreifenmarkierung und permanente Lenkrad-Anzeige dargestellt	Aktive Querführung nur durch Fahrstreifenmarkierung dargestellt	Ja <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Green</td><td>Yellow</td></tr></table>	Var A	Var B	Green	Yellow
Var A	Var B							
Green	Yellow							
15	Mode changes are displayed	Schnelle und klare Rückmeldung durch Hinzukommen/Wegfall des Lenkrads bei Aktivierung/Deaktivierung QF	Verzögerte Aktivierung, Deaktivierung weniger auffällig durch fehlendes Lenkrad	Ja <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Green</td><td>Red</td></tr></table>	Var A	Var B	Green	Red
Var A	Var B							
Green	Red							
16	Visual information is easy to perceive and easy to read • Color of indicators contrasts adequately from background color	guter Kontrast zwischen Icon-Farben (v. a. blauton) und Hintergrund (schwarz)	Schlechter Kontrast zwischen Vorder- und Hintergrund durch sehr dunklen Blauton	Nein <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Yellow</td><td>Yellow</td></tr></table>	Var A	Var B	Yellow	Yellow
Var A	Var B							
Yellow	Yellow							
17	The presented information is easy to process and understand	Gleiche Symbolik in beiden MMS-Varianten verwendet – keine expliziten Variationen		Ja <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Green</td><td>Yellow</td></tr></table>	Var A	Var B	Green	Yellow
Var A	Var B							
Green	Yellow							
18	System promotes an easy and correct use of relevant control elements • Controls (e.g. buttons) are grouped together • Functions and labels are easy to understand	Separate Tasten für Längs- und Querführung liegen auf der linken Seite direkt übereinander; Tastenbeschriftung entspricht den dargestellten Icons für die Systeme	Querführungs-Taste auf rechter Seite, On-Off Taste auf rechter Seite, Aktivierung der Längsführung auf linker Seite Tastensbeschriftung durch andere Symbolik als in Anzeige	Ja <table border="1"><tr><td>Var A</td><td>Var B</td></tr><tr><td>Green</td><td>Red</td></tr></table>	Var A	Var B	Green	Red
Var A	Var B							
Green	Red							

Tab. 2-2: Fortsetzung

Var	Fulfilled	Satisfactory	Not-acceptable
A	11	3	0
B	0	8	6

Tab. 2-3: Gemeinsames Urteil aller Anforderungen auf 3-stufiger Bewertungsskala

belle 2-3). Dies führte zu einer Gesamtbewertung der Variante A, die von den Ratern aufgrund eines 3-stufigen Urteils nach ihrem eigenen Ermessen vorgenommen wurde, als „acceptable with options for improvement“ und der Variante B als „not acceptable“. Dieses Ergebnis impliziert, dass die beiden MMS-Varianten durch die Checkliste gut voneinander

getrennt werden können und in der bestehenden Form (mit leichten Modifikationen) zur Validierung der im Anschluss durchgeführten Nutzerstudie (siehe Kapitel 3.3.2) genutzt werden können. Durch Rückfragen ergab sich jedoch, dass die Rater das Urteil „very good – no improvements necessary“ nur für ein absolut perfektes System vergeben würden, d. h. 100 % „fulfilled“ Bewertungen. Da dies ein eher unwahrscheinliches Ergebnis darstellt, würden sich die Rater eine 4-stufige Abstufung durch eine weitere Unterteilung der aktuell mittleren Kategorie „acceptable with options for improvement“ wünschen, um ihre Bewertung besser differenzieren zu können, was in der finalen Version der Checkliste berücksichtigt wurde.



Die Berechnung der Interrater-Reliabilität zum Zeitpunkt vor der Diskussion ergab, dass gerade auf der Ebene der Unteritems (ja-/nein-Bewertung) eine hohe Übereinstimmung der beiden Rater bestand. Für Variante A betrug die Übereinstimmung nach Brennan  $k = 0.898$ , für Variante B  $k = 0.795$ . Für die dreistufige Bewertung auf Ebene der jeweiligen Anforderung lag die Übereinstimmung für Variante A ebenfalls in einem sehr guten Bereich von Brennan  $k = 0.679$ . Für die Bewertung der Variante B war die Übereinstimmung mit Brennan  $k = 0.250$  weniger zufriedenstellend. Mögliche Gründe bestehen darin, dass Rater A bei der Itembewertung generell kritischer vorging und häufiger „not-acceptable“ Ratings auf der Anforderungsebene vergab. Zudem folgte Rater A nicht immer der Regel zur Vergabe des 3-stufigen Urteils, sondern vergab das „not-acceptable“-Urteil auch in Fällen, in denen nur ein Unteritem mit „nein“ bewertet wurde, wenn er die Verletzung dieses Items als besonders gravierend einstuft (z. B. problematische Deaktivierung, mangelnder Kontrast). Im Gegensatz dazu wurden von Rater B einige Aspekte wie die Verzögerung im Aktivierungsprozess nicht bemerkt und demnach nicht in den Bewertungsprozess einbezogen. Nach der gemeinsamen Diskussion konnten diese Unstimmigkeiten zur Zufriedenheit beider Rater ausgeräumt werden. Die Diskussionsdauer für Variante A betrug 20 Minuten, für Variante B 30 Minuten. Die Ergebnisse zeigen, dass ein Vier-Augen-Prinzip bei der Bewertung sinnvoll ist und mit einem akzeptablen Aufwand umgesetzt werden kann. Dass eine Diskussion im Rahmen einer heuristischen Evaluation durch ein Checklistenverfahren unerlässlich ist, zeigten auch WIEDEMANN, SCHÖMIG, NAUJOKS, HERGETH & KEINATH (2018).

Die zur Verfügung gestellten Materialien wurden von den Ratern als umfassend und verständlich empfunden. Die in der Checkliste formulierten Items erwiesen sich größtenteils als eindeutig und verständlich formuliert. An anderen Stellen wurde die Formulierung der Items für ein klareres Verständnis präzisiert. Die Regel für die Bewertung der übergeordneten Anforderung nach der Häufigkeit von „nein“-Urteilen in den Unteritems erwies sich als sinnvoll, um eine einheitliche Bewertung zu erreichen. An einigen Stellen konnte Verbesserungsbedarf aufgedeckt werden: Zum Beispiel sollte auf dem Bogen vermerkt werden können, dass ein Use Case nicht erlebt wurde, weil er nicht explizit ausgelöst oder instruiert werden können (z. B. Systemrückmeldung im Fall eines abgebrochenen Fahr-

streifenwechsels). Wenn Fälle nicht erlebt werden können, weil sie nicht im Systemumfang enthalten sind, wie z. B. ein Spurwechselassistent, wird dies vor Beginn der Testung auf dem Bewertungsbogen vermerkt. Für eine bessere Trennschärfe von Items erfolgte eine Reduzierung der Anforderungen von 18 auf nun insgesamt 14. Weitere Rückmeldungen der Rater wurden genutzt, um die Checkliste hinsichtlich Anwendbarkeit, Verständlichkeit und Vollständigkeit zu verbessern, bevor sie einer finalen Überprüfung unterzogen wurde.

### 2.5.3 Finale Erprobung im Realverkehr

Die finale Überprüfung der überarbeiteten Checkliste fand im Realverkehr mithilfe eines Tesla Model 3 mit Autopilotfunktion (Beta-Version) statt (Mai 2019). Der Autopilot läuft im Gegensatz zur sog. Spurhalteassistent-Funktion nur in Verbindung mit einer Zieleingabe im Navigationssystem und macht zusätzlich zur kontinuierlich assistierenden Quer- und Längsführung auch Fahrstreifenwechsel-Vorschläge. Diese werden – nachdem der Fahrer sie über den Blinker bestätigt hat – systemseitig ausgeführt. Die Zustandsanzeige des Autopiloten wird ausschließlich auf dem in der Mittelkonsole angebrachten Touch-Display angezeigt (siehe Bild 2-3). Die aktive L2-Funktion wird dabei über eine pfadähnliche Anzeige einer Trajektorie visualisiert. Nach jeder Aktivierung des Systems wird der Fahrer über eine Textmeldung aufgefordert, die Hände am Lenkrad zu lassen (Hände bitte am Lenkrad behalten). Detektiert das System während des Fahrens mit dem System zu wenige Fahrereingaben am Lenkrad, fordert es den Fahrer mit der Textmeldung „Lenkrad leicht drehen“ und einem eingeblendeten Lenkradsymbol mit Händen daran dazu auf. Die



Bild 2-3: Mensch-Maschine-Schnittstelle des Tesla Model 3

Querführung kann jederzeit über ein Übersteuern des Lenkrads deaktiviert werden. Daraufhin ist nur noch das ACC aktiv, was über eine blaue Umrandung des ACC-Icons dargestellt ist. Über ein erneutes zweimaliges Herunterdrücken des Lenkradhebels kann der Autopilot reaktiviert werden. Ein zu langes Wegnehmen der Hände wird zunächst über ein blau pulsierendes Leuchten der Systemanzeige, dem Lenkrad-Symbol und der Textanzeige „Lenkrad leicht drehen“ dargestellt. Mit zunehmender Dauer ohne Hände am Lenkrad erhöht sich die Frequenz des Pulsierens, die Hände am Lenkrad werden dann rot dargestellt und ein Warnton kommt hinzu. In der letzten Stufe der Hands-Off-Warnung wird der Fahrer über eine dringliche Warnung aufgefordert, die Hände wieder ans Lenkrad zu nehmen. Reagiert er auch darauf nicht, wird das System abgeschaltet und kann für den Rest der Fahrt nicht mehr genutzt werden, es sei denn, ein Neustart erfolgt.

Als Prüfstrecke diente derselbe Autobahnabschnitt wie bei der ersten Erprobung (siehe Kapitel 2.5.1). Im Unterschied dazu fanden während dieses Tests Bauarbeiten auf der Strecke statt, durch die der rechte Fahrstreifen durch Warnbaken abgesperrt war und der Verkehr auf den linken Fahrstreifen geleitet wurde (Beschränkung auf 80 km/h). Während des Tests zeigte das getestete System praktisch keine Ausfälle in der Querführung, bei der Einfahrt in die Baustelle musste der Fahrer allerdings eingreifen, da das System dort das Fahrzeug ohne Eingriff in die Warnbaken gesteuert hätte. Der Tunnel wurde vom System als bevorstehende Systemgrenze angekündigt, mit der Text-Meldung, dass in 200 m der Autopilot aufgrund des Tunnels nicht nutzbar sein würde.

Im Zuge dieser Erprobung wurden auch gezielt die Anforderungen zur Bedienungsanleitung überprüft, die im Rahmen der Simulatorstudie nicht berücksichtigt werden konnten. Durch die im Anschluss an die Simulatorerprobung erfolgten Anpassungen an Wortlaut und Inhalt der Items ließ sich die Übereinstimmung vor der Diskussion erhöhen sowie die Diskussionszeit der Rater verkürzen. Auch in dieser Erprobung machte die kritische Auseinandersetzung zweier Human-Factors-Experten des WIVW deutlich, an welchen Stellen Items zur Vermeidung von Unklarheiten noch deutlicher formuliert werden mussten. Außerdem wurde deutlich, dass die Rater manche Kritikpunkte am System noch nicht in der Checkliste unterbringen konnten. Die Items wurden nochmals angepasst oder ergänzt. Vereinzelt er-

folgte noch eine Zusammenführung von Items. Das Ergebnis dieser finalen Erprobung ist die finale Checkliste, wie sie in Kapitel 2.1 beschrieben und in Kapitel 8.2 im Anhang eingefügt ist.

## 2.6 Gesamtbewertung und Limitationen des Verfahrens

Gegenstand der entwickelten Methode ist die Bewertung der visuell-auditiven Komponenten der Mensch-Maschine-Schnittstelle teilautomatisierter Fahrzeugsysteme, der zur Systeminteraktion benötigten Bedienelemente sowie Systeminformationen, die dem Fahrer im Vorfeld durch das Benutzerhandbuch vermittelt werden. Das Checklisten-Verfahren erlaubt eine Gesamteinschätzung über den Erfüllungsgrad von Anforderungen an die MMS sowie die Bedienungsanleitung, die auf allgemein akzeptierten Normen und Gestaltungsrichtlinien basiert und von Experten getroffen wird. Auf diese Weise kann überprüft werden, ob die Voraussetzungen für ein korrektes Verständnis und einen sicheren Systemgebrauch von L2-Systemen anhand der Betriebsanleitung und der Nutzerschnittstelle vorliegen. Das Verfahren wurde in einem iterativen Prozess entwickelt und optimiert. Anerkannte Gütekriterien von subjektiven Einschätzungsverfahren wurden ebenfalls geprüft und können als erfüllt angesehen werden.

Die Methode hat nicht zum Ziel, konkrete Designvorgaben zu machen (z. B. welche Farben oder Icons verwendet werden sollen). Die Anforderungen aus der Checkliste sind vielmehr als generische Richtlinien zu verstehen, die von Herstellern individuell umgesetzt werden können. Eine Bewertung hinsichtlich der Gebrauchssicherheit von Systemen im Falle von Systemgrenzen oder -fehlern ist ebenfalls nicht Gegenstand der Methode. Weiterhin macht das Verfahren explizit keine Angaben zur Gestaltung des Regelungsverhaltens. Es werden keine Vorgaben oder Empfehlungen zum Systemverhalten in konkreten Situationen gemacht. Grund hierfür ist, dass auf Basis des vorliegenden Wissensstands keine abschließenden Aussagen allein aufgrund von Expertenbeurteilungen getroffen werden können. Dies bedeutet auch, dass Systeme, die sich zwar im Systemverhalten (z. B. bezüglich der Häufigkeit von Systemausfällen), nicht aber in der MMS unterscheiden, mittels der vorliegenden Checkliste vergleichbar bewertet werden würden. Zudem muss herausgestellt werden, dass durch ein

Checklistenverfahren nur beschränkt Aussagen über den intuitiven Erstkontakt durch naive Nutzer gemacht werden können. Einerseits handelt es sich bei den durchführenden Ratern um Experten im Bereich der MMI, andererseits ist das Lesen der Bedienungsanleitung und der dazugehörigen Kapitel zur Systeminteraktion Teil des Testprotokolls. Die Methode ermöglicht die Identifikation von Unzulänglichkeiten in der Gestaltung der MMS, die ein korrektes Verständnis und eine korrekte Interaktion mit L2-Systemen behindern können. Wie sich derartige Unzulänglichkeiten letztlich im Verhalten von Fahrern niederschlagen und wie gravierend die Auswirkungen sind, ist jedoch ein weitaus komplexer Prozess, der nur bedingt durch Expertenverfahren präzisiert werden kann. Für diese Aspekte stellen Nutzerstudien durch eine direkte, intuitive Interaktion von naiven Nutzern mit dem System eine sinnvolle Ergänzung des Verfahrens dar. Weiterhin ist die Nutzerstudie in der Lage, Effekte aufzuzeigen, die die Checkliste nicht vorhersagen kann. Darunter fallen beispielsweise Übungs- und Lerneffekte in der Systembedienung und im Umgang mit Systemausgaben (z. B. nehmen Fahrer die Hände nicht mehr vom Lenkrad, nachdem sie eine Hands-Off-Warnung erhalten haben) und Reihenfolgeeffekte (z. B. vorheriges Erleben von Systemgrenzen wirkt sich auf nachfolgende Reaktionen aus). Um die gesamte Bandbreite des Verhaltens abzudecken, wird bewusst eine größere Anzahl von Fahrern getestet.

Im Unterschied zu Nutzerstudien ergeben sich Vorteile des Experten-Bewertungsverfahrens in folgenden Punkten: Die konkrete Aufgabe der Experten besteht darin, die MMS zu bewerten, weswegen sie verstärkt auf die Systemausgaben achten und Use Cases so lange wiederholen können, bis sie alle Ausgaben bewerten können. In Nutzerstudien kann es vorkommen, dass das Erleben der Use Cases für eine Bewertung nicht ausreichend ist, beispielsweise, wenn die Anzeigen nicht gesehen werden. Beispielsweise ist davon auszugehen, dass Nutzer nicht die vollständige Warnkaskade einer Hands-Off-Warnung erleben. Weiterhin sind Experten in der Lage, aufgrund der Checkliste auch ohne einen direkten Vergleich ein gutes von einem schlechten System zu unterscheiden, wenn die Items und Anforderungen nicht erfüllt sind. Ein Nutzer, der über keinerlei Vorerfahrung mit einem System verfügt und keinen Vergleich zu anderen Systemen hat, bewertet unter Umständen auch ein mangelhaftes System als gut, wenn er selbst damit zurechtkommt.

Auch falsche Vorannahmen über die Funktionsweise eines L2-Systems, die die Ergebnisse von Nutzerstudien beeinflussen können, sollten bei der Expertenbewertung eine geringere Rolle spielen als bei Probandenstudien.

### **3 Szenarienbasiertes Beobachtungsinstrument für die Bewertung der Mensch-Maschine-Interaktion in Nutzerstudien**

Neben der Methode der Expertenbewertung wurde die Methode der Nutzerstudie als zweiter Baustein einer ganzheitlichen Bewertung der Effizienz und Sicherheit der MMI beim teilautomatisierten Fahren definiert. Mit ihr ist es möglich, interaktionsbezogene Aspekte zu erfassen, die Aufschluss darüber geben, ob das System bzw. die MMS dem Fahrer ein angemessenes System- und Situationsbewusstsein vermitteln kann, effizient bedient werden kann und der Fahrer an Systemgrenzen sicher eingreifen kann. Als geeignete Prüfmaße wurden die Beobachtung und Befragung der Fahrer in den definierten Use Cases (siehe Kapitel 1.5) festgelegt. Im folgenden Kapitel wird ein Bewertungsinstrument vorgestellt, das den Versuchsleiter bei dieser szenarienbasierten Fahrerbeobachtung und -befragung in Nutzerstudien unterstützt. Es beschreibt die Entwicklung und Anwendung dieser Methode, die als Applikation zur Nutzung mittels eines Tablets verwendet werden kann.

#### **3.1 Anforderungen an das Bewertungsinstrument**

Die Bewertung von MMI beim automatisierten Fahren kann sowohl in der Fahrsimulation als auch auf Teststrecken oder im Realverkehr stattfinden. Die Prüfung in der Fahrsimulation oder auf der Teststrecke hat den Vorteil, dass dort für alle Probanden vergleichbare Prüfbedingungen hergestellt werden können, die Prüfsituationen klar spezifiziert sind, zu erwartende Einflussfaktoren auf die Bewertung gezielt kontrolliert bzw. aktiv manipuliert werden können (z. B. Reihenfolgeeffekte), das Fahrverhalten hochzeitauflösend und detailliert analysiert werden kann (z. B. Reaktionszeiten auf eine Übernahme-situation) und auch Systeme, die noch nicht in Serie

sind in einem frühen Entwicklungsstand untersucht werden können. Es kann ein fester Prüfparcours definiert werden, den alle Nutzer in gleicher Form absolvieren, sodass eine Gesamtbewertung des Systems auf derselben Grundlage beruht. Die Prüfung im Realverkehr stellt höhere Anforderungen an das Bewertungsinstrument. Aufgrund schwer kontrollierbarer Ausgangsbedingungen durch unterschiedliche Verkehrskonstellationen (z. B. können manche Use Cases gar nicht erlebt werden oder die Reihenfolge innerhalb einer Prüffahrt ändert sich beispielsweise durch das plötzliche Auftreten einer Systemgrenze) muss es flexibler in der Anwendung sein. Außerdem steht in der Regel kein instrumentiertes Fahrzeug zur Verfügung, mit dem unmittelbar Daten des Fahrzeugverhaltens oder Fahrereingaben erfasst werden können.

Um die Methode in den verschiedenen Testumgebungen möglichst effizient einzusetzen, sollte sie ohne großen Implementierungsaufwand einsetzbar sein. Außerdem sollte die Analyse der Daten mit geringem Zeit- und Kostenaufwand erfolgen können. Für die hier zugrundeliegende Fragestellung erscheint es nicht notwendig, eine aufwendige zeitverlaufsbasierte Datenanalyse vorzunehmen, z. B. bezüglich exakter Reaktionszeiten auf Übernahme-situationen. Vielmehr ist es empfehlenswert, die Erfassung des Fahrerverhaltens über eine Beobachtung der Interaktion durch einen geschulten Rater vornehmen zu lassen. Anstelle einer rein beschreibenden Dokumentation dieses Verhaltens wird empfohlen, normativ vorzugehen, d. h. das beobachtete Verhalten direkt anhand vorgegebener Kategorien zu klassifizieren. Darüber hinaus werden nur fehlerhafte Verhaltensweisen entsprechend kodiert, um Probleme in der Interaktion mit dem System bzw. Optimierungsbedarf aufzudecken.

Zudem sollte keine aufwendige Transkription von Daten, die in nicht-digitaler Form erfasst werden, notwendig sein. Eine computergestützte Dateneingabe, beispielsweise mittels einer Tablet-Anwendung, die auch problemlos in ein Testfahrzeug transportiert werden kann, ist daher empfehlenswert. Im Optimalfall kann die Analyse der mit dieser Anwendung erhobenen Daten zumindest teilweise automatisiert und unmittelbar nach der Durchführung der Nutzerstudie eine Bewertung des Systems erfolgen. Dies gilt unter der Voraussetzung, dass die Anwendung Zugriff auf quantitative Daten hat, wie es in der Fahrsimulation der Fall sein kann.

Eine weitere Anforderung ist die standardisierte Anwendbarkeit der Methode. Für den Vergleich zwischen Systemen sollte die Bewertung der MMS beim teilautomatisierten Fahren unter vergleichbaren Bedingungen vorgenommen werden, was die Prüfumgebung, die Prüfscenarien und die Prüfkriterien anbelangt. Ziel muss daher ein reliables und objektives Urteil über die MMS sein, d. h. das Urteil muss zu einer anderen Zeit, mit anderen Nutzern und anderen Versuchsleitern wieder genauso ausfallen. Dies erfordert eine klare Spezifikation des Prüfablaufs, der Prüfscenarien und der Bewertungskriterien. Die Bewertungskriterien müssen objektiv beobachtbar und anhand klarer Regeln zu vergeben sein und sollten nicht auf subjektiven Mutmaßungen über potenziell zugrundeliegende Ursachen beruhen. Wenn ein Fahrer bei der Aktivierung des Systems beispielsweise sehr lange braucht, bis er die Aktivierung erfolgreich abschließen kann, wird dieses Verhalten als zögerlich bewertet. Die Ursache dafür, z. B. dass er wahrscheinlich nicht genau weiß, welche Bedienhandlungen dafür nötig sind, weil die Tasten nicht eindeutig beschriftet sind, sollte dabei zunächst nicht relevant sein. Umgekehrt bedeutet dies aber auch, dass über diese Methode die Ursachen für ein Fehlverhalten nicht aufgedeckt werden können. Dies erfordert weitere Befragungen nach dem „Warum“, die im Nachgang der Studie durchgeführt werden können.

Nachdem sich die zu bewertenden Systeme in Bezug auf die Systemauslegung und die Gestaltung der MMI unterscheiden können, sollte eine Bewertung auf Basis derselben Grundlage eines festen Sets von Szenarien erfolgen, anhand derer alle Systeme geprüft werden. Dazu dienen die typischen Prüffälle, die in Kapitel 1.5 hergeleitet wurden. Um zu verhindern, dass in einer nachträglichen Gesamtbewertung eines Systems einzelne Aspekte über- oder unterrepräsentiert werden, bzw. Erinnerungseffekte (man erinnert sich nicht mehr an einzelne Situationen) oder sog. Primacy-Recency-Effekte wirksam werden (Primacy-Effekt: der erste Kontakt mit dem System überstrahlt die nachfolgenden Interaktionen bzw. Recency-Effekt: vor allem das zuletzt erlebte Ereignis wird besonders stark in die Bewertung miteinbezogen), sollte das Bewertungsinstrument die Interaktion unmittelbar in oder nach der Situation erfassen. Zudem ermöglicht diese Vorgehensweise, dass gezielter Optimierungsbedarf in einzelnen Use Cases identifiziert werden kann und ein System nicht unspezifisch aufgrund eines Gesamteindrucks als „schlecht“ oder „gut“ bewertet wird.

Zusammenfassend erfüllt das Bewertungsinstrument folgende Anforderungen:

- Flexibel einsetzbar in verschiedenen Testumgebungen,
- einfach implementierbar,
- unaufwändige, effiziente Dateneingabe und -analyse,
- standardisiertes Vorgehen bzgl. Prüfablauf, Prüfscenarien und Prüfkriterien,
- objektive und normative Bewertung anhand klar definierter Vorgaben,
- in Situ, d. h. Online-Erfassung der Interaktion und des Systemerlebens,
- szenarienbezogene Analyse der Interaktion.

### 3.2 Definition der Bewertungskategorien

Als nächstes müssen die relevanten Bewertungskategorien definiert werden, die innerhalb der Nutzerstudie geprüft werden sollen. Für eine standardisierte, objektive und effiziente Erfassung der Beobachtungsaspekte werden untergeordnete Problem- und Fehlerkategorien definiert. Diese Kategorien wurden basierend auf der langjährigen Erfahrung der Autorinnen aus vielfältigen Studien (Simulationsstudien und Realverkehrsstudien; z. B. in Bezug auf die Bewertung von Übernahmesituationen; siehe NAUJOKS et al., 2018a – TOC-Rating Methode) mit derartigen Systemen abgeleitet und spiegeln die häufigsten zu erwartenden Probleme mit L2-Systemen wieder.

#### 3.2.1 Beobachtungskategorien

Die Beobachtungskategorien gliedern sich in folgende drei Bereiche:

- Bedienprobleme (vor allem für fahrerinitiierte Bedienhandlungen relevant)
  - Nichts bemerkt: Fahrer bemerkt die Änderung eines Systemzustands nicht.
  - Unsicher/verzögert: Fahrer zeigt Unsicherheiten bei der Bedienung, z. B. sucht nach richtiger Taste, überlegt lange bevor er die richtige Aktion ausführt.

- Unangemessen: Fahrer zeigt unangemessene Bedienhandlung; z. B. aktiviert ein System/Teilfunktion in einem Abschnitt, wo es unangemessen ist, z. B. Querführung auf der Landstraße, bzw. deaktiviert das System, wo es eigentlich automatisch reagiert hätte.
- Bedienfehler: Fahrer führt eine Bedienaktion aus, die nicht unmittelbar zum Ziel führt, z. B. drückt die Taste nicht fest genug, will System in einem Abschnitt aktivieren, in dem es nicht verfügbar ist, drückt z. B. fälschlicherweise erneut eine Taste und deaktiviert dadurch ein Teilsystem, obwohl er eigentlich ein anderes Teilsystem aktivieren wollte.
- Hilfestellung nötig: Fahrer führt bis zu einem definierten Zeitpunkt keine Aktion aus, die erfolgreich zum Ziel führt, weshalb der Versuchsleiter eingreifen und Hilfestellung geben muss, damit der Fahrer in den erwarteten Systemzustand gelangt.

- Probleme in der Fahrzeugführung (für fahrerinitiierte Bedienhandlungen und systeminitiierte Transitionen relevant)
  - Keine Reaktion: Fahrer zeigt in einer bestimmten Situation, die eigentlich eine Handlung von ihm erfordert, keinerlei Reaktion.
  - Verzögert: Fahrer reagiert verzögert auf ein Ereignis, z. B. mit einem Bremsmanöver.
  - Fahrstreifenabkommen: das Fahrzeug gerät mit den Reifen einer Fahrzeugseite über die Spurmarkierung.
  - Zu stark: Fahrer reagiert zu stark; z. B. mit Übersteuern des Lenkrads oder mit einer unnötigen Notbremsung.
  - Schlechte Spurhaltung: Fahrer schwankt innerhalb seines Fahrstreifens nach links und/oder rechts.
  - Unzureichend gesichert: Fahrer führt bei einem Fahrstreifenwechsel keinen Kontrollblick aus.
  - Gefährdung: Fahrzeug unterschreitet einen Mindestabstand von 1 m zur Seite bzw. 1 Sekunde nach vorne oder hinten zu einem anderen Fahrzeug.
  - Kollision: Fahrzeug kollidiert mit einem anderen Verkehrsteilnehmer oder Objekten in der Verkehrsumgebung z. B. einer Leitplanke.

- Auffälligkeiten im Überwachungsverhalten (für das längerfristige Fahren in L2 relevant)
  - Unsicher Hände dran: Fahrer zeigt deutliche Unsicherheiten darin, ob er die Hände am Lenkrad lassen soll oder nicht, z. B. nimmt mehrfach kurz die Hände weg und wieder dran, ohne dass Hands-Off-Warnung ausgelöst wird.
  - Nicht aufmerksam genug: Fahrer zeigt deutliche Anzeichen von Unaufmerksamkeit; z. B. schaut sich verstärkt die Landschaft an, kontrolliert deutlich zu selten den Systemzustand über Blicke auf die MMS.
  - Hands-Off-Warnung: Hands-Off-Warnung hat ausgelöst.
  - Stufe der Hands-Off-Warnung (HO).

### 3.2.2 Befragungskategorien

Neben den Beobachtungskategorien gibt es Befragungsvariablen, die je nach Art des Szenarios (system- vs. fahrerinitiierte Transition) eingesetzt werden können. Diese soll der Proband während der Fahrt unmittelbar nach einer bestimmten Situation beantworten. Folgende Fragen erscheinen für die Bewertung der MMS durch den Fahrer relevant:

- Frage zur Verständlichkeit der notwendigen Fahrerreaktion: Weiß der Fahrer, was er in einer bestimmten Situation tun muss, z. B. um das System zu aktivieren, zu deaktivieren, angemessen auf eine Systemgrenze zu reagieren?
- Frage zur Nachvollziehbarkeit des Systemverhaltens: Versteht der Fahrer, warum sich das System in einer bestimmten Situation derart verhält, z. B. die Querführung ausfällt?
- Frage zur Verständlichkeit der Systemausgaben: Versteht der Fahrer, was eine Systemanzeige und/oder eine akustische Ausgabe bedeutet?
- Frage zur wahrgenommenen Situationskritikalität: Wie kritisch empfindet der Fahrer eine Situation als Resultat aus der Kombination der objektiven Situation und der von ihm geforderten Verhaltensreaktion in dieser Situation?

Darüber hinaus ergeben sich je nach konkreter Fragestellung weitere Befragungskategorien, wie beispielsweise die wahrgenommene Sicherheit beim Fahren mit dem System oder das wahrgenommene

-3 Sehr wenig <input type="checkbox"/>	-2 Wenig <input type="checkbox"/>	-1 Eher wenig <input type="checkbox"/>	0 Weder noch <input type="checkbox"/>	1 Eher stark <input type="checkbox"/>	2 Stark <input type="checkbox"/>	3 Sehr stark <input type="checkbox"/>
----------------------------------------------	-----------------------------------------	----------------------------------------------	---------------------------------------------	---------------------------------------------	----------------------------------------	---------------------------------------------

Bild 3-1: 7-stufige bipolare Ratingskala zur Beantwortung der Fragen

Systemvertrauen, die optional zusätzlich abgefragt werden können.

Als Antwortskala wird eine 7-stufige, verbalisierte, bipolare Ratingskala mit Skalenwerten von -3 bis +3 vorgeschlagen, die flexibel an die Fragenkategorien angepasst werden kann (siehe Bild 3-1). In Übersichtsarbeiten kommen KROSNICK et al. (KROSNICK & FABRIGAR, 1997; KROSNICK & PRESSER, 2010 aus MENOLD & BOGNER, 2015) zum Schluss, dass eine optimale Messung — in Bezug auf die Reliabilität, Validität und den Differenzierungsgrad — mit fünf bis sieben Kategorien erreicht werden kann. Für die Online-Anwendung während der Fahrt hat die begrenzte Anzahl an Kategorien zusätzlich den Vorteil, dass sie von den Fahrern schnell erfasst und leicht erlernt werden kann. Alle Skalenwerte sind verbalisiert, was nach MENOLD & BOGNER (2015) die Reliabilität (SARIS & GALLHOFER, 2007) und die Validität (KROSNICK & FABRIGAR, 1997) erhöht. Außerdem berichten sie, dass Personen vollverbalisierte Ratingskalen bevorzugen (z. B. WALLSTEN, BUDESCU, & ZWICK, 1993; ZALLER, 1988). Die zusätzlichen numerischen Werte dienen der schnelleren Beantwortung der Fragen, die online während der Fahrt beantwortet werden sollen. Zur Unterstützung der Beantwortung wird empfohlen, die Skalen so im Fahrzeug anzubringen, dass der Fahrer diese mittels eines kurzen Blicks erfassen kann (z. B. Befestigung auf dem Lenkrad). Zudem sollten dem Fahrer die Fragen, die ihm während der Fahrt gestellt werden, vor der Fahrt kurz erklärt werden.

Für die Frage zur wahrgenommenen Situationskritikalität wird auf eine im Rahmen von Kontrollierbarkeitsstudien bewährte Ratingskala, die sog. SBFV (Skala zur Bewertung von Fahr- und Verkehrssituationen, NEUKUM et al., 2008; abgeleitet von der Störungsbewertungsskala nach NEUKUM & KRÜGER, 2003 zur Bewertung der Kontrollierbarkeit von Lenkeingriffen; siehe Bild 3-2) zurückgegriffen. Hierbei soll zweistufig vorgegangen werden: Der Fahrer entscheidet sich zunächst für eine der fünf Verbalkategorien. Die drei mittleren Stufen werden in einem zweiten Schritt durch ein Zahlenurteil feiner abgestuft. Der Fahrer soll die Situationen so be-

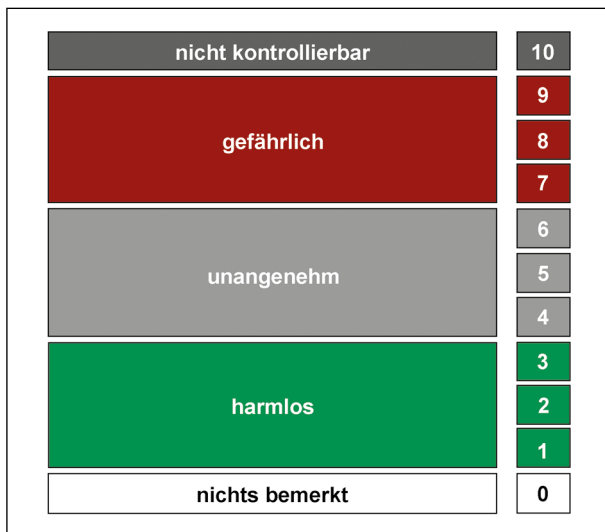


Bild 3-2: SBFV-Skala zur Bewertung der Situationskritikalität (nach NEUKUM et al., 2008)

urteilen, als ob sie im realen Straßenverkehr aufgetreten wären.

- Eine „0“ wird vergeben, wenn der Fahrer die Situation als solche (d. h. das Auftreten eines besonderen Ereignisses) nicht wahrgenommen/bemerkt hat und entsprechend kein Urteil der Kritikalität abgeben kann.
- Als „harmlos“ (Urteile 1-3) werden Situationen gekennzeichnet, in denen kein oder nur ein geringer Aufwand des Fahrers zur Bewältigung der Fahraufgabe notwendig ist.
- Die Kategorie „unangenehm“ (Urteile 4-6) umfasst Situationen, die einen deutlichen, aber vom Fahrer vertretbar bewerteten Aufwand erfordern.
- Die Situation wird als „gefährlich“ (Urteile 7-9) eingestuft, wenn die Anforderungen an den Fahrer hoch sind und die Kritikalität der entstandenen Situation vom Fahrer als nicht mehr tolerierbar beurteilt wird.

### 3.2.3 Gesamturteil

Zusätzlich soll der Versuchsleiter pro Szenario auch ein Gesamturteil abgeben, wie gut der Fahrer ein jeweiliges Szenario bewältigt hat. Dieses soll sich aufgrund der beobachteten Fehler bzw. Probleme im Umgang mit dem System ergeben und als ein globales Maß zur Bewältigung der Situation dienen.

Dieses ist angelehnt an die Fitness-to-Drive-Skala (FtD-Skala; nach KAUSSNER et al., 2013 sowie KENNTNER-MABIALA et al., 2015; in Anlehnung

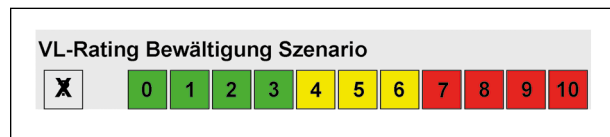


Bild 3-3: Ratingskala zur Vergabe eines globalen Versuchsleiter-Urteils zur Bewältigung einer Situation

an NEUKUM & KRÜGER, 2003; siehe Bild 3-3), die in der S.A.F.E.-Anwendung eingesetzt wird (siehe KAUSSNER; HOFFMANN & KENNTNER-MABIALA, 2013). Für diese Ratings konnte im Rahmen von Schulungen eine Interrater-Reliabilität zwischen  $r = .833$  und  $r = .944$  erreicht werden. Für die Anwendung auf die Fragestellungen beim automatisierten Fahren wurden die dort verwendeten Kategorien entsprechend angepasst.

Grundprinzip des Versuchsleiterratings ist es, ein Gesamturteil basierend auf den beobachteten Fehlern spezifisch für jedes Szenario zu vergeben. Je nach Szenario können unterschiedliche Kriterien zu einer Einstufung in die verschiedenen Kategorien führen, z. B. wird die 10 vergeben, wenn ein Fahrer auch nach wiederholter Hilfestellung nicht in der Lage ist, das System in einen geforderten Zustand zu versetzen bzw. wenn er in Folge einer Systemgrenze eine Kollision verursacht hat. Entsprechend müssen die Urteile für jeden Szenariotyp gesondert interpretiert werden (z. B. Übernahmeszenarien getrennt von Aktivierungsszenarien).

Die Definition der Schweregrade der Fehler sind als grobe Empfehlungen zu verstehen, in welcher Kategorie auf der Ratingskala der Versuchsleiter das Verhalten einordnen sollte. Weiterhin muss nach eigenem Ermessen berücksichtigt werden, welche Konsequenzen aus dem Fehlverhalten resultieren und diese in das Urteil miteinbezogen werden. Entsprechend kann innerhalb einer Verbalkategorie das Urteil weiter verfeinert werden (d. h. ob eher eine 4 oder eine 6 vergeben wird). In Tabelle 3-1 sind Empfehlungen zur Vergabe des Versuchsleiter-Ratings dargestellt.

## 3.3 Umsetzung der Methode in eine Tablet-Applikation

Zur Unterstützung des Versuchsleiters in der szenarienbasierten Fahrerbeobachtung und -befragung in Nutzerstudien anhand der im vorigen Kapitel definierten Kriterien wurde die sog. S.A.D.E.-App entwickelt. Das Akronym S.A.D.E. steht für Standardized Application for Automated Driving Evaluation



Fehlerart	Verbalkategorie	Numeralkategorie
<ul style="list-style-type: none"> <li>Im Bedienverhalten: Auch nach Hilfestellung keine erfolgreiche Bedienung</li> <li>Im Fahrverhalten: Auftreten einer Kollision mit einem vorausfahrenden Fahrzeug oder einem Fahrzeug auf dem Nachbarfahrstreifen</li> <li>Im Überwachungsverhalten: Hands-Off-Warnung auf Stufe 3 (d. h. Abschaltung der Querführung aufgrund Nicht-Beachten der Hands-Off-Warnungen)</li> </ul>	Nicht bewältigt	10
<ul style="list-style-type: none"> <li>In der Bedienung: nichts bemerkt in Bezug auf Situationen, in denen eine Transition auf eine niedrigere Stufe notwendig ist (z. B. bei einer Systemgrenze), Hilfestellung nötig, aber anschließend erfolgreiche Bedienung</li> <li>Im Fahrverhalten: keine oder verzögerte Reaktion mit der Konsequenz einer Gefährdung des vorausfahrenden Verkehrs oder des Verkehrs auf dem Nachbarfahrstreifen</li> <li>Im Überwachungsverhalten: mehr als eine Hands-Off-Warnung und/oder eine Hands-Off-Warnung auf Stufe 2</li> </ul>	Nicht-akzeptable Probleme	9
		8
		7
<ul style="list-style-type: none"> <li>In der Bedienung: Fehlbedienung (z. B. falsche Taste) und/oder unangemessene Bedienung oder nicht erfolgreiche Bedienung</li> <li>Im Fahrverhalten: verzögerte oder zu starke Reaktion, vor allem in Bezug auf ein verzögertes Bremsverhalten und/oder Spurabkommen</li> <li>Im Überwachungsverhalten: Eindeutige Anzeichen für unaufmerksames Fahren, mindestens eine Hands-Off-Warnung in Stufe 1</li> </ul>	Fehlerhafte, aber akzeptable Bewältigung	6
		5
		4
<ul style="list-style-type: none"> <li>In der Bedienung: Kategorie „unsicher/verzögert“</li> <li>Im Fahrverhalten: Fahrfehler Spurschwankung oder unzureichendes Sichern</li> <li>Unsicherheiten im Überwachungsverhalten des Fahrers: unsicher, ob Hände am Lenkrad gelassen werden müssen</li> </ul>	gute Bewältigung	3
		2
		1
o Kein Fehler	perfekte Bewältigung	0

Tab. 3-1: Übersicht über die Empfehlungen zur Vergabe des Versuchsleiter-Ratings basierend auf den pro Szenario beobachteten Fehlern

und ist eine Applikation der Software SILAB® zur Durchführung von Fahrverhaltensbeobachtungen und Fahrerbefragungen beim teilautomatisierten Fahren auf Level 2 anhand eines Tablets.

Die App wurde in Anlehnung an ein entsprechendes Verfahren zur Fahrverhaltensbeobachtung im manuellen Fahren, dem sog. S.A.F.E.-Verfahren entwickelt (siehe KAUSSNER, 2013; KAUSSNER; HOFFMANN & KENNTNER-MABIALA, 2013). Der Testleiter kann mit der Applikation die im vorigen Kapitel beschriebenen Fehler bei der Systembedienung, im Fahrverhalten sowie im Überwachungsverhalten über die Touch-Oberfläche protokollieren, die Bewältigung einer Situation als Ganzes einstufen sowie standardisierte Befragungen während einer Messfahrt durchführen. Durch die anschauliche und übersichtliche visuelle Darstellung der Fehlerkategorien und der Befragungen inklusive Rating-Skalen sowie – falls die Durchführung im Simulator erfolgt – durch eine teilweise automatisierte Fehlerzählung für bestimmte Kategorien wird der Versuchsleiter in der Beobachtung unterstützt und die Objektivität des Verfahrens erhöht.

### 3.3.1 App-Oberfläche

Basierend auf den beiden Bewertungsdimensionen der Beobachtung und der Befragung ist die Oberfläche der S.A.D.E.-App in diese beiden Bereiche aufgeteilt (siehe Bild 3-4):

- Beobachtungsvariablen (links),
- Befragungsvariablen (rechts).

Die Beobachtungsvariablen sind dabei in die drei Bereiche „Bedienprobleme“, Probleme in der Fahrzeugführung (Kurzbezeichnung in App: „Fahrzeugführung“) und Auffälligkeiten im Monitoringverhalten (Kurzbezeichnung in App: „Monitoring“) unterteilt. Der Großteil der Beobachtungskategorien ist in numerischer Form hinterlegt (mit Ausnahme der Stufe der Hands-Off-Warnung), sodass der Versuchsleiter manuell durch Anklicken des entsprechenden Buttons den Wert jeweils nach oben bzw. nach unten verändern kann. Unten links ist die Skala für das Versuchsleiterrating (VL-Rating) zur Bewältigung des Szenarios angezeigt. Aus Platzgründen sind hier nur die numerischen Kategorien der Skala dargestellt.



Szenario		1_Aktivierung_a (Quer- und Längs)					
S.A.D.E. ist aktiv / Szenario-ID 11	Bedienprobleme	nichts bemerkt	unsicher/verzögert	unangemessen		<b>Instruktion:</b> Bitte aktivieren Sie das komplette System, d.h. die Längs- und die Querführung	
		0	1	0			
	Bedienfehler	Hilfestellung nötig			1. Wie klar war Ihnen, was zu tun ist?		
	1	0			<input type="text" value="?"/> -3 -2 <del>X</del> 0 1 2 3		
	Fahrzeugführung	keine Reaktion	verzögert	zu starke Reaktion	Spurabkommen	2. Wie nachvollziehbar war das Systemverhalten?	
		0	0	0	0	<input checked="" type="text" value="X"/> -3 -2 -1 0 1 2 3	
		schlechte Spurhaltung	unzur. gesichert	Gefährdung	Kollision	3. Wie verständlich waren die Systemausgaben?	
		0	0	0	0	<input checked="" type="text" value="X"/> -3 -2 -1 0 1 2 3	
	Monitoring	unsicher, Hände dran	nicht aufmerksam	Handsoff-Warnung	Stufe HO-Warnung	4. Wie sicher fühlen Sie sich beim Fahren mit dem System?	
		0	0	0		<input checked="" type="text" value="X"/> -3 -2 -1 0 1 2 3	
VL-Rating Bewältigung Szenario					5. Wie kritisch war die Situation?		
<input type="text" value="?"/> 0 1 2 3 <del>X</del> 5 6 7 8 9 10					<input checked="" type="text" value="X"/> 0 1 2 3 4 5 6 7 8 9 10		

Bild 3-4: Tablet-Benutzeroberfläche mit Beobachtungs- und Befragungsvariablen (Quelle: WIVW GmbH)

Auf der rechten Seite sind die Befragungsvariablen mit den dazugehörigen Ratingskalen (für die Fragen 1-5: Skala von -3 bis +3; für die Frage 6 elf-stufige SBFV) dargestellt, auf die der Versuchsleiter klicken kann. Folgende Befragungsvariablen wurden in der App umgesetzt:

- Frage 1: Wie klar war Ihnen, was zu tun ist?
- Frage 2: Wie nachvollziehbar war das Systemverhalten?
- Frage 3: Wie verständlich waren die Systemausgaben?
- Frage 4: Optionale Zusatzfrage: kann je nach konkreter Fragestellung frei genutzt werden, um spezifische Fragen an den Probanden zu stellen – im Beispielbild ist die Frage „Wie sicher fühlen Sie sich beim Fahren mit dem System?“ eingesetzt
- Frage 5: Wie kritisch war die Situation?

Oben rechts kann pro Szenario eine Instruktion für den Versuchsleiter eingefügt werden, wie der Proband anzuleiten ist (d. h. welche Aufgabe er ausführen soll). Für eine gute Orientierung des Versuchsleiters wird der aktuelle Use Case als Titel oben mittig in der App eingeblendet. In der Fahrsimulation kann hier ein Szenariename sowie eine Szenari-

enummer hinterlegt werden. Die Buttons rechts oben auf der Benutzeroberfläche dienen der Bedienung der App.

S.A.D.E. bietet eine übersichtliche Testdokumentation: Alle Eingaben des Testleiters sowie die automatisch registrierten Fehler werden am Ende einer Messfahrt zusammen mit probanden- oder studienbezogenen Daten (Probanden-Code, experimentelle Bedingung etc.) in einer übersichtlichen Aufzeichnungsdatei zusammengefasst. Diese Datei kann problemlos in gängige Statistikprogramme importiert werden. Zudem wird pro Szenario ein Screenshot mit dem ausgefüllten Formular (Bewertungsbogen der Tablet-App) abgelegt.

### 3.3.2 Anwendung der App in der Fahrsimulation

Im Folgenden wird die generelle Anwendung der Tablet-App für Untersuchungen im Fahrsimulator beschrieben. Für die Einarbeitung in die Methode liegen Schulungsmaterialien vor, die sowohl eine Präsentation als auch eine ausgearbeitete schriftliche Dokumentation beinhalten, anhand deren Versuchsleiter für entsprechende Studien, in denen die Tablet-App zum Einsatz kommt, vorbereitet werden können.

In Simulatorstudien ist die Reihenfolge und Dauer der Use Cases (bzw. Szenarien) vorab definiert und entsprechend in der App hinterlegt. Ein Prüfzenario ist dabei so definiert, dass es eine bestimmte Vorlaufzeit vor dem Abschnitt hat, in dem das zu bewertende Verhalten in Reaktion auf ein Ereignis oder einer Instruktion erfolgen soll. In einem sich daran anschließenden Abschnitt kann der Versuchsleiter die Befragung durchführen. Es wird empfohlen, bereits während des zu beobachtenden Ereignisses die aufgetretenen Fehler zu kodieren. Um alle relevanten Fehler beobachten zu können, muss der Versuchsleiter die Möglichkeit haben, zum einen die Bedienhandlungen des Fahrers am Lenkrad einsehen zu können als auch das Blickverhalten des Fahrers beobachten zu können, um Aussagen zum angemessenen Überwachungs- und Sicherungsverhalten machen zu können.

Im Optimalfall ermöglichen ihm dies zwei Kameras, die gut positioniert im Simulator-Mock-Up angebracht sind und sowohl das Fahrergesicht von vorne, als auch das Lenkrad mit Bedienelementen und die Handhaltung des Fahrers zeigen. Für eine Einschätzung der Fahrerleistung, vor allem bezüglich der Spurhaltegröße sollte die Fahrszenerie zusätzlich in Vogelperspektive auf einem Monitor gezeigt werden (siehe Bild 3-5). Im Anschluss an die Situation sollte das Versuchsleiter-Rating auf der Skala abgegeben werden. Erst danach sollte die Befra-

gung des Fahrers erfolgen. Die Längen der einzelnen Prüfzenarien sind so zu gestalten, dass die Beobachtung und die Befragung in einem angemessenen Zeitrahmen möglich sind. Dennoch sollten die Probanden vorab instruiert werden, möglichst zügig und spontan auf die Fragen zu antworten und ausführliche Kommentare erst in der Befragung nach der Fahrt zu äußern.

Nur während eines laufenden Szenarios sind Eingaben durch den Versuchsleiter möglich. Ist das Ende eines jeweiligen Szenarios erreicht, sind keine Eingaben mehr möglich, die App wechselt automatisch zum neuen Szenario, in dem wieder neue Eingaben gemacht werden können. Nachdem alle Prüfzenarien absolviert wurden, kann mithilfe einer Rückschaufunktion durch die Fehlerprotokolle aller Szenarien durchgeblättert werden, um nachträglich Korrekturen vorzunehmen oder fehlende Ratings zu ergänzen.

Zur Erleichterung der Beobachtung und Befragung sind folgende Features in der App integriert:

- Automatische Fehlerzählung: Neben den Fehlern, die der Versuchsleiter manuell kodiert, gibt es Fehler, die automatisch von der App gezählt werden, um dem Versuchsleiter die Beobachtung zu erleichtern. Diese automatische Fehlerzählung funktioniert allerdings nur für Simulatorstudien, in denen die App Zugriff auf die von



Bild 3-5: Optimale Bildschirm-Anordnung zur Mitschau für den Versuchsleiter während des Versuchs im Simulator (Quelle: WIVW GmbH)

SILAB® aufgezeichneten Daten hat, nicht jedoch in Realfahrtstudien. Folgende Fehler werden automatisch gezählt:

- Spurabkommen: Wird immer dann gezählt, wenn das Fahrzeug die Fahrstreifenmarkierungen mit einem Reifen überfährt. Geschieht dies willentlich, z. B. im Fall eines beabsichtigten Fahrstreifenwechsels, muss dieser Wert manuell nach unten korrigiert werden.
- Gefährdung: Wird automatisch gezählt, wenn ein Mindestabstand von 1 m zur Seite bzw. 1 Sekunde nach vorne oder hinten zu einem anderen Fahrzeug unterschritten wird.
- Kollision: Wird automatisch gezählt, wenn das Fahrzeug mit einem anderen Verkehrsteilnehmer kollidiert. Bei feststehenden Objekten in der Verkehrsumgebung, z. B. Leitplanken, erfolgt keine automatische Fehlerzählung. In diesem Fall ist eine manuelle Kodierung notwendig.
- Hands-Off-Warnung: Dieser Fehler wird automatisch gezählt, wenn die Hands-Off-Warnung ausgelöst hat.
- Stufe HO-Warnung: Die Stufe der Hands-Off-Warnung wird automatisch erfasst (Anzahl und Ausgestaltung der Stufen sind systemspezifisch; jeweils die höchste Stufe innerhalb eines Abschnitts wird aufgezeichnet).
- Der Versuchsleiter erhält ein akustisches Signal, wenn ein neues Szenario beginnt.
- Beobachtungsvariablen, die in einem bestimmten Szenario nicht beobachtet werden müssen bzw. Befragungsvariablen, die in einem bestimmten Szenario nicht sinnvoll sind, werden ausgegraut.
- Das bevorstehende Ende eines Szenarios, bis zu dem die Beobachtung bzw. Befragung abgeschlossen sein sollte, wird farbig visualisiert.
- Ist sich der Versuchsleiter bei der Bewertung unsicher oder ist die Zeit nicht mehr ausreichend, um das VL-Rating abzugeben oder eine bestimmte Frage zu stellen, kann er den Wert auf dem Fragezeichen stehenlassen und gegebenenfalls nach der Fahrt ausfüllen.

### 3.3.3 Anwendung der App im Realverkehr

Neben dem Einsatz in der Fahrsimulation kann die Methode auch im Realfahrzeug im Realverkehr oder auf einer Teststrecke eingesetzt werden. Grundsätzlich ist dabei der Einsatz auf der Teststrecke zu empfehlen, da kontrollierte Bedingungen hergestellt und sicherheitskritische Systemgrenzen ohne den Probanden zu gefährden, erlebbar gemacht werden können. Wenn aus ökonomischen Gründen oder der fehlenden Verfügbarkeit eines abgesperrten Prüfgeländes auf die Untersuchung im Realverkehr zurückgegriffen werden muss, sind einige Anpassungen der Methode erforderlich.

Im Rahmen einer Realfahrtstudie werden die Use Cases nicht experimentell erzeugt, sondern müssen dem Fahrer zur Bewertung erlebbar gemacht werden. Für fahrerinitiierte Transitionen oder Fahren in einer Automationsstufe kann dies durch entsprechende Aufforderungen zu Bedienhandlungen erfolgen. Systeminitiierte Transitionen können nur in begrenztem Maß durch eine geeignete Auswahl der Strecke „erzeugt“ werden (z. B. Ausfälle der Querführung an Stellen, in denen die Sensorik die Fahrstreifenmarkierung schlecht erkennt). Vor Beginn der Testfahrt muss der Proband zunächst im Stand die Bedienungsanleitung bzw. eine vorbereitete schriftliche Kurzform durchlesen. Der Versuchsleiter erklärt anschließend die zu benutzenden Skalen und gibt generelle Instruktionen über den Testablauf. Der Proband sollte dabei darauf hingewiesen werden, dass die Gewährleistung der Fahrsicherheit immer oberste Priorität haben muss und er nur wenn er sich dabei sicher fühlt, entsprechende Use Cases auf Anweisung durchführen soll. Nachdem auf die Autobahn aufgefahren wurde, kann mit der Instruktion und Durchführung der Use Cases begonnen werden.

In Bezug auf die zu bewertenden Use Cases muss betont werden, dass der Fahrer aus ethischen Gründen nicht absichtlich und unvorbereitet in kritische Situationen wie plötzliche Systemgrenzen gebracht werden sollte. Für derartige Use Cases eignen sich Studien auf abgesperrten Teststrecken. Als unproblematisch werden geplante, sicherheitsunkritische Systemgrenzen, wie das Abfahren von der Autobahn, in denen der Fahrer das System ohnehin abschalten muss, betrachtet.

Folgende Use Cases werden in der hier vorgestellten Reihenfolge als sinnvoll erachtet:

1. Aktivierung des Systems, d. h. der kombinierten Quer- und Längsführung: Dies kann je nach System unterschiedlich erfolgen:
  - D. h. entweder über eine schrittweise Aktivierung von zunächst der Längsführung, anschließend der Querführung oder
  - über ein Hinzuschalten der Längsführung zu einer bereits standardmäßig aktiven Querführung oder
  - über eine direkte Aktivierung des Gesamtsystems.
  - Die Anpassung der Set Speed ist nicht Teil des Use Cases.
2. Längeres Fahren mit aktivem L2-System, wobei hier nach Möglichkeit mit einer empfohlenen Geschwindigkeit zwischen 80 und 120 km/h (je nach Verkehrsdichte) konstant auf dem rechten Fahrstreifen gefahren werden sollte, was u. U. ein Folgefahren hinter einem langsameren Fahrzeug bedeutet. Fahrstreifenwechsel sollten nicht durchgeführt werden.
3. Deaktivierung des kompletten Systems, d. h. sowohl der Quer- als auch der Längsführung.
4. Durchführung eines instruierten Überholvorgangs eines langsamer fahrenden Fahrzeugs (bestehend aus einem Fahrstreifenwechsel nach links und eines Zurückscherens nach rechts), nach vorheriger Re-Aktivierung des Systems, wobei der Proband im Vorfeld instruiert wird, die Set Speed des Systems auf eine entsprechend hohe Geschwindigkeit zu stellen, sodass er den Fahrstreifenwechsel zügig ausführen kann; Bewertet werden sollen beide Fahrstreifenwechsel in Kombination.
5. Erleben eines kurzfristigen Ausfalls der Querführung, wenn dieser Fall beobachtet wird, z. B. aufgrund fehlender oder schlechter Fahrstreifenmarkierungen; d. h. von der Reihenfolge flexibel in die anderen Use Cases einzufügen.
6. Erleben einer geplanten Systemgrenze, wobei diese sich aus vorsehbaren Streckenbedingungen ergeben und keine zeitkritische Fahrerreaktion erfordern sollte, wie das Abfahren von der Autobahn an einer Autobahnabfahrt.
7. Erleben einer Hands-Off-Warnung (optional): im Realverkehr höchstens bis zur ersten Stufe der

Hands-Off-Warnung zu instruieren, um keine Sicherheitsgefährdungen zu riskieren.

8. Erleben einer unvorhergesehenen Systemgrenze, wenn sich dieser Fall ereignet; diese kann systembedingt sehr unterschiedlich sein und stark von der individuellen Verkehrssituation abhängen, weshalb diese Fälle als Einzelfälle auftreten werden und als solche entsprechend individuell dokumentiert und interpretiert werden müssen.

Diese Use-Case-Reihenfolge ist in der App hinterlegt. Da im Gegensatz zur Simulatorstudie die Dauer und Reihenfolge der Use Cases aber nicht vordefiniert ist, wechselt der Testleiter selbstständig mithilfe der App-Tasten weiter zum nächsten Use Case. Auch hier kann er mittels der Rückschweifunktion zu einem später platzierten Use Case oder zu einem in der Reihenfolge weiter zurückliegenden Use Case wechseln und diesen bearbeiten. Bereits erfolgreich absolvierte und entsprechend bewertete Use Cases werden durch eine blaue Hinterlegung der VL-Rating-Skala angezeigt. So kann der Versuchsleiter überprüfen, welche Use Cases bereits bearbeitet wurden und welche noch durchzuführen sind. Da während der Fahrt im Realverkehr sichergestellt werden muss, dass der Fahrer durch die Befragung nicht zu stark abgelenkt wird, sollte die Fahrerbefragung generell stark reduziert werden (im Optimalfall eine, maximal zwei relevante Fragen pro Use Case). Die Bewertung der Use Cases kann mehrfach erfolgen. Dazu ist in der App ein Wiederholungssset der Use Cases hinterlegt, um eventuell aufgrund von notwendigen Unterbrechungen nicht vollständig durchgeführte Use Cases wiederholen oder Übungs-Effekte aufdecken zu können.

### 3.4 Prüfzenarien

#### 3.4.1 Definition von Prüfzenarien für Simulatorstudien

Die in Kapitel 1.5 hergeleiteten generischen Use Cases können als Prüfzenarien in einer experimentell erzeugten Prüfanordnung z. B. im Simulator umgesetzt werden. Hierfür gibt es häufig verschiedene Variationsmöglichkeiten, z. B. im Hinblick darauf, wie eine systeminitiierte Transition plausibilisiert wird. Für eine geeignete Auswahl können die Systembeschreibungen aus Bedienungsanleitungen gängiger L2-Seriensysteme verwendet

werden. Für Systemgrenzen können diejenigen Prüfzenarien ausgewählt werden, die in den meisten Systemen Systemgrenzen darstellen (z. B. System kann keine Baustellen durchfahren, schafft keine übermäßig scharfen Kurven, sollte nicht bei schlechtem Wetter oder schlechten Straßenverhältnissen genutzt werden). Wie genau die Transition gestaltet ist, d. h. ob der Fahrer eine Warnung erhält bzw. ob das ganze System vs. nur eine Teilfunktion abgeschaltet wird, ist dabei systemabhängig.

Im Rahmen einer Simulatorstudie werden die Szenarien dann zu einem zusammenhängenden Prüfparcours aneinandergesetzt. Die Reihenfolge der Szenarien (z. B. wann während einer Fahrt eine Systemgrenze auftritt bzw. wie viele Systemgrenzen insgesamt während einer bestimmten Dauer der Prüffahrt erlebt werden) kann sich auf die Interaktion des Fahrers mit dem System auswirken. Um derartige Reihenfolge-Effekte, z. B. Übungseffekte zu kontrollieren, empfiehlt es sich, die Szenariereihenfolge entweder explizit zu variieren oder sich basierend auf einer plausiblen Begründung (z. B. Überprüfung des Erstkontakts anhand der kritischsten Systemgrenze) für eine Reihenfolge zu entscheiden und diese entsprechend in der Interpretation der Ergebnisse zu berücksichtigen.

Für eine Standardisierung der Methodik, die es ermöglicht, Studienergebnisse in anderen Versuchsumgebungen zu replizieren und untereinander vergleichbar zu machen, sollte die Spezifikation der Prüfzenarien dokumentiert werden und eine grafische Visualisierung beinhalten. Es sollten folgende Angaben gemacht werden:

- Straßentyp (inklusive Bebauung, Anzahl Fahrstreifen, Straßenbreite),
- Umgebungsbedingungen (Bepflanzung, Wetterbedingungen, Sichtbedingungen, Tag/Nacht),
- Besonderheiten wie Baustelle, Hindernisse auf der Fahrbahn (inklusive Längen von Abschnitten, Positionen von Objekten etc.),
- Höhen- und Lageprofil (Kurvigkeit der Strecke, Steigungen, Gefälle),
- Geschwindigkeiten (Set Speed des Systems, vorgegebene Höchstgeschwindigkeiten z. B. über Beschilderung),
- Verkehrssituation (Verkehrsdichte auf eigenem Fahrstreifen, anderen Fahrstreifen, Gegenverkehr).

### 3.4.2 Auswahl einer geeigneten Prüfstrecke im Realverkehr

Als geeignete Prüfstrecke im Realverkehr wird empfohlen, eine möglichst wenig befahrene, vierstreifige Autobahnstrecke (mit je zwei Fahrstreifen pro Richtung) zu wählen, um den Einfluss des umgebenden Verkehrs zu reduzieren und ein längeres Fahren in L2 zu ermöglichen ohne selbst ein Verkehrshindernis darzustellen.

Grundsätzlich ist es nicht nachteilig, wenn die gewählte Strecke geschwindigkeitsbeschränkt ist. Im Optimalfall enthält die Strecke einzelne Abschnitte mit besonderen baulichen Gegebenheiten, die im Regelfall den Ausfall der Querführung bedingt, z. B. ein Abschnitt mit schlechten oder fehlenden Fahrstreifenmarkierungen. Um vergleichbare Bedingungen für alle Probanden der Studie zu schaffen, sollten sich die Gegebenheiten der Strecke während des Zeitraums der Datenerhebung nicht ändern. Daher empfiehlt es sich auch, im Vorfeld einer geplanten Studie Erkundigungen vorzunehmen, ob dort in absehbarer Zeit Baumaßnahmen geplant sind, die zu einer Änderung des Streckenverlaufs bzw. der Anforderungen an den Fahrer führen würden. Nicht vorhersehbar bleibt das Auftreten von Tagesbaustellen, die eine Systemgrenze für einen Teil der Stichprobe darstellen und separat erfasst und berichtet werden sollten.

Aus organisatorischen Gründen wird empfohlen, eine Prüfstrecke auszuwählen, die vom Standort des durchführenden Forschungsinstituts der Studie und den Studienteilnehmern leicht zu erreichen ist, sodass dieser als Ausgangs- und Endpunkt der Studie genutzt werden kann. Alternativ kann als Ausgangs- und Endpunkt auch ein geeigneter Parkplatz direkt an der Auf- und Abfahrt von einer Autobahn, z. B. ein Pendlerparkplatz, gewählt werden, um die Zeit bis zum Beginn der Prüfstrecke möglichst gering zu halten. Es sollte sichergestellt sein, dass sich der Versuchsleiter bzw. das VL-Team dort längerfristig aufhalten kann, die Probanden direkt empfangen werden können und Möglichkeiten für einen Toilettengang sowie eine abgeschiedene Möglichkeit für eine Vor- und Nachbefragung geboten ist.

Die Strecke selbst sollte eine ca. halbstündige Fahrt auf der Autobahn in eine Fahrtrichtung ermöglichen (bei einer Durchschnittsgeschwindigkeit von 100 km/h entspricht dies einer Strecke von ca. 50 km). Dann sollte man die Autobahn an einer Abfahrt verlassen können, um in der Gegenrichtung zurück

zum Ausgangspunkt zu gelangen und diese als Use Case der geplanten Systemgrenze erleben zu können. Im Optimalfall können bis zu diesem Zeitpunkt alle sieben geplanten Use Cases (Use Case 8 kann nicht geplant werden) erlebt werden, sodass die Rückfahrt für eine Wiederholung von Use Cases genutzt werden kann. Falls die Fahrdauer aus verkehrstechnischen Gründen verlängert werden muss, um alle Use Cases zu absolvieren, sollte die nächste Ausfahrt bzw. Wendemöglichkeit nicht weit entfernt sein. Zwischen den einzelnen Testfahrten sollte ausreichend Zeit für unvorhergesehene Staus eingeplant werden, wobei die Zeiten der Prüffahrten generell außerhalb der Rush-Hour geplant werden sollten (z. B. erst ab 9 Uhr bis ca. 16 Uhr).

### 3.5 Erprobung der Bewertungsmethode

Zur Erprobung und Validierung der Tablet-App und deren Bewertungskategorien wurden zwei Nutzerstudien im Fahrsimulator mit unterschiedlichen Fragestellungen durchgeführt. Eine Studie wurde im dynamischen Fahrsimulator der WIVW GmbH durchgeführt. Diese sollte evaluieren, inwieweit sich in der Expertenbewertung identifizierte Gestaltungsprobleme der MMS einer teilautomatisierten Fahrfunktion im Verhalten und Erleben von Nutzern widerspiegeln. Die Ergebnisse dieser Studie wurden zur weiteren Optimierung der Bewertungsmethode verwendet und in der Anwendung der App im Realverkehr berücksichtigt. Die zweite Studie wurde im Pkw-Simulator der Bundesanstalt für Straßenwesen durchgeführt und hatte zum Ziel, den Einfluss unterschiedlicher Ausprägungen des Systemverhaltens, speziell der Güte der Querführung, auf das Erleben und Verhalten der Fahrer zu prüfen. Die Ergebnisse dieser Studie können aufzeigen, inwieweit sich Aspekte des Systemverhaltens neben der Gestaltung der MMS auf eine sichere und effiziente Interaktion mit dem System auswirken. Anschließend wurde das Tool im Realverkehr erprobt. Die dafür verwendete App-Version enthielt die in den vorherigen Kapiteln beschriebenen Anpassungen.

#### 3.5.1 Nutzerstudie zum Vergleich verschiedener MMS-Konzepte im WIVW Simulator

Die Nutzerstudie in der WIVW-Simulation hatte zwei Ziele. Zum einen sollte die entwickelte standardisierte Beobachtungs- und Befragungsmetho-

de S.A.D.E. evaluiert werden. Dabei sollte geprüft werden, ob die über die Tablet-Anwendung erfassten Maße des Fahrerverhaltens und der Fahrerbewertung geeignet sind, um die Gestaltung der MMS beim teilautomatisierten Fahren zu bewerten und wie gut das Tool, zunächst für den Anwendungsbereich der Fahrsimulation, anwendbar ist. Erst in einem weiteren Schritt wurde die App derart angepasst, dass sie auch im Realverkehr flexibel und effizient eingesetzt werden kann. Darüber hinaus sollten die Ergebnisse der Studie genutzt werden, um die im Rahmen des Projekts entwickelte Checkliste als Methode zu validieren. Hierfür wurde geprüft, inwieweit sich Unterschiede zwischen zwei MMS-Varianten, die durch eine Expertenbewertung von zwei Ratern bewertet wurden (siehe Kapitel 2.5.2), auch im Verhalten und Erleben von Nutzern aufzeigen lassen.

#### Teilautomatisierte Fahrfunktion und Varianten der Mensch-Maschine-Schnittstelle

Das verwendete System inklusive der beiden MMS-Varianten, die sich in wesentlichen Gestaltungsaspekten unterschieden („gute“ Variante A vs. eine „schlechte“ Variante B), war identisch zu dem unter Kapitel 2.5.2 beschriebenen System. Eine Übersicht über die verschiedenen MMS-Varianten findet sich in Kapitel 2.5.2.

#### Prüfstrecke

Die Prüfstrecke setzte sich aus folgenden Use Cases (hier Szenarien genannt) zusammen, die für die vorliegende Fragestellung als relevant erachtet wurden. Statt Fahrstreifenwechsel wird hier der Begriff „Spurwechsel“ verwendet:

- Systemaktivierung (instruiert),
- Systemdeaktivierung (instruiert),
- Fahren mit aktivem L2-System,
- Spurwechsel hinter langsamem Vorderfahrzeug (instruiert),
- automatische Anpassung der Set Speed durch System,
- Ausfall Querführung (geht in Stand-by) aufgrund fehlender Fahrstreifenmarkierungen,
- Systemausfall an Pannenfahrzeug; stehendes Fahrzeug wird durch Sensorik nicht erkannt, System bleibt aktiv; Fahrer muss eingreifen.



- Systemgrenze bei scharfer Kurve: System kann Situation nicht bewältigen, Querführung schaltet sich ab (geht in Stand-by-Zustand).

Tabelle 3-2 zeigt den Ablauf der Strecke mit den Längen der einzelnen Streckenabschnitte. Die Gesamtlänge der Prüfstrecke betrug ca. 35 km. An manchen Stellen ergeben sich Wiederholungen von Szenarien aus dem Ablauf (konkret: vier Phasen des Fahrens mit aktivem L2-System sowie eine zweite Systemaktivierung am Ende der Fahrt nach instruierter Deaktivierung). Die Fahrt startete auf der Landstraße, anschließend sollte auf eine vierstreifige Autobahn (zwei Fahrstreifen pro Richtung) aufgefahen werden. Die Richtgeschwindigkeit betrug 120 km/h.

Die Fahrer sollten mit Ausnahme des instruierten Fahrstreifenwechsels und des Ausweichens um das Pannenfahrzeug auf dem rechten Fahrstreifen bleiben. Auf dem linken Fahrstreifen herrschte eine mittlere Verkehrsdichte. In den speziellen Fahr- und Fahrstreifenwechselsituationen waren keine weiteren Fahrzeuge auf den Nachbarfahrstreifen, um vergleichbare Bedingungen für alle Probanden herzustellen.

### Versuchsplan

Im Rahmen des Versuchs wurden die zwei unter Kapitel 2.5.2 beschriebenen MMS-Varianten miteinander verglichen. Dieser Faktor wurde als between-subject-Faktor variiert, d. h. eine Gruppe von 12 Probanden erlebte die MMS-Variante A, die andere die MMS-Variante B. Ein weiterer Faktor war die Reihenfolge der beiden kritischen Situationen „Systemausfall bei Pannenfahrzeug“ (Szenario 5) sowie „Systemgrenze bei scharfer Kurve“ (Szenario 10), da zu erwarten war, dass der Erstkontakt mit einer Systemgrenze/-ausfall besonders schwierig werden könnte, was die Entscheidung für diesen Faktor begründet. Die Position der anderen Szena-

		Reihenfolge Situation Hindernis und scharfe Kurve permutiert (between-subject Faktor)	
		1 (erst Hindernis)	2 (erst scharfe Kurve)
MMS-Variante (between-subject Faktor)	A	n = 6	n = 6
	B	n = 6	n = 6

Tab. 3-3: Versuchsplan der WIVW-Simulatorstudie

Nr.	Szenario	Beschreibung	Länge [m]
1	Systemaktivierung (1) nach Autobahnauffahrt	Fahrer wird aufgefordert L2 zu aktivieren	4.000
2	Fahren mit aktivem System auf Autobahn (1)	Fahrer fährt mit L2	2.000
3	Spurwechsel hinter langsamem Fahrzeug	Fahrer muss SW initiieren, während Blinker gesetzt ist, Querführung in Stand-by	4.000
4	Automatische Anpassung der Set Speed bei Speed Limit	System passt Geschwindigkeit automatisch an (in der Auswertung nicht berücksichtigt)	2.500
5	Systemausfall: Hindernis (RF permutiert mit Szenario 10: scharfe Kurve)	Vorderfahrzeug wird nicht erkannt (Sensorfehler, der vom System nicht vorhersehen werden kann)	3.000
6	Systemreaktivierung (1)	Falls Fahrer vorab System deaktiviert, Aufforderung, es wieder zu aktivieren	2.000
7	Fahren mit aktivem System auf Autobahn (2)	Fahrer fährt mit L2	2.000
8	Schlechte Spurmarkierungen	Querführung fällt aus für 500 m	3.000
9	Fahren mit aktivem System auf Autobahn (3)	Fahrer fährt mit L2	2.000
10	Scharfe Kurve – Fahrer muss mitlenken	Systemgrenze: System kann Querführung nicht ausführen, deaktiviert sich.	3.000
11	Fahren mit aktivem System auf Autobahn (4)	Fahrer fährt mit L2	2.000
12	Systemdeaktivierung – instruiert	Fahrer wird aufgefordert L2 zu deaktivieren (L2->ACC->MAN)	2.000
13	Systemreaktivierung (2)	Fahrer wird aufgefordert L2 zu aktivieren (MAN->ACC->L2)	2.000
14	Abfahrt von Autobahn	Abfahrt von der Autobahn	2.000
Gesamtlänge			35.500

Tab. 3-2: Strecke der Prüffahrt mit den einzelnen Szenarien und Streckenlängen

rien wurde nicht verändert. Dadurch erlebten die Fahrer in Reihenfolge 1 nach den unkritischen Situationen „Spurwechsel“ und „Anpassung der Set Speed“ den „Systemausfall beim Pannenfahrzeug“ als erste kritische Situation, die eine zeitnahe Reaktion erforderte, um nach dem Erleben eines eher unkritischen Ausfalls der Querführung im Szenario „schlechte Spurmarkierungen“ zu einem späteren Zeitpunkt den kritischen Ausfall der Querführung im Szenario „Scharfe Kurve“ zu erleben.

In Reihenfolge 2 erlebten die Fahrer als erste kritische Situation den Ausfall der Querführung in der scharfen Kurve und erst anschließend den unkritischen Ausfall im Szenario „schlechte Spurmarkierungen“, gefolgt von der kritischen Situation „Pannenfahrzeug“. Auch dieser Faktor wurde als between-subject-Faktor realisiert, d. h. innerhalb der beiden MMS-Varianten erlebten jeweils  $n = 6$  Fahrer den Prüfparcours in Reihenfolge 1, die anderen  $n = 6$  in Reihenfolge 2. Der Versuchsplan der Studie war somit ein  $2 \times 2$  Versuchsplan mit den zwei unabhängig variierten Faktoren MMS-Variante (2 Stufen) und Szenarien-Reihenfolge (2 Stufen).

### Stichprobe

An dem Versuch nahmen insgesamt  $N = 24$  Probanden (11 weiblich) aus dem Testfahrerpanel des WIVW teil. Zur Aufnahme in das Panel nehmen die Fahrer an einem mind. 2.5-stündigen Trainingsprogramm im Simulator teil. Das mittlere Alter betrug  $M = 41.2$  Jahre ( $SD = 14.3$  Jahre; Altersspanne: 21-64 Jahre). Die teilnehmenden Probanden fahren im Durchschnitt  $M = 21.200$  km ( $SD = 18.200$  km). Die Fahrer hatten im Schnitt eine Erfahrung von  $M = 3.4$  h ( $SD = 4.0$  h) mit hochautomatisierten Systemen (SAE-Level 3) aus anderen Simulatorstudien so-wie  $M = 0.48$  h mit teilautomatisierten Systemen ( $SD = 0.5$  h). Je  $n = 12$  Probanden fuhren die Prüfstrecke mit Systemvariante A, die anderen  $n = 12$  fuhren mit Systemvariante B. Die beiden Gruppen waren bzgl. der Charakteristika Alter, Geschlechterverteilung, Fahrleistung sowie Automatisierungserfahrung vergleichbar, wie Tabelle 3-4 zeigt.

### Versuchsablauf

Nach der Begrüßung und des Ausfüllens der Datenschutzerklärung wurden die Probanden über den Inhalt der Studie aufgeklärt. Anschließend erfolgte eine kurze manuelle Eingewöhnungsfahrt (Dauer: 5 Minuten). Danach wurden die Probanden gebeten, eine kurze Anleitung zur Systembedienung durchzulesen. Diese enthielt eine Beschreibung der grundlegenden Teilfunktionen des Systems, d. h. der Funktion der kontinuierlichen Abstands- und Geschwindigkeitsregelung (kurz: ACC, Adaptive Cruise Control) sowie der Spurhalteassistenten, deren grundsätzliche Funktionsweise, die Systemgrenzen und die Bedienung dieser Systeme. Darin wurde explizit darauf hingewiesen, dass der Fahrer weiterhin auf das Verkehrsgeschehen und das Systemverhalten zu achten hat und die Hände jederzeit am Lenkrad behalten muss, da das Fahrzeug nicht die Überwachung der Fahrzeugumgebung übernimmt. Die Systemanleitung war an die jeweilige Bedienlogik der MMS-Varianten A und B angepasst.

Vor Beginn der Prüffahrt wurden die Probanden auf die visuelle Anzeige im Kombiinstrument hingewiesen, auf der zu erkennen war, ob bzw. welche Funktion des Systems gerade aktiv ist. Zudem wurde ihnen die Position der relevanten Lenkradtasten gezeigt sowie die Antwortskalen zur Online-Befragung während der Fahrt (7-stufige Skala von -3 bis +3 für die Fragen 1-5 auf der Tablet-App sowie die 11-stufige Skala zur Bewertung der Situationskritikalität) und deren Anwendung erklärt. Die Fahrer wurden gebeten, möglichst spontan und zügig auf die Fragen zu antworten und weitere Anmerkungen und Kommentare in der Befragung nach der Fahrt zu äußern. Weiterhin erhielten sie generelle Instruktionen zum Ablauf der Prüffahrt. Sie wurden gebeten, zunächst selbst auf die Autobahn aufzufahren und erst auf konkrete Instruktion des Versuchsleiters hin das System zu aktivieren. Während der Fahrt auf der Autobahn sollten sie nach Möglichkeit immer mit aktivem System fahren (d. h. mit aktivem ACC und aktivem Spurhalteassistenten). Sie sollten immer auf dem rechten Fahrstreifen bleiben, 120 km/h

	Anteil weiblich	Alter (Jahre)	Fahrleistung (Jahr)	L3-Erfahrung (min)	L2-Erfahrung (min)
MMS-Variante A	5	M = 41.6 SD = 14.5	M = 21.200 SD = 14.100	M = 191 SD = 243	M = 28.8 SD = 34
MMS-Variante B	6	M = 41.0 SD = 14.7	M = 18.200 SD = 16.700	M = 222 SD = 246	M = 30.0 SD = 50.0

Tab. 3-4: Stichprobencharakteristika der beiden MMS-Gruppen



fahren und sich an die Straßenverkehrsordnung halten. Die Probanden wurden nochmals daran erinnert, dass sie die Verkehrssituation jederzeit überwachen und ihre Hände am Lenkrad behalten müssen. Anschließend wurde die Simulation gestartet und der Fahrer absolvierte den Prüfparcours, in dem er alle relevanten Use Cases erlebte. Während der Fahrt fand unmittelbar nach den Use Cases die Befragung anhand der vorab pro Szenario definierten Fragen statt. Zur Erleichterung der Beantwortung wurde die Skala auf das Lenkrad geklebt. Der Versuchsleiter protokollierte während der Fahrt mittels der S.A.D.E.-Anwendung aufgetretene Bedien-, Fahr- und Überwachungsfehler. Für den Fall, dass er sich bei einer Bewertung unsicher war, etwas nicht eindeutig beobachten konnte, einen Teil der Fragen aus Zeitgründen nicht stellen konnte oder fehlerhafte Eingaben korrigieren wollte, hatte er einen ausgedruckten Protokollbogen, auf dem er derartige Sonderfälle pro Szenario händisch protokollieren konnte.

Nach Absolvieren aller Use Cases pausierte der Versuchsleiter die Simulation, um alle eingegebenen Formulare pro Szenario mittels der Vor- und Zurück-Blätter-Funktion in der App auf Vollständigkeit zu überprüfen, fehlende Eingaben zu ergänzen und fehlerhafte Eingaben zu korrigieren. Abschließend erfolgte eine Nachbefragung, in der die Probanden das System bzw. die MMS anhand vorgegebener Aussagen beurteilten sowie ein Gesamturteil über das System abgaben. Danach erhielten die Probanden eine Aufwandsentschädigung von 20 Euro ausbezahlt und wurden verabschiedet. Insgesamt dauerte der Versuch ca. 1 h.

### Abhängige Variablen und statistische Analysen

Als abhängige Variable wurden folgende Parameter ausgewertet:

- VL-Rating zur Bewältigung des Szenarios (Skala von 0-10) pro Szenario
- Auftreten einzelner Bedien-, Überwachungsprobleme und Fahrfehler pro Szenario (0/1)
- Analyse der Zeitverlaufdaten aus SILAB für die Szenarien „Hindernis“ und „Scharfe Kurve“
  - Szenario „Hindernis“: Minimale Time-To-Collision (TTC) vor Pannenfahrzeug
  - Szenario „Scharfe Kurve“: Maximale laterale Abweichung in Kurve

- Subjektive Bewertung der MMS mittels spezifisch pro Szenario definierter Online-Befragung (-3...+3)
  - Frage 1: „Wie klar war Ihnen, was zu tun ist?“ (im Szenario Aktivierung, Deaktivierung, Hindernis, scharfe Kurve)
  - Frage 2: „Wie nachvollziehbar war das Systemverhalten?“ (im Szenario Spurwechsel, Hindernis, scharfe Kurve, Ausfall Querführung)
  - Frage 3: „Wie verständlich waren die Systemausgaben?“ (Im Szenario Spurwechsel, Hindernis, scharfe Kurve, Ausfall Querführung)
  - Frage 4a: „Wie schwierig war die Systembedienung?“ (im Szenario Aktivierung, Deaktivierung) – diese Frage wurde später aus dem Variablenkatalog entfernt.
  - Frage 4b: „Wie schwierig war die Reaktion auf das Systemverhalten?“ (im Szenario Hindernis, scharfe Kurve) – diese Frage wurde später aus dem Variablenkatalog entfernt.
  - Frage 5 (optionale Frage): „Wie gut waren die Systemzustände voneinander unterscheidbar; aufgrund der bewusst unterschiedlichen Gestaltung der MMS-Anzeigen im Bezug auf die Kontrastgüte und die Symbolik für die aktive Spurhalteassistenten?“ (im Szenario Aktivierung, Deaktivierung, Ausfall Querführung)
- Subjektive Bewertung der Situationskritikalität
  - Frage 6: „Wie kritisch bewerten Sie die Situation?“ (im Szenario Hindernis und scharfe Kurve)

- Nachbefragung: Bewertung der MMS anhand verschiedener Dimensionen aus der Checkliste (-2...+2 + Kategorie „kann ich nicht beurteilen“)

Die Online-Verhaltensbeobachtung und -befragung erfolgte mittels der in Kapitel 3.3 beschriebenen App S.A.D.E. Der Versuchsleiter war vorab anhand der Schulungsmaterialien in der Anwendung der Benutzeroberfläche und der Vergabe der Fehler und des Ratings geschult worden. Zudem wurden in Vorversuchen spezifisch für die in der Simulatorstudie zu erwartenden Bedien- und Fahrprobleme

Regeln zur Vergabe der Fehlerkategorien aufgestellt (z. B. was in einem speziellen Fall als Bedienfehler zu werten ist). Ebenso wurde vereinbart, dass Fahrfehler, die aufgrund der Versuchsanordnung auftraten, z. B. weil die Probanden während einer manuellen Fahrphase ihre Aufmerksamkeit für die Beantwortung der Fragen sehr lange auf die auf dem Lenkrad aufgeklebte Skala richteten, nicht gezählt werden sollten. Lediglich Fahrfehler, die in Verbindung mit der Reaktion auf das Systemverhalten oder Bedienaktionen auftraten, sollten in die Bewertung eingehen. Der Rater hatte in mehreren Vorversuchen die Gelegenheit, geeignete Prozesse für Beobachtung vs. Befragung einzustudieren (z. B. was soll zuerst anhand welcher Monitore beobachtet werden, wann muss er spätestens mit der Befragung beginnen, um diese innerhalb des Szenarios rechtzeitig abzuschließen etc.).

### Inhaltliche Ergebnisse

Die Datenauswertung erfolgte mittels des IBM-Statistik-Programms SPSS in der Version 24. Für die Analyse der subjektiven Daten wurden univariate ANOVAs mit zwei unabhängigen Faktoren gerechnet (Faktor 1: MMS-Variante; Faktor 2: Reihenfolge der kritischen Situationen „Hindernis“ und „Scharfe Kurve“). Die beschriebenen Kennwerte der Fahrdatenanalyse wurden ebenfalls mittels ANOVAs ausgewertet. Für die Analyse der Häufigkeitsverteilungen der beobachteten Bedienprobleme und Fahrfehler pro Szenario wurden Chi<sup>2</sup>-Tests für die Effekte der MMS-Variante gerechnet.

Die erwarteten Unterschiede zwischen den MMS-Varianten, die so gestaltet waren, dass sie sich in wesentlichen Gestaltungsprinzipien unterschieden, zeigten sich im Erleben und Verhalten von Nutzern auf verschiedenen Ebenen: Das Fehlen einer visuellen und/oder akustischen Warnung bewirkte eine schlechtere Reaktion in der Spurhaltung sowie ein schlechteres Systemverständnis: Dies zeigte sich im Szenario „Scharfe Kurve“, in dem sich die beiden MMS-Varianten dadurch voneinander unterschieden, dass in Variante A eine Sekunde vor dem Ausfall der Querführung eine visuelle (Lenkrad wird rot) und eine akustische Warnung (wiederholter Warnton) ausgegeben wurde, während das System in Variante B „still“, d. h. ohne explizite Warnung ausfiel. In beiden Gruppen trat bei einer sehr hohen Anzahl von Probanden zumindest eine schlechte Spurhaltung auf. In der MMS-Variante A erlitten tendenziell weniger Fahrer Spurverlassens-Ereignisse

( $p = .132$ ). In der nachfolgenden Online-Befragung gaben die Fahrer der MMS-Variante A übereinstimmend an, die Systemausgaben signifikant besser verstanden zu haben ( $p < .000$ ). Zudem war das Systemverhalten durch die zusätzliche Warnung besser nachvollziehbar ( $p = .003$ ). Im Einklang mit den Ergebnissen der dichotomen Beobachtungskategorien war die maximale laterale Abweichung in der Kurve in der MMS-Variante B tendenziell höher als in Variante A ( $p = .058$ ). Die Situation wurde zudem von der Fahrergruppe mit MMS-Variante B kritischer wahrgenommen (Var A:  $M = 3.75$ ;  $SD = 1.71$ ; Var B:  $M = 5.17$ ;  $SD = 2.1$ ). Dieser Unterschied wurde allerdings statistisch nicht signifikant ( $p = .119$ ).

Die Unterschiede in der Bedienlogik (Anzahl nötiger Bedienschritte, Positionierung und Benennung der Bedienelemente sowie Unmittelbarkeit des Systemfeedbacks bei Bedienvorgängen) wirkten sich auf die Häufigkeit von Bedienproblemen dergestalt aus, dass die Bedienhandlungen unsicher oder verzögert bzw. nicht vollständig ausgeführt wurden, so dass zur erfolgreichen Ausführung zum Teil Hilfestellung nötig war: Dies zeigte sich vor allem bei der Analyse der Fehlerhäufigkeiten im Szenario „Aktivierung 1“ (erste Systemaktivierung nach Auffahrt auf die Autobahn) daran, dass die Fahrer beim Aktivierungsprozess der MMS-Variante B häufiger Unsicherheiten zeigten und erst verzögert aktivierten, häufiger nicht erfolgreiche Bedienhandlungen ausführten und häufiger Hilfestellung benötigten, um das System komplett zu aktivieren. Signifikant werden die Unterschiede zwischen den MMS-Varianten für die Kategorie „nicht erfolgreich“ ( $p = .049$ ) und „Hilfestellung nötig“ ( $p = .019$ ). In beiden MMS-Varianten resultierten während des Aktivierungsprozesses Probleme mit der Spurhaltung, die jedoch vergleichbar häufig auftraten. Zudem schien den Probanden in der MMS-Variante A klarer zu sein, was sie tun müssen. Die Aktivierung wurde als leichter erlebt und die Systemzustände schienen besser unterscheidbar zu sein. Statistisch ergibt sich aufgrund der großen Streuungen zwischen den Probanden nur für die Variable der subjektiv wahrgenommenen Schwierigkeit eine Tendenz zugunsten der Variante A ( $p = .061$ ). Bei der zweiten Aktivierung des Systems zeigten sich insgesamt weniger Probleme als bei der ersten Aktivierung. Dennoch blieben die Fahrer der MMS-Variante B tendenziell unsicherer bei der Aktivierung und machten mehr Bedienfehler. Diese größere Problematik in der Bedienung wirkte sich in diesem Fall in einer häufigeren Zahl von Spurverlassens-Ereignissen in der

MMS-Variante B aus. Subjektiv gesehen, gibt es keine signifikanten Unterschiede zwischen den beiden MMS-Varianten trotz augenscheinlich größerer wahrgenommener Bedienschwierigkeit und schlechterer Unterscheidbarkeit der Systemzustände in MMS-Variante B. Im Szenario „Deaktivierung“ hatte der überwiegende Teil der Probanden keine Bedienprobleme. Wenn Probleme auftraten, dann in Einzelfällen in der MMS-Variante B, was zu tendenziell signifikanten Unterschieden in den Bewertungskategorien „unsicher“ ( $p = .140$ ), „nicht erfolgreich“ ( $p = .140$ ) und „Hilfestellung notwendig“ ( $p = .140$ ) führte. Tendenziell traten dort auch vermehrt Spurverlassens-Ereignisse auf ( $p = .132$ ).

Die Unterscheidbarkeit der Systemzustände, die durch den visuellen Kontrast und die gewählte Symbolik für die aktive Spurhalteassistenten definiert waren, wurden von den Fahrern der beiden MMS-Varianten entsprechend unterschiedlich gut bewertet, allerdings ohne Auswirkungen auf das Fahrverhalten. Dies zeigte sich zumindest in unkritischen Situationen wie in dem hier realisierten Prüfzenario „Ausfall Querführung“: Insgesamt resultierten bei gut einem Drittel der Fahrer dort Probleme in der Spurhaltung, trotz der Tatsache, dass diese hier aufgrund der fehlenden Spurmarkierungen weniger streng bewertet wurden. Die beobachteten Fahrfehler waren allerdings in beiden MMS-Gruppen gleich verteilt, die unterschiedlich gute Bemerkbarkeit des Ausfalls der Querführung in Variante B im Vergleich zu Variante A wirkte sich also nicht auf das Fahrverhalten aus. Subjektiv merkten die Fahrer der MMS-Variante B jedoch tendenziell an, dass die Systemzustände schlechter voneinander unterscheidbar seien. Statistisch gesehen vergleichbar wurde die Nachvollziehbarkeit des Systemverhaltens und die Verständlichkeit der Systemausgaben bewertet. Letztere wurde in beiden MMS-Varianten zwar nicht negativ, aber eher neutral bewertet.

Das globale VL-Rating spiegelt die spezifischen Probleme auf den untergeordneten Kategorien weitgehend wieder, weshalb es gut geeignet erscheint, um einen schnellen und effizienten Gesamteindruck über ein System zu bekommen. In folgenden Szenarien gab es signifikante bzw. tendenziell signifikante Unterschiede zwischen den MMS-Varianten:

- Aktivierung 1 ( $p = .111$ ),
- Aktivierung 2 ( $p = .024$ ),
- Hindernis ( $p = .041$ ),
- Deaktivierung ( $p = .064$ ).

Während sich in den Aktivierungs- und Deaktivierungsszenarien die Unterschiede jeweils in die erwartete Richtung realisierten, dass die Probanden, die mit der MMS-Variante A gefahren sind, die Szenarien besser bewältigten, zeigte sich für das Szenario „Hindernis“ der Unterschied in die entgegengesetzte Richtung. In den Szenarien „Fahren mit aktivem System“, „Lane Change“, „Ausfall Querführung“ und „Scharfe Kurve“ ergaben sich keine globalen Unterschiede in Abhängigkeit der MMS-Variante.

Die identifizierten Unterschiede im beobachteten Fahrverhalten auf dichotomen Kategorien decken sich mit der Analyse von Zeitverlaufs-Fahrdaten. Dies zeigte sich im Szenario „scharfe Kurve“ daran, dass passend zu den Ergebnissen der dichotomen Beobachtungskategorien, die maximale laterale Abweichung in der Kurve in der MMS-Variante B tendenziell höher als in Variante A ( $p = .058$ ) war. Die Situation wurde zudem von der Fahrergruppe mit MMS-Variante B kritischer wahrgenommen (Var A:  $M = 3.75$ ;  $SD = 1.71$ ; Var B:  $M = 5.17$ ;  $SD = 2.1$ ). Dieser Unterschied wurde allerdings statistisch nicht signifikant ( $p = .119$ ). Im Szenario „Hindernis“ bestätigten sich die durch die dichotomen Beobachtungskategorien ermittelten Unterschiede zwischen den MMS-Varianten auch, wenn man die Zeitverlaufs-Fahrdaten aus SILAB analysiert. Die minimale Time-to-Collision fiel in MMS-Variante A signifikant niedriger aus als in MMS-Variante B, d. h. in MMS-Variante A resultierten engere kritische Abstände zum stehenden Hindernis ( $p = .058$ ). Auf Basis dieser Ergebnisse lässt sich schlussfolgern,

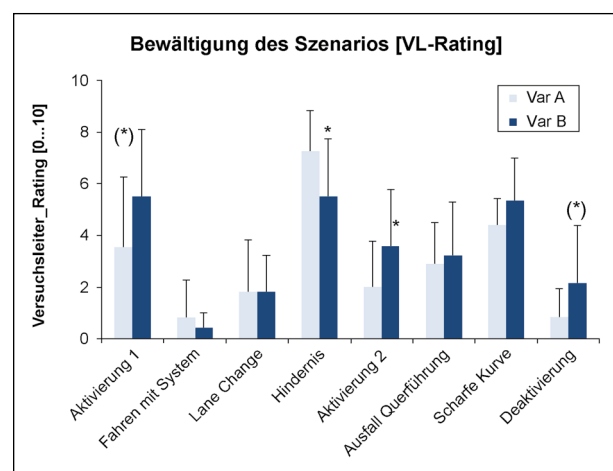


Bild 3-6: VL-Rating zur Bewältigung des Szenarios pro Szenario im Vergleich der MMS-Varianten. Dargestellt sind Mittelwerte und Standardabweichungen. (\*) markiert tendenziell signifikante Unterschiede ( $p < .10$ ), \* markiert signifikante Unterschiede ( $p < .05$ ).

dass die häufig sehr zeit- und ressourcenaufwändigen Analysen von Zeitverlaufsdaten nicht mehr zwingend gebraucht werden. Dadurch wird das Bewertungstool gerade für die Beobachtung im Realverkehr mit nicht-instrumentierten Fahrzeugen zu einem sehr nützlichen und effizienten Instrument.

Keine MMS-Effekte zeigten sich im Szenario „Spurwechsel“, u. U. weil die Situation generell von allen Fahrern gut gelöst wurde. Eine zusätzliche Textanzeige über den Systemstatus während des Spurwechsels hatte hier kaum Auswirkungen. Nach Eindruck des Versuchsleiters wurde die Anzeige eventuell von vielen Probanden gar nicht wahrgenommen, weil deren Aufmerksamkeit auf die Absicherung und Ausführung des Spurwechsels konzentriert war. Die geringe Häufigkeit von Hands-Off-Warnungen in den Abschnitten mit aktivem System lässt sich mit der expliziten Instruktion, die Hände immer am Lenkrad zu lassen, erklären.

Unerwartet waren die Unterschiede im Szenario „Hindernis.“ Hier zeigte sich trotz identischem Systemverhalten und MMS-Ausgaben in den beiden System-Varianten eine schlechtere Fahrerreaktion der Fahrer mit der MMS-Variante A: Der Versuchsleiter bewertete alle Fälle in der MMS-Variante A sowie die Hälfte aller Fälle in Variante B als verzögerte Bremsreaktion, was dadurch definiert war, dass der Fahrer nicht unmittelbar auf das Warndreieck, das 170 m vor dem Pannenfahrzeug positioniert war, mit einer Bremsung reagierte. Daraus resultierte in zwei Dritteln der Fälle in Variante A eine Gefährdung (Abstand nach vorne bzw. zur Seite weniger als 1 m), in Variante B in einem Drittel der Fälle. Nur ein Fahrer der Gruppe A produzierte eine Kollision mit dem Pannenfahrzeug. Dieser Befund ist insofern auffällig, als dass objektiv gesehen in dem Szenario keinerlei Unterschiede im Systemverhalten und in den MMS-Anzeigen vorhanden waren. In beiden Gruppen wurde das Vorderfahrzeug nicht detektiert, entsprechend nicht angezeigt und das Fahrzeug fuhr mit aktiver Längs- und Querführungsassistenz weiter. Dieser Effekt lässt sich nicht auf einer grundlegend anderen Stichprobenzusammensetzung der beiden MMS-Gruppen zurückführen. Wie Tabelle 3-4 zeigt, waren die beiden MMS-Gruppen bzgl. Alter, Geschlechterverteilung, Fahrerfahrung, sowie Erfahrung im Umgang mit teil- und hochautomatisierten Systemen vergleichbar. Auch eine post-hoc Analyse dieser Stichprobencharakteristika aufgeteilt in die Gruppe der Fahrer, die in der Situation eine Gefährdung verursacht hatten, vs. die Gruppe an Fahrern, die keine Gefährdung

verursacht hatten, ergab kein plausibles Erklärungsmuster. Potenzielle Erklärungsmodelle werden in der Diskussion angesprochen. Subjektiv gesehen wurden die MMS-Varianten vergleichbar bewertet. Auch die Situationskritikalität wurde mit einem Mittelwert von  $M = 5.33$  ( $SD = 2.19$ ) für Variante A vs.  $M = 5.08$  ( $SD = 1.88$ ) für Variante B vergleichbar hoch in die Kategorie „unangenehm“ eingestuft.

An dieser Stelle können daher nur Überlegungen über mögliche alternative Erklärungsansätze angestellt werden. So könnte das Erleben der „guten“ MMS-Variante zu einem stärkeren Sicherheitsgefühl und zu einem erhöhten Vertrauen in das Gesamtsystem geführt haben, sodass die Fahrer dem System unbewusst eine höhere Leistungsfähigkeit zuwiesen und entsprechend länger abwarteten, bevor sie von dem Systemausfall beim Pannenfahrzeug überrascht wurden und erst stark verzögert selbst eingriffen. Allerdings kann diese Hypothese aus den Analysen der bestehenden Daten nicht eindeutig belegt werden. Zudem fand keine weitere Fahrerbefragung statt.

Ebenfalls deutlich wurde in der Studie, dass bei der Prüfung von Systemen durch Nutzer, die das System im Erstkontakt weitgehend intuitiv und ohne explizite Übung kennenlernen, Reihenfolge-Effekte wirksam werden, die bei der Analyse der Ergebnisse beachtet werden müssen. So zeigten sich in der vorliegenden Studie Lerneffekte bzgl. der Hands-Off-Warnungen. Im ersten der vier Abschnitte des Fahrens mit aktivem System erlebten insgesamt drei Fahrer, in den drei weiteren Abschnitten nur noch jeweils ein Fahrer eine Hands-Off-Warnung. Diese insgesamt sehr geringe Häufigkeit von Hands-Off-Warnungen ist mit der eindeutigen Instruktion, die Hände am Lenkrad zu lassen, zu erklären. Die Fahrer, die eine Hands-Off-Warnung erhielten (in den anderen Szenarien zeigen sich auch vereinzelt Fälle mit Hands-Off-Warnungen), hatten die Hände zu locker am Lenkrad für eine Sensordektion und korrigierten dies zeitnah.

Außerdem zeigte sich, dass das vorherige Erleben einer kritischen Situation die Probleme in einer späteren unkritischen Situation verringert. Umgekehrt wirkte sich das Erst-Erleben einer unkritischen Situation dann problematischer auf eine nachfolgende kritische Situation aus. Dies kann dadurch erklärt werden, dass die Probanden bestimmte Erwartungen über die Konsequenzen eines Systemausfalls aufbauen und dann von schwerwiegenderen Kon-

sequenzen überrascht werden, wenn die Situation an sich kritischer ist. Diese Interpretation ergibt sich aus den signifikanten Reihenfolge-Effekten für das Szenario „Ausfall Querführung“ ( $p = .047$ ) sowie die „Scharfe Kurve“ ( $p = .08$ ). Dabei ergab sich folgendes Muster: Wurde die kritische Situation „Scharfe Kurve“ als erstes erlebt, reagierte man bei der nachfolgenden unkritischen Situation „Ausfall Querführung“ besser (besseres VL-Rating in RF2). Wurde die kritische Situation „Scharfe Kurve“ nach einer unkritischen Situation erlebt, reagierte man dort schlechter (schlechteres VL-Rating in RF2). Dies ist ein Indikator, dass vorherige Erfahrungen mit einem System in Abhängigkeit der Art der Situation das Erleben eines nachfolgenden Ereignisses in unterschiedlicher Weise beeinflussen können.

Insgesamt ist festzustellen, dass in der vorliegenden Studie häufig nur tendenziell statistische Effekte aufgezeigt werden können. Dies liegt an der hohen Varianz zwischen den Probanden, sowohl auf der Verhaltens- als auch auf subjektiver Ebene, die sich durch die geringe Stichprobengröße stark auf die Ergebnisse auswirkt.

### Methodische Ergebnisse

Neben den beschriebenen inhaltlichen Ergebnissen können aus der durchgeführten Studie methodische Erkenntnisse gewonnen werden und Optimierungsbedarf für das Bewertungsinstrument abgeleitet werden. Generell waren bei der Fehlerkodierung und der Abgabe des VL-Ratings nur wenige Korrekturen notwendig. Dies macht das Verfahren (nachdem die Versuchsleiter zuvor eine intensive Schulung erhalten haben) zu einem effizienten Bewertungsinstrument. Zudem war der Versuchsleiter durch die klaren Vorgaben und Vergaberegeln in der Schulung sehr gut in der Lage, die einzelnen Fehlerkategorien sinnvoll zu einem Gesamtrating zusammenzufassen. Dies macht das unidimensionale Rating zu einem effektiven Maß zur Bewertung eines Systems. Die Fehlerzählung als vorausgehender Schritt und klare Vergaberegeln sind allerdings notwendig, um eine hohe Reliabilität und Objektivität der Methode zu gewährleisten. Einige Optimierungsvorschläge ergaben sich aus Anmerkungen des Versuchsleiters während der Beobachtung der Fehlerkategorien. Nachdem es dem Versuchsleiter eher schwer fiel, die Kategorien „verzögert“ vs. „unsicher“ aus dem Bereich „Bedienprobleme“ zu differenzieren, und es nicht zwingend zur Bewertung des Systems notwendig scheint, diese Fehler-

typen voneinander zu trennen, wird empfohlen diese beiden Kategorien zusammenzufassen. Dasselbe gilt für die Kategorien „Fehlbedienung“ und „nicht erfolgreich“. Diese Optimierungen wurden in den Anpassungen der App für den Realverkehr umgesetzt (siehe Kapitel 3.5.3).

Bezüglich der subjektiven Befragung wurde deutlich, dass manche Befragungsvariablen von den Probanden sehr unterschiedlich verstanden wurden und dadurch zum Teil eine hohe interindividuelle Varianz resultierte. Dies lag sicherlich zum einen an den Effekten, die für subjektive Maße generell typisch sind, z. B. Testpersoneneffekte, wie das Bestreben „ein guter Proband“ sein zu wollen, die dazu führen, dass die Probanden aufgetretene Probleme eher weniger negativ einschätzen (d. h. „die MMS war zwar schlecht, aber ich habe die Situation trotzdem gut gelöst“). Generell ist nur schwer zu trennen, ob Probanden eher eine objektive Bewertung der MMS-Ausgabe vornehmen oder die Interaktion aus ihrer eigenen Reaktion auf das System und die MMS bewerten. Zum anderen können die Probanden durch das between-subjects Design, in dem jeder Proband nur eine MMS-Variante erlebt, nur das System bewerten, das sie erlebt haben. Durch den fehlenden Vergleich wissen die Probanden nicht, dass es auch besser oder schlechter hätte sein können und bewerten das eigene System unter Umständen eher wohlwollend.

Des Weiteren führte die Frage „Wie verständlich waren die Systemanzeigen?“ zu einer uneinheitlichen Beantwortung, wenn Systemanzeigen von den Fahrern gar nicht bemerkt wurden (manche vergeben Rating „0“, manche bewerten aufgrund einer Vermutung, wie es in anderen Fällen war) oder diese nicht vorhanden waren, wie in der Situation „Hindernis“. Für den ersten Fall wäre die Einführung einer zusätzlichen Kategorie „kann ich nicht beurteilen/habe ich nicht bemerkt“ empfehlenswert. Im zweiten Fall sollte auf die Frage verzichtet werden. Ein weiteres Problem entstand durch die – im Vergleich zu den übrigen Fragen – umgekehrte Polung der Antwortskala bei der Beantwortung der Frage nach der Schwierigkeit einer Bedienung bzw. einer Reaktion auf das Systemverhalten. Eine positive Bewertung, d. h. „wenig schwierig“ war hier durch negative Zahlen repräsentiert, während für die anderen Fragen (z. B. „Wie verständlich war etwas?“) eine positive Bewertung durch positive Zahlen auf der Skala angegeben werden sollte. Dies führte bei manchen Probanden zu einer anfänglichen Verwirrung, die durch Nachfragen gelöst wer-

den konnte, bzw. zu langen Blicken auf die Lenkradskalen. Anderen Tendenzen subjektiver Methoden, wie die Beantwortung nach sozialer Erwünschtheit (man möchte als möglichst wissend wahrgenommen werden und gibt daher ungern zu, etwas nicht richtig verstanden zu haben), kann teilweise mit einer geeigneten Instruktion entgegengewirkt werden. Dabei sollte betont werden, wie wichtig es für die Untersuchung ist, dass der Nutzer das System ehrlich bewertet, da das Ziel der Studie ist, Verbesserungspotenzial zu generieren. Insgesamt wird für die Befragungsvariablen empfohlen, diese nur auf wenige Fragen und Use Cases zu konzentrieren und in Relation zu den Verhaltenskategorien nachrangig zu gewichten.

### 3.5.2 Nutzerstudie zur Auswirkung der Spurhaltegüte im BAST-Simulator

#### Fragestellung

Neben Bedien- und Gestaltungsaspekten der Mensch-Maschine-Schnittstelle wirkt sich auch das eigentliche Systemverhalten auf die Sicherheit von teilautomatisierten Systemen aus (VICTOR, TIVESTEN, GUSTAVSSON, JOHANSSON, SANGBERG & LJUNG AUST, 2018). Ein wesentlicher Aspekt ist dabei die Gestaltung der Querführung. An dieser Stelle wird die Hypothese aufgestellt, dass objektiv gesehen „schlechte“ Systeme, d. h. Systeme, bei denen der Fahrer stärker in die Lenkung eingebunden ist, weil die Spurhaltegüte entweder relativ schlecht ist oder das System häufiger ausfällt, eventuell den Vorteil haben, dass sie den Fahrer stärker im Loop halten. Da sie ihn dazu veranlassen, das System und die Verkehrssituation stärker zu überwachen, könnte dies im Notfall z. B. bei einem Systemausfall, zu einer schnelleren bzw. besseren Reaktion beitragen. Im Gegensatz dazu könnten „sehr gute“ Systeme, die über eine nahezu perfekte, zentrierte Spurführung und eine geringe Ausfallquote verfügen, ein übermäßiges Systemvertrauen und damit ein reduziertes Überwachungsverhalten zur Folge haben, was sich im Fall eines Systemausfalls negativ auswirken könnte. Die Überprüfung dieser Hypothesen erfolgte in der hier beschriebenen Nutzerstudie im Fahrsimulator der Bundesanstalt für Straßenwesen (BAST), in der gezielt Systemcharakteristika der Querführung variiert und daraufhin das Fahrerverhalten an Systemgrenzen untersucht wurde.

#### Versuchsumgebung

Die Nutzerstudie fand im statischen Pkw-Simulator der BAST statt, der mit dem Softwarepaket SILAB® der WIVW GmbH betrieben wird. Zusätzlich wurde eine zweite Kamera zur Beobachtung von Bedieneingaben des Fahrers am Lenkrad und eine Videoaufzeichnung implementiert. Das standardisierte Bewertungsinstrument S.A.D.E. wurde auf einem Windows-Tablet implementiert und in den Rechnerverbund eingebunden, um die Online-Verhaltensbeobachtung und -befragung durchzuführen.

#### Teilautomatisierte Fahrfunktion

Als teilautomatisiertes System wurde das in Kapitel 3.5.1 beschriebene System für die Studie verwendet. Es kombiniert eine kontinuierliche Abstands- und Geschwindigkeitsregelung mit einer kontinuierlichen Querführungsassistenz. Das System wird über Tasten am Lenkrad bedient. Die Teilfunktionen können getrennt voneinander aktiviert und deaktiviert werden. Die Geschwindigkeit kann in 10 km/h-Schritten angepasst werden. Das System passt zudem seine Geschwindigkeit automatisch an vorgegebene Geschwindigkeitslimits an. Der Abstand zu einem vorausfahrenden Fahrzeug kann prinzipiell verstellt werden, war aber in der vorliegenden Studie standardmäßig auf der mittleren Stufe eingestellt und sollte vom Fahrer nicht verändert werden.

#### Versuchsplan

Für die Prüfung der Fragestellung wurde die Güte der Querführung auf zwei Dimensionen variiert:

- Faktor 1: Variabilität der Spurmittenzentrierung: zentriert vs. variabel.
- Faktor 2: Reliabilität der Querführung: keine vs. regelmäßige Ausfälle.

Diese Faktoren wurden miteinander gekreuzt und als between-subject-Faktoren untersucht, d. h. jeder Fahrer erlebte nur eine von vier möglichen Kombinationen aus Faktor 1 und 2 und somit eine der folgenden vier Systemvarianten:

- System bewältigt alle Situationen, perfekte Spurmittenzentrierung
- System bewältigt alle Situationen, variable Spurführung
- häufige Ausfälle in der Spurführung, aber perfekte Spurmittenzentrierung

- häufige Ausfälle in der Spurführung, variable Spurführung

Für die Umsetzung der Prüfbedingungen wurden streckengebundene Trigger in die Strecke implementiert, die bewirken, dass das System in den Bedingungen „variable Spurführung“ regelmäßige Spurversätze von der Fahrstreifenmitte aus nach links und rechts ausführt, die gerade so stark sind, dass der Fahrer sie zwar als unangenehm wahrnimmt und das Gefühl bekommt, mitlenken zu müssen, aber noch nicht dermaßen die Lenkung übersteuern muss, um ein Verlassen des Fahrstreifens zu verhindern. Um als weiteren Faktor Systemausfälle zu erzeugen, wurden zusätzlich in regelmäßigen Abständen streckengebundene Trigger gesetzt, die einen Ausfall der Querführung für ca. 3 – 4 sec bewirken. Lenkt der Fahrer hier nicht aktiv mit, kommt er vom Fahrstreifen ab. Als Häufigkeit dieser Systemausfälle wurde eine Frequenz von ca. 1.000 m, d. h. alle 30 s bei 120 km/h definiert. Dies führte bei der definierten Streckenlänge zu insgesamt 22 Systemausfällen bis zum Erreichen der jeweiligen Prüfsituation. Bezüglich der Längsführung verhielten sich alle Systeme gleich.

### Mensch-Maschine-Schnittstelle

Basierend auf Benchmark-Analysen wurde ein prototypisches Anzeige- und Bedienkonzept entwickelt, das die verschiedenen Systemzustände über entsprechend voneinander unterscheidbare Symboliken und Farbigkeit abbilden kann. Diese sind in ein Grafical User Interface (GUI) integriert, das im Kombi-Display des Simulator-Mock-Ups angezeigt wird. Es ist modular aufgebaut (z. B. Anzeigebereich für Systemzustände, Bereich für Geschwindigkeitsanzeigen, Box für Textmeldungen, farbig

gestalteter und animierbarer Rahmen zur Unterstützung der Unterscheidbarkeit der Systemzustände).

Bild 3-7 zeigt beispielhaft die visuelle Anzeige im Systemzustand „L2 aktiv“. Sie besteht zum einen aus klassischen Anzeige-Elementen wie einem analogen Drehzahl- und Geschwindigkeitsmesser. Hinzu kommen Anzeigen, die die verschiedenen Assistenzfunktionen anzeigen. Die horizontalen Balken symbolisieren das ACC. Im Stand-by-Zustand werden die Balken grau angezeigt, im Aktiv-Zustand blau. Ein erkanntes Vorderfahrzeug wird als Fahrzeug innerhalb des Fahrstreifens angezeigt. Die aktive Unterstützung in der Querführung wird über das Lenkradsymbol in blau angezeigt. Die vertikalen Linien entlang der symbolisierten Spur wechseln zwischen Grau und Blau, je nachdem, ob das System aktuell regelt oder nicht. Werden keine Fahrstreifen erkannt, sind keine Fahrstreifenbegrenzungen dargestellt. Im oberen Bereich der Anzeige findet sich in digitaler Form links die eingestellte „Set Speed“ sowie die aktuell gefahrene Geschwindigkeit (mittig). Regelt das Fahrzeug auf ein Geschwindigkeitslimit, wird dies durch ein entsprechendes Verkehrszeichen rechts angezeigt.

Die Hands-Off-Warnung wird visuell über eine Veränderung der Farbe des Lenkrads angezeigt sowie akustisch über einen Warnton. Die Hands-Off-Warnung wird in zwei Stufen ausgegeben. Die erste Stufe warnt den Fahrer durch ein gelbes Lenkrad-Symbol nach 15 Sekunden, wenn er die Hände nicht am Lenkrad hat. Die zweite Stufe gibt nach weiteren 15 Sekunden ein rotes Lenkrad-Symbol sowie eine zusätzliche akustische Warnung aus.

Für das Bedienkonzept wurde am Lenkrad ein Bedienelement von BMW verbaut und durch entsprechende Aufkleber auf den Tasten mit den Funktio-



Bild 3-7: Visuelle Anzeige des Systemzustands im Kombi-Instrument (links: Screenshot; rechts: implementiert im BASt-Simulator)



nen beschriftet. Die Lenkradtastenzuordnung wurde wie in Bild 3-8 festgelegt.

### Prüfstrecke

Die Prüfstrecke bestand aus zwei relativ langen Abschnitten, in denen mit dem automatisierten System in Level 2 gefahren werden sollte, je nach Bedingung mit den unterschiedlichen Varianten der Querführung (ca. 15 Minuten). An diese längeren Fahrten mit aktivem System schloss sich jeweils eine Prüfsituation an, d. h. es erfolgte eine Systemgrenze (ohne Warnung), die dann von den Fahrern bewältigt werden musste. Um zu prüfen, ob die Art der Systemgrenze (in Querführung: Spurerkennung fällt aus bzw. in der Längsführung: stehendes Vorderfahrzeug kann nicht erkannt werden) einen Einfluss auf die Fahrerreaktion hat, wurden zwei Prüf-szenarien implementiert und deren Reihenfolge zwischen den Probanden permutiert. Als Prüf-szenario für den Ausfall der Querführung wurde eine

scharfe Rechtskurve mit einem Radius von 300 Meter umgesetzt (identisch zum Szenario „Scharfe Kurve“ in der WIVW-Studie), in der sich die Querführung in den Stand-by schaltete, und der Fahrer in die Lenkung eingreifen musste, um nicht vom Fahrstreifen abzukommen. Als Prüf-szenario für den Ausfall der Längsführung diente ein liegengeliebtes Pannenfahrzeug im eigenen Fahrstreifen. 170 m vor diesem Fahrzeug stand ein Warndreieck, das auf das Hindernis hinwies. In dem Szenario selbst war kein zusätzlicher Verkehr, die Streckenführung war gerade ohne Sichteinschränkungen. Das System erkannte das stehende Vorderfahrzeug nicht, weshalb sowohl die Quer- als auch die Längsregelung auf die eingestellte Geschwindigkeit aktiv blieben und der Fahrer eingreifen und die Fahraufgabe übernehmen musste, um durch Ausweichen und/oder Bremsen eine Kollision zu vermeiden. Tabelle 3-5 zeigt den resultierenden Prüfparcours für diese Fragestellung.

### Versuchsablauf

Nach einer kurzen Begrüßung des Probanden wurde er zunächst über den Inhalt der Studie aufgeklärt und darüber informiert, welche Daten von ihm erhoben werden. Anschließend wurde er gebeten, die Datenschutzerklärung auszufüllen. Anschließend wurde allen Probanden die Bedienungsanleitung für das System mit der Bitte diese zu lesen ausgehändigt. In dieser Anleitung wurden die grundsätzlichen Funktionen in der Quer- und Längsführung, die Bedienung der Systeme sowie die Grenzen der



Bild 3-8: Bedienlogik und verwendete Lenkradtastenzuordnung im BAST-Simulator

Lauf. Nr.	Szenario	Beschreibung	Länge [m]
1	Systemaktivierung nach Autobahnauffahrt	Fahrer wird aufgefordert L2 zu aktivieren	4.000
2	Fahren mit aktivem System auf Autobahn – 1	Wiederholt/keine Ausfälle der Querführung und/oder variable Spurführung	24.000
3	Prüfsituation 1: Scharfe Kurve (RF permutiert mit Situation 2)	Querführung fällt aus (ohne Warnung)	2.000
4	Systemreaktivierung nach Übernahme 1	Fahrer kann System wieder aktivieren	4.000
5	Fahren mit aktivem System auf Autobahn – 2	Wiederholt/keine Ausfälle der Querführung und/oder variable Spurführung	24.000
6	Prüfsituation 2: Hindernis (RF permutiert mit Situation 1)	Vorderfahrzeug wird nicht erkannt (ohne Warnung)	2.000
7	Systemreaktivierung nach Übernahme 2	Fahrer kann System wieder aktivieren	4.000
8	Fahren mit aktivem System auf Autobahn - 3	Wiederholt/keine Ausfälle der Querführung und/oder variable Spurführung	2.000
9	Abfahrt von Autobahn	Fahrer sollte System deaktivieren	2.000
Gesamtlänge (ca. 35 Minuten)			68.000

Tab. 3-5: Ablauf Prüfstrecke für Studie „Spurhaltung“



Systeme exemplarisch erläutert. Beide im Rahmen des Versuchs simulierte Systemgrenzen (Hindernisse und Kurve) wurden als Beispiel in der Bedienungsanleitung genannt. Zudem wurde explizit darauf hingewiesen, dass der Fahrer weiterhin auf das Verkehrsgeschehen und das Verhalten des Systems achten muss, da das Fahrzeug nicht die vollständige Überwachung der Fahrzeugumgebung übernimmt und der Fahrer die Hände jederzeit am Lenkrad behalten muss. Der Proband durfte Nachfragen stellen, falls ihm während des Durchlesens etwas unklar war. Unabhängig der jeweiligen Versuchsbedingung erhielten alle Probanden die gleiche Bedienungsanleitung. Das Lesen der Anleitung fand direkt im Simulator-Mock-Up statt, sodass die Probanden die Bedienelemente sowie Displays direkt vor Augen hatten.

Nach spätestens 15 Minuten wurden die Probanden gebeten, das Lesen der Anleitung zu beenden und einen Fragebogen zur Technikaffinität, zum Vertrauen in das System auf Basis der Informationen der Bedienungsanleitung sowie zur aktuellen Müdigkeit zu beantworten. Anschließend erfolgte eine kurze Eingewöhnungsfahrt von 5 Minuten, in denen sich die Fahrer wieder an das Fahren im Simulator gewöhnen konnten sowie die Bedienung der Systeme (Aktivierung, Übersteuerung, Deaktivierung, etc.) einmal unter Anleitung des Versuchsleiters erproben konnten.

Im Anschluss erfolgte die eigentliche Prüffahrt im Simulator. Hierzu erhielten die Probanden zunächst die Einweisung in die Online-Befragung während der Fahrt. Dazu wurden den Fahrern die verwendeten Skalen (7-stufige Skala von -3 bis +3 für die Fragen 1-5 auf der Tablet-App sowie die 11-stufige Skala zur Bewertung der Situationskritikalität) und deren Anwendung erklärt. Die Fahrer wurden gebeten, möglichst spontan und zügig auf die Fragen zu antworten und sich weitere Anmerkungen und Kommentare auf die Nachbefragung nach der Fahrt aufzusparen. Weiterhin erhielten sie generelle Instruktionen zum Ablauf der Prüffahrt. Sie wurden gebeten, zunächst selbst auf die Autobahn aufzufahren und erst dann das System zu aktivieren. Während der Fahrt auf der Autobahn sollte nach Möglichkeit immer mit aktivem System gefahren werden (d. h. mit aktivem ACC und aktivem Spurhalteassistenten). Die Probanden sollten immer auf dem rechten Fahrstreifen bleiben, 120 km/h fahren und sich an die Straßenverkehrsordnung halten. Sie wurden nochmals explizit daran erinnert, dass sie die Verkehrssituation jederzeit überwachen müssen und

ihre Hände am Lenkrad behalten müssen. Anschließend wurde die Simulation gestartet und der Fahrer absolvierte den Prüfparcours, in dem er alle relevanten Use Cases erlebte.

Während der Fahrt wurde er nach den Use Cases unmittelbar anhand der vorab pro Szenario definierten Fragen befragt. Der Versuchsleiter protokollierte mittels der Fahrt anhand der S.A.D.E.-Anwendung aufgetretene Bedien-, Fahr- und Überwachungsfehler. Für den Fall, dass er sich bei einer Bewertung unsicher war, etwas nicht eindeutig beobachten konnte, einen Teil der Fragen aus Zeitgründen nicht stellen konnte oder fehlerhafte Eingaben korrigieren wollte, hatte er einen zusätzlichen Protokollbogen, auf dem er diese Sonderfälle pro Szenario händisch protokollieren konnte.

Nach Absolvieren der Fahrt wurde der Fahrer gebeten, kurz im Simulator sitzen zu bleiben. Der Versuchsleiter pausierte die Simulation, um alle eingegebenen Formulare pro Szenario mittels der Vor- und Zurück-Blätter-Funktion in der App kurz auf Vollständigkeit zu überprüfen, fehlende Eingaben zu ergänzen und fehlerhafte Eingaben zu korrigieren. Erst dann wurde die Simulation vollständig gestoppt und der Fahrer konnte aussteigen. Zum Abschluss erfolgte eine Nachbefragung, in der die Probanden erneut das Vertrauen in das System auf Basis der Eindrücke der Versuchsfahrt, die aktuelle Müdigkeit sowie die Beanspruchung der Versuchsfahrt bewerten sollten. Danach wurden die Probanden über den Versuch aufgeklärt und verabschiedet. Insgesamt dauerte der Versuch ca. 1.5 h.

### Stichprobe

An der Studie nahmen 56 Probanden (davon 24 weiblich) teil. Die Probanden waren zwischen 18 und 62 Jahren alt. Das durchschnittliche Alter betrug  $M = 37.1$  Jahre ( $SD = 13.4$ ). Ausgewählt wurden ausschließlich naive Nutzer, also Personen, die nicht regelmäßig mit Längs- und/oder Querführungsassistenten fahren oder gefahren sind, um einen herstellerbezogenen Gewöhnungseffekt auszuschließen. Ein weiteres Kriterium war, dass die Probanden eine jährliche Fahrleistung von 3.000 km überschreiten sollten, um eine gewisse Routine und Fahrerfahrung voraussetzen zu können.

Jeder Proband durchlief zunächst ein neunzigminütiges Simulatortraining mit dem Ziel, sowohl die Verträglichkeit als auch die Fahrzeugführung, Geschwindigkeiten und Distanzen beim simulierten

Fahren zu erproben und zu testen. Dieses Training fand ohne Verwendung der teilautomatisierten Systeme (ACC, Spurhalteassistent) statt. Die Probanden wurden vor dem Versuch randomisiert den jeweiligen Versuchsbedingungen zugeteilt.

### Abhängige Variablen und statistische Analysen

Als abhängige Variable wurden für diesen Bericht folgende Parameter ausgewertet:

- VL-Rating zur Bewältigung des Szenarios (Skala von 0-10) pro Szenario,
- Auftreten einzelner Bedien-, Überwachungsprobleme und Fahrfehler pro Szenario (0/1),
- Analyse der Zeitverlaufsdaten aus SILAB für die Szenarien „Hindernis“ und „Scharfe Kurve“
  - Szenario „Hindernis“: Minimale TTC vor dem Pannenfahrzeug (TTC\_min), Anzahl der Kollisionen mit dem Pannenfahrzeug,
  - Szenario „Scharfe Kurve“: Standardabweichung der lateralen Position in der Kurve, Anzahl Spurabkommen (rechts oder links).
- Subjektive Bewertung der MMS mittels spezifisch pro Szenario definierter Online-Befragung (-3..+3).
- Subjektive Bewertung der Situationskritikalität (Skala von 0-10)
  - „Wie kritisch bewerten Sie die Situation?“ (im Szenario Hindernis und scharfe Kurve)

Die Online-Verhaltensbeobachtung und -befragung erfolgte mittels der in Kapitel 3.3 beschriebenen App S.A.D.E. Mittels der App kann das Auftreten von Fahrfehlern, das VL-Rating zur Bewältigung des Szenarios sowie die subjektive Bewertung der Situationskritikalität in den beiden Prüfzuszenarien in Abhängigkeit der unterschiedlichen Systemvarianten ausgewertet werden. Außerdem können die Zeitverlaufsdaten geeigneter Fahrdaten aus SILAB analysiert werden. Der Versuchsleiter war vorab anhand der Schulungsmaterialien in der Anwendung der Benutzeroberfläche und der Vergabe der Fehler und des Ratings allgemein geschult worden. Zudem wurden in Vorversuchen spezifisch für die in der vorliegenden Simulatorstudie zu erwartenden Bedien- und Fahrprobleme Regeln zur Vergabe der Fehlerkategorien aufgestellt (z. B. was in einem speziellen Fall als Bedienfehler, was als nicht-erfolgreiche Bedienung zu werten ist).

### Ergebnisse

Die Datenauswertung erfolgte mittels des IBM-Statistik-Programms SPSS in der Version 21. Für die Analyse der kontinuierlichen Daten wurden univariate ANOVAs mit drei unabhängigen Faktoren gerechnet (Faktor 1: Reliabilität der Querführung; Faktor 2: Variabilität der Querführung; Faktor 3: Reihenfolge der kritischen Situationen „Hindernis“ und „Scharfe Kurve“). Für die Analyse der Häufigkeitsverteilungen der beobachteten Bedienprobleme und Fahrfehler pro Szenario wurden Chi<sup>2</sup>-Tests jeweils getrennt für die drei Faktoren gerechnet.

Die Ergebnisse der Studie zeigen, dass alle Probanden mit der Systemaktivierung nach Autobahnauffahrt keine Probleme hatten. Das VL-Rating zur Bewältigung dieses Szenarios lag bei allen Probanden über alle Bedingungen hinweg zwischen 0 und 4. Somit haben alle Probanden dieses Szenario gut bis zufriedenstellend bewältigt. Signifikante Unterschiede zwischen den Bedingungen wurden nicht festgestellt ( $p > .05$ ). Da alle Probanden das hinsichtlich der Aktivierung gleiche System erprobt haben, wurden hier jedoch auch keine Unterschiede erwartet.

In den beiden jeweils 15-minütigen Phasen, in denen die Probanden mit aktivem System fahren sollten, gab es lediglich in der ersten Phase einen signifikanten Unterschied ( $p < .05$ ) hinsichtlich der Häufigkeit des Spurabkommens: Probanden in der Bedingung mit Ausfällen (geringe Reliabilität) sind hier im Mittel  $M = 3.04$  ( $SD = 3.84$ ) mal vom Fahrstreifen abgekommen, Probanden ohne Ausfälle hingegen nur durchschnittlich  $M = 1.15$  ( $SD = 2.70$ ) mal. In der zweiten 15-minütigen Phasen nach dem Erleben der ersten Systemgrenze reduzierten sich die Spurabkommen in beiden Gruppen im Vergleich zur ersten Phase. Ein signifikanter Effekt konnte in dieser zweiten Phase dann nicht mehr festgestellt werden ( $p > .05$ ). Es kann also davon ausgegangen werden, dass Probanden in der Bedingung mit Ausfällen im Laufe der Fahrt gelernt haben, dass das System ausfällt und ein Abkommen vom Fahrstreifen durch manuellen Eingriff verhindert werden muss. Signifikante Unterschiede in der Anzahl an Hands-Off-Warnungen, der Bewertung der Aufmerksamkeit der Probanden während der Fahrt oder hinsichtlich des VL-Ratings konnten für beide Phasen nicht festgestellt werden ( $p > .05$ ).

In der Prüfsituation 1, in welcher die Querführung in der Kurve ausfällt beeinflusst die Reliabilität des

Systems die Stabilität der manuellen Spurführung nach Ausfall der Querführungsassistenten tendenziell signifikant ( $p = .06$ ). Probanden mit Ausfällen (geringe Reliabilität) weisen im Durchschnitt ( $M = 0.35$ ,  $SD = 0.30$ ) eine geringere mittlere Standardabweichung der lateralen Position unmittelbar nach Ausfall der Querführung auf als Probanden ohne Systemausfälle ( $M = 0.50$ ,  $SD = 0.29$ ). Keinen signifikanten Unterschied gibt es zwischen den experimentellen Bedingungen der Variabilität ebenso wie der Reihenfolge ( $p > .05$ ). Hinsichtlich der Anzahl der Spurabkommen konnte ein signifikanter Effekt für den Faktor Reliabilität für Spurabkommen nach rechts ( $p < .05$ ), nicht jedoch für Spurabkommen nach links festgestellt werden ( $p > .05$ ). Probanden ohne Ausfälle kamen signifikant häufiger rechts vom Fahrstreifen ab ( $n = 26$ ) als Probanden mit Ausfällen ( $n = 20$ ). Für die anderen beiden Faktoren konnten keine signifikanten Effekte festgestellt werden ( $p > .05$ ).

Das VL-Rating der S.A.D.E.-App zeigt einen signifikanten Effekt für den Faktor Reliabilität in dieser Prüfsituation ( $p = .04$ ). In der Bedingung ohne Ausfälle lag das durchschnittliche VL-Rating bei  $M = 3.3$  ( $SD = 2.1$ ), in der Bedingung mit Ausfällen hingegen bei  $M = 2.2$  ( $SD = 1.8$ ). Für die anderen beiden Faktoren konnte kein signifikanter Effekt gefunden werden ( $p > .05$ ). Die Korrelation zwischen der Standardabweichung der lateralen Position und dem VL-Rating liegt bei  $r = .80$  ( $p < .01$ ). Es kann daher davon ausgegangen werden, dass die Qualität der Fahrerreaktion in der Prüfsituation 1 insgesamt äußerst zufriedenstellend durch das VL-Rating gemessen und bewertet werden kann.

In der Prüfsituation 2, in welcher zunächst ein Warn-dreieck und kurz darauf ein Pannenfahrzeug als Hindernis auf dem rechten Fahrstreifen erscheint, zeigt sich, dass der Faktor Reliabilität einen signifikanten Einfluss auf die TTC\_min hat ( $p = .03$ ). Bei einer hohen Reliabilität in der Bedingung ohne Ausfälle kann eine signifikant niedrigere TTC\_min ( $M = 0.88$ ,  $SD = 0.73$ ) festgestellt werden als bei einer niedrigen Reliabilität ( $M = 1.36$ ,  $SD = 0.86$ ). Keinen signifikanten Unterschied gibt es zwischen den experimentellen Bedingungen der Variabilität ebenso wie der Reihenfolge ( $p > .05$ ). Hinsichtlich der Anzahl an Kollisionen mit dem Pannenfahrzeug konnte für den Faktor Reliabilität des Systems festgestellt werden, dass Probanden in der Bedingung ohne Ausfälle ( $n = 5$ ) deskriptiv mehr als doppelt so häufig mit dem Hindernis kollidieren als Probanden mit Ausfällen ( $n = 2$ ). Dieser Unterschied ist statis-

tisch jedoch nicht signifikant ( $p > .05$ ). Für die anderen beiden Faktoren konnten ebenfalls keine signifikanten Effekte gefunden werden ( $p > .05$ ).

Das VL-Rating zur Bewältigung der Prüfsituation 2 der S.A.D.E.-App zeigt, dass das Szenario mit einem hoch reliablen System tendenziell kritischer gelöst wurde als mit einem System mit niedriger Reliabilität ( $p = .08$ ). Beim System mit niedriger Reliabilität haben 50.0 % der Probanden das Szenario gut, 38.5 % akzeptabel und 11.5 % gefährlich gelöst. Beim System mit einer hohen Reliabilität hingegen haben 29.6 % der Probanden das Szenario gut, 33.8 % akzeptabel und 37.0 % gefährlich gelöst. Die Korrelation zwischen der TTC\_min und dem VL-Rating liegt bei  $r = -.56$  ( $p < .01$ ). Es kann daher ebenso davon ausgegangen werden, dass die Qualität der Fahrerreaktion in der Prüfsituation 2 insgesamt zufriedenstellend durch das VL-Rating gemessen und bewertet werden kann.

Hinsichtlich der Variablen der Online-Befragung der S.A.D.E.-App konnten signifikante Effekte für die Bewertung der Situationskritikalität der Prüfsituation 2 für die Faktoren Reliabilität und Reihenfolge gefunden werden. Probanden in der Bedingung ohne Ausfälle (hohe Reliabilität) bewerten die Situation signifikant kritischer ( $M = 6.52$ ,  $SD = 2.33$ ) als Probanden in der Bedingung mit Ausfällen (geringe Reliabilität) ( $M = 5.08$ ,  $SD = 2.65$ ,  $p = .05$ ). Darüber hinaus konnte ein signifikanter Effekt für den Faktor Reihenfolge festgestellt werden, insofern, dass Probanden die Situation signifikant kritischer bewerten, wenn diese zuerst erlebt wird ( $p = .04$ ). Für die anderen Faktoren ebenso wie für die anderen Befragungs-Items des S.A.D.E.-Tools konnten keine signifikanten Effekte gefunden werden ( $p > .05$ ). Auch die Prüfsituation 1 zeigte keine signifikanten Unterschiede hinsichtlich der Befragung der Probanden ( $p > .05$ ).

Die Ergebnisse der Studie zeigen insgesamt, dass das VL-Rating des S.A.D.E.-Tools in beiden Prüfsituationen ein valider Indikator zur Vorhersage des objektiv sicherheitsrelevanten Fahrerhaltens ist. Gleichzeitig wird auch deutlich, dass ein hoch reliables System ohne Ausfälle zu einer schlechteren Reaktion der Probanden in den Prüfsituationen führt. Die Variabilität in der Querführung scheint hingegen keinen Effekt zu haben. Hier sollte jedoch angemerkt werden, dass die Spurversätze von der Spurmitte aus aufgrund der statischen Fahrsimulation für die Probanden nur rein visuell und teils schwierig wahrnehmbar waren. Die Befragungs-

Items des S.A.D.E.-Tools liefern keine validen Informationen hinsichtlich der experimentellen Manipulation in dieser Studie und sollten daher für zukünftige Studien angepasst werden. Hier wurde deutlich, dass die Items durch die Probanden teils sehr unterschiedlich verstanden und beantwortet wurden, sodass keine eindeutige Interpretation der Ergebnisse möglich ist.

### 3.5.3 Erprobung im Realverkehr

#### Vorgehen bei der Erprobung

In diesem Kapitel wird die Erprobung der Methode mit dem System „Autopilot“ im Tesla Model 3 im Realverkehr dokumentiert. Der Einsatz im Realverkehr hat das Ziel, dass Nutzer reale Seriensysteme im alltäglichen Gebrauch erleben und bewerten können. Die Erprobung der App wurde im Mai 2019 durchgeführt. Ziel dieses Tests war es, die Durchführbarkeit der Use Cases zu überprüfen bzw. eine geeignete Reihenfolge der Use Cases festzulegen. Zudem sollte die Anwendbarkeit der App durch den Versuchsleiter während des Fahrens evaluiert werden. Eine Bewertung des für den Test verwendeten Systems stand nicht im Vordergrund der Erprobung, weswegen das System und dessen Bedienung und Anzeigen im Folgenden nur kurz vorgestellt wird.

Als Testfahrzeug wurde ein Tesla Model 3 verwendet. Eine Beschreibung des Systems sowie der Teststrecke findet sich in Kapitel 2.5.3.

#### Erkenntnisse

Während der Tests zeigte sich, dass der Versuchsleiter die Fahrereingaben am besten beobachten

kann, wenn er auf dem Beifahrersitz sitzt und das Tablet auf den Knien hält. Von dort kann er auch die Verkehrssituation am besten beobachten und entscheiden, ob diese es zulässt, einen Use Case durchzuführen oder gegebenenfalls abzuwarten oder zu unterbrechen. Aus datenschutzrechtlichen Gründen wird empfohlen, keine Kameras zur Testdokumentation im Fahrzeug zu montieren, sondern Besonderheiten während des Tests schriftlich zu protokollieren.

Ob der Fahrer das System erfolgreich aktiviert oder deaktiviert hat und somit den entsprechenden Systemzustand erreicht hat, ist im Tesla durch die große und mittig positionierte Anzeige auf dem Tablet vom Beifahrersitz aus gut erkennbar. In der Regel sind die Systemzustandsanzeigen allerdings eher im Kombiinstrument dargestellt, die der Versuchsleiter eventuell schlecht einsehen kann. Auch können Bedieneingaben des Fahrers teilweise durch eine entsprechende Handhaltung verdeckt sein. Daher ist es durchaus legitim, den Probanden zu bitten, entweder prinzipiell bevor er einen Use Case durchführt, zunächst zu beschreiben, wie er plant, dies zu tun und ihn erst dann die Handlung ausführen zu lassen. Alternativ kann er auch nach einer ausgeführten Handlung befragt werden, welche Bedienelemente er gerade genutzt hat. Auf diese Weise kann er ebenfalls gebeten werden, beobachtete Systemanzeigen, Texte etc. zu beschreiben, auf die er sich bei der Beantwortung der Fragen bezieht.

Bezüglich der Instruktion zu den Use Cases zeigte sich, dass während des aktiven Fahrens mit dem System die Durchführung von Fahrstreifenwechseln möglichst vermieden werden sollte, wenn der jeweilige Erstkontakt mit dem System erfasst werden soll. Daher sollte der Proband instruiert wer-

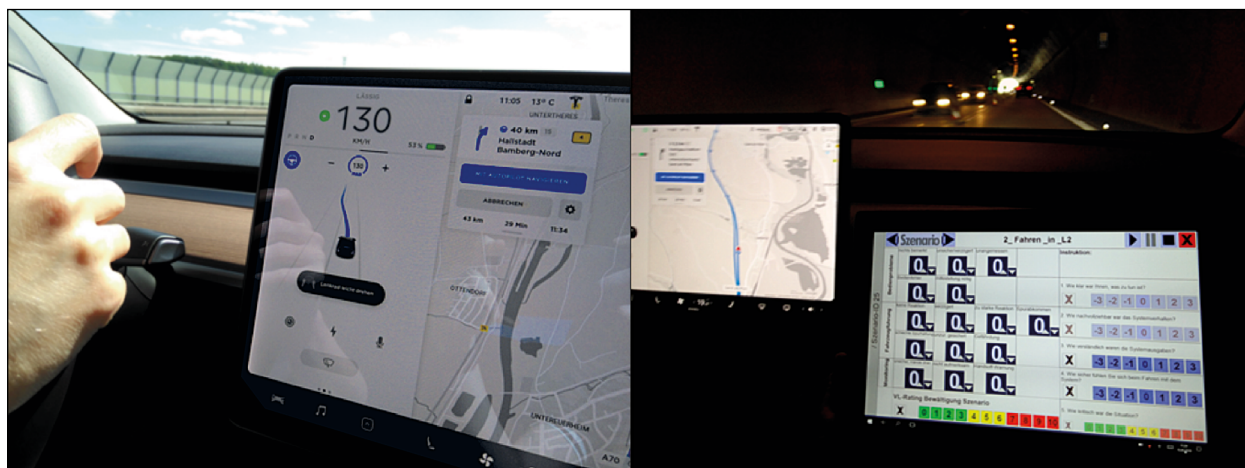


Bild 3-9: Tests im Realverkehr mit Tesla Model 3 (Quelle: WIVW GmbH)

den, nach einer ersten Systemaktivierung zunächst in gemäßigem Tempo auf dem rechten Fahrstreifen zu bleiben. Von der Streckenplanung her sollten daher auch vorhersehbare Systemgrenzen nicht unmittelbar am Anfang der Fahrt vorkommen, um den Fahrereindruck nicht zu stark zu beeinflussen, wenn er in nachfolgenden Use Cases-Systemverständnis, -Anzeigenverständlichkeit oder die -Nachvollziehbarkeit des Systemverhaltens bewerten soll. Auch wenn nicht zu vermeiden ist, dass die Fahrer bei der subjektiven Bewertung des Systems unbewusst ihre vorherigen Erfahrungen aus anderen Use Cases miteinbeziehen, sollten sie explizit dazu instruiert werden, die Bewertung ausschließlich auf die konkret erlebte Situation zu beziehen. Wenn Interesse an einem subjektiven Gesamteindruck der Fahrer zum Systemverhalten und der Interaktion mit dem System besteht, kann der Use Case „Fahrt mit aktivem System“ im zweiten Durchgang etwas länger gestaltet werden. So kann der Fahrer das System erst einmal frei erleben, bevor anschließend eine Befragung in Bezug auf den Gesamteindruck durchgeführt wird.

Insgesamt zeigte sich, dass der geplante Testablauf gut durchgeführt werden kann. Die Use Cases können innerhalb einer überschaubaren Zeitdauer absolviert werden, die Beobachtung und Befragung ist einfach anzuwenden. Der Versuchsleiter kann durch Vor- und Zurückschalten im Pause-Modus leicht überprüfen, welche Use Cases bereits durchgeführt wurden und welche noch fehlen. Es wurde erkennbar, dass einige Verbesserungen an der Reihenfolge und konkretere Instruktionen für die Probanden notwendig sind, die letztendlich zu den in Kapitel 3.3.3 beschriebenen Use Cases und deren Ablauf geführt haben.

### 3.6 Gesamtbewertung und Limitationen des Verfahrens

Die Methode der Nutzerstudie soll eine Bewertung der MMI von teilautomatisierten Fahrfunktionen ermöglichen. Mit ihr sollen vor allem interaktionsbezogene Aspekte des Fahrerhaltens und des Fahrerlebens erfasst werden, die sich darauf beziehen, wie ein Nutzer mit dem System interagiert, d. h. ob er das Systemverhalten und die Systemanzeigen versteht, das System effizient bedienen kann und angemessen und sicher auf Systemänderungen, z. B. an Systemgrenzen, reagiert. Vor diesem Hintergrund wurde ein entsprechendes Bewertungs-

tool entwickelt, mit dessen Hilfe man die definierten Bewertungskategorien des Fahrerhaltens und des Fahrerlebens effizient erfassen kann. Dieses Verfahren stellt einen von mehreren Methodenbausteinen dar, die zu einer ganzheitlichen Bewertung eines teilautomatisierten Systems notwendig sind. Durch die Tatsache, dass in Nutzerstudien tatsächliche spätere Nutzer eines Systems untersucht werden, ergänzt diese Methode in sinnvoller Weise die Expertenbewertung der MMS mittels einer Checkliste, in der wahrscheinlich auftretende Probleme nur hypothetisch von Experten definiert werden, die mit der Thematik und derlei Systemen hoch vertraut sind (siehe Kapitel 2).

Die Methode hat den Vorteil, dass durch die Erfassung des Erstkontakts sowohl das intuitive Verhalten eines Nutzers mit einem System abgebildet werden kann, aber auch Lernerfahrungen durch mehrmaliges Erleben eines Use Cases (z. B. nehmen Fahrer ihre Hände fester ans Lenkrad, nachdem sie Hands-Off-Warnungen erlebt haben) erfasst werden können. Zudem kann die Methode Reihenfolge-Effekte aufdecken (beispielsweise wirkt sich das vorherige Erleben von Systemgrenzen auf nachfolgende Reaktionen aus). Weiterhin kann die Methode den Einfluss eines spezifischen Systemverhaltens (z. B. die Güte der Spurführung) auf Erleben und Verhalten des Fahrers abbilden. Somit erweitert sie die Methode der Expertenbewertung, die sich explizit auf Gestaltungsaspekte der MMS fokussiert.

Durch die Prüfung von tatsächlichen Fahrszenarien (wie z. B. Systemgrenze), können die Auswirkungen des Systems auf die Fahrsicherheit ermittelt werden. Eine größere Anzahl an Probanden, die eine gewisse Bandbreite des Verhaltens abdecken, ermöglicht eine statistische Schätzung der Auftretenswahrscheinlichkeit eines bestimmten beobachteten Problems (sei es im Bedienenverhalten, im Überwachungsverhalten oder im Fahrverhalten) in der Gesamtbevölkerung. Im Response-Code-of-Practice (Response, 2006) beispielsweise wird zur Prüfung der Kontrollierbarkeit eines Assistenzsystems die Testung von 20 Probanden gefordert, von denen alle ein definiertes Kriterium erfüllen müssen, damit mit einer Fehlerwahrscheinlichkeit von unter 5 % auf eine Kontrollierbarkeit für 85 % der Gesamtbevölkerung geschlossen werden kann. Welches absolute Pass/Fail-Kriterium bzw. welche Kombination von Entscheidungskriterien festgesetzt werden müsste, anhand dessen entschieden werden kann, ob ein System die notwendigen An-

forderungen an Usability und Sicherheit erfüllt, muss weiterführend diskutiert werden. Für das entwickelte Bewertungstool bietet sich das Versuchsleiter-Rating an, das sich aus der Beobachtung der verschiedenen Fehlertypen in den Prüfscenarien zusammensetzt und zur Bewertung des Gesamtsystems genutzt werden kann. Eventuell könnte man als Pass-Fail-Kriterium formulieren, dass nur ein bestimmter Prozentsatz einer getesteten Stichprobe in der Nutzerstudie in einer der Prüfscenarien einen Wert  $> x$  auf dem VL-Rating haben darf, damit das System den Test besteht.

Potenzielle Einschränkungen des Verfahrens betreffen vor allem die Befragungskomponente der Methode. Beispielsweise sind Probanden nicht darin geschult, einzelne System- oder MMS-Charakteristika zu betrachten und andere auszublenden, sondern betrachten ein System eher als eine Gesamtheit. Insofern ist davon auszugehen, dass z. B. das Systemverhalten, aber auch die eigenen Vorannahmen über ein System in eine Bewertung miteinbezogen werden, auch wenn dies in dem Fall der Bewertung der MMS eher störend ist. Dies ist zum Teil auch durch die Instruktion beeinflusst, die darauf fokussiert, dass der Proband im Folgenden ein bestimmtes System kennenlernen wird und dieses dann bewerten soll. Im Gegensatz dazu bekommt der Experte durch die Items der Checkliste wesentlich eindeutiger die Anweisung, auf welche Charakteristika er zu achten hat und kann seine Aufmerksamkeit entsprechend ausrichten. Zudem darf er die Use Cases mehrmals ausprobieren und erleben, um zu einer Beurteilung zu kommen. Außerdem sind die Probanden durch die Versuchsanordnung eventuell so stark gefordert, dass sie ihre Aufmerksamkeit nicht ausschließlich auf die Systemrückmeldungen konzentrieren können, und dadurch bestimmte Aspekte evtl. gar nicht wahrnehmen. Gerade, wenn Use Case nur einmal erlebt werden, kann eine Bewertung unter diesen Umständen nicht sinnvoll vorgenommen werden.

Aufgrund der Einschränkungen der Methode, was die Bewertungsmöglichkeiten von Gestaltungsaspekten der MMS durch Nutzer angeht, ist generell zu hinterfragen, inwieweit die subjektive Bewertung der Probanden in die Gesamtbewertung des Systems miteingehen sollte. Zwar sind subjektive Eindrücke im Umgang mit dem System eine wertvolle Informationsquelle zur Ableitung von Verbesserungspotenzial der MMS. Es erscheint jedoch für eine Gesamtbewertung sinnvoll, zusätzlich zu Nutzerstudien, bzw. im Vorfeld derer, die Expertenbeur-

teilung mittels der Checkliste durchzuführen. Diese bewertet, ob es für jeden Systemzustand überhaupt einen entsprechenden Indikator gibt, anhand dessen abgeleitet werden kann, welche Aufgaben das System gerade übernimmt bzw. welche der Fahrer ausführen muss, und ob diese so gestaltet sind, dass der Fahrer sie gut wahrnehmen, deren Bedeutung verstehen und die relevanten notwendigen Handlungen daraus ableiten kann. Zudem wird evaluiert, ob das System gut bedient werden kann und ob dem Fahrer über die Bedienungsanleitung alle wesentlichen Vorinformationen in verständlicher Form präsentiert werden. Diese Aspekte werden separat über einzelne Items der Checkliste von zwei Ratern in einem vorgegebenen Set von Use Cases (die denen der Nutzerstudie weitgehend entsprechen) dahingehend bewertet, ob sie akzeptabel umgesetzt sind und zu einem Gesamturteil integriert. Nachdem die Rater als Human Factors Experten entsprechend darin geschult sind, einzelne Aspekte der MMS zu beurteilen, scheint das Verfahren der Checkliste für diese Aspekte das geeignetere Instrument zu sein.

Weitere Einschränkungen betreffen die Anwendung der Methode im Realverkehr. Aufgrund von Sicherheitsaspekten ist es hier nicht möglich, potenziell gefährliche Situationen, wie unvorhersehbare Systemgrenzen oder Systemausfälle und deren Auswirkungen auf das Fahr- und Fahrerverhalten zu untersuchen. Für die Gesamtbewertung des Systems wären solche Szenarien allerdings durchaus relevant, da sonst gerade die Sicherheitsaspekte der Systeme nicht ausreichend betrachtet werden können. In Simulator- oder Teststreckenstudien können derartige Effekte ohne echte Gefährdung der Fahrsicherheit zumindest abgeschätzt werden. Für eine realistische Bewertung der Kontrollierbarkeit ist es allerdings empfehlenswert, Tests im realen Fahrzeug auf einer abgesperrten Teststrecke durchzuführen.

Ebenfalls problematisch für die Anwendung der Methode im Realverkehr ist, dass dort keine kontrollierten Bedingungen herrschen, die sicherstellen, dass alle Fahrer die Use Cases in gleicher Form erleben. Dies betrifft beispielsweise unterschiedliche Verkehrskonstellationen und unterschiedliche Geschwindigkeiten während der Durchführung eines Use Cases, die zu einer unterschiedlich starken Beanspruchung des Fahrers führen und sich somit unterschiedlich auf das Fahrerverhalten auswirken können. Auch unter diesem Aspekt wäre eine Anwendung der Methode auf der Teststrecke, in der

die Bedingungen deutlich besser kontrolliert werden können, empfehlenswert. Unter Umständen könnte man eine Fahrt im Realverkehr anschließen, in der der Fahrer eine längere Fahrt mit dem aktiven L2-System erleben kann, um den Gesamteindruck zum System abzufragen.

## 4 Fazit Gesamtbewertungskonzept

Der Bericht stellt zwei standardisierte methodische Ansätze zur Bewertung der Effizienz und Sicherheit der Mensch-Maschine-Interaktion (MMI) bei der Nutzung von teilautomatisierten Fahrfunktionen vor. Dazu wurden Prüfkriterien definiert, welche die Erfüllung der dazu notwendigen Anforderungen an die Vermittlung eines angemessenen Systemwissens, eines adäquaten Mode- und Situationsbewusstseins, einer wenig beanspruchenden Systembedienung und einer sicheren Fahrerreaktion an Systemgrenzen erfassen sollen.

Mittels einer Expertenbewertung wird geprüft, ob die Mensch-Maschine-Schnittstelle (MMS) die dazu notwendigen Mindestanforderungen bezüglich Wahrnehmbarkeit und Verständlichkeit von Systemzuständen erfüllt. Der Fokus liegt dabei auf visuellen Anzeigen und akustischen Systemausgaben. Zusätzlich werden Aspekte der Systembedienung sowie die Vermittlung eines angemessenen Systemverständnisses über das Benutzerhandbuch bewertet. Dazu wurde eine Checkliste mit Anforderungen entwickelt, die aus bestehenden Standards und Gestaltungsrichtlinien abgeleitet wurden. Während einer Fahrt im Realverkehr, im Simulator oder auf der Teststrecke mit dem zu prüfenden System wird ein Set an Use Cases durchfahren und die Erfüllung der Anforderungen durch eine Bewertung der zugehörigen Unteritems beurteilt. Die Anwendung der Checkliste wurde im Rahmen von zwei Realfahrtstudien und einer Fahrsimulatorstudie erprobt. Ziel war die Überprüfung der Validität, (d. h. kann die Checkliste unterschiedlich gute Mensch-Maschine-Interaktionskonzepte voneinander unterscheiden) sowie der Reliabilität (d. h. inwieweit kommen unterschiedliche Rater bei der Bewertung eines Systems auf vergleichbare Ergebnisse). Die Ergebnisse der Erprobungsphase fanden Eingang in die Optimierung der Checkliste und des Anwendungsverfahrens.

Als weitere Methode wurde ein standardisiertes, szenarienbasiertes Beobachtungs- und Befragungsinstrument für den Einsatz in Nutzerstudien entwickelt. Als Bewertungskriterien werden zum einen Aspekte des Fahrerverhaltens genutzt, die über eine Beobachtung des Fahrers in der Interaktion mit dem System über die MMS und der Kodierung von Problemen in der Bedienung, des Fahrverhaltens und des Überwachungsverhaltens erfasst werden und schließlich in ein übergeordnetes Versuchsleiterring pro Use Case münden. Zum anderen werden Aspekte des Systemverständnisses über eine Befragung des Fahrers zu einzelnen MMI-Anzeigen und Systemreaktionen erfasst. Diese Bewertungskategorien sind in anschaulicher und leicht anzuwendender Form in einer Tablet-Anwendung, der sog. S.A.D.E.-App umgesetzt. Die Methode wurde in zwei Simulatorstudien und während einer Fahrt im Realverkehr mit einem Seriensystem erprobt und validiert.

Für beide Methoden wurden jeweils Vor- und Nachteile identifiziert: Die Methode der Expertenbewertung erlaubt es, die Voraussetzungen für ein korrektes Verständnis von L2-Systemen anhand der Betriebsanleitung und der Nutzerschnittstelle zu evaluieren und Unzulänglichkeiten in der Gestaltung zu identifizieren. Ihr Fokus liegt auf der Bewertung der Ausgaben der MMS als Voraussetzung für einen sicheren Systemgebrauch.

Die Methode hat den Vorteil, dass durch (wiederholtes) Ausprobieren alle Use Cases vollständig erlebbar sind, wodurch alle Items bewertet werden können, was in Nutzerstudien, die das natürliche Verhalten im Umgang mit dem System untersuchen, nicht sichergestellt werden kann. Außerdem können Experten besser als Nutzer Gestaltungsaspekte bewerten, da sie sich ausschließlich auf diese Aufgabe konzentrieren können, durch ihre Erfahrung auch ohne einen direkten Vergleich ein gutes von einem schlechten System zu unterscheiden, typische Testpersoneneffekte (z. B. das Bestreben, ein „guter“ Proband sein zu wollen) unwahrscheinlich sind und sie sich nicht von falschen Vorannahmen über die Funktionsweise eines L2-Systems beeinflussen lassen sollten.

Als Nachteil der Methode muss herausgestellt werden, dass durch das Checklistenverfahren nur beschränkt Aussagen über den intuitiven Erstkontakt durch naive Nutzer möglich sind. Zudem kann nicht geprüft werden, ob und wie stark sich identifizierte

Unzulänglichkeiten in der Gestaltung letztlich im Verhalten von Fahrern niederschlagen. Dazu bedarf es der Durchführung von Nutzerstudien, in denen die direkte Interaktion von Nutzern mit dem System abgebildet werden kann. Eine Nutzerstudie kann zudem den Einfluss der Interaktion aus Systemverhalten und MMS auf Verhalten und Erleben des Fahrers abbilden (z. B. Systemeinflüsse, die über die MMS hinausgehen und durch die Checkliste nicht abgedeckt sind wie etwa die Güte der Spurführung).

Mit der Methode der Nutzerstudie sollen vor allem interaktionsbezogene Aspekte des Fahrerverhaltens und des Fahrererlebens erfasst werden, die sich darauf beziehen, wie ein „echter“ Nutzer mit dem System interagiert, d. h. ob er das Systemverhalten und die Systemanzeigen versteht, das System effizient bedienen kann und angemessen und sicher auf Systemänderungen, z. B. an Systemgrenzen, reagiert.

Die Methode hat den Vorteil, dass sowohl das intuitive Verhalten eines Nutzers mit einem System abgebildet werden kann, aber zudem auch Lernerfahrungen durch mehrmaliges Erleben eines Use Cases (z. B. nehmen Fahrer ihre Hände fester ans Lenkrad, nachdem sie Hands-Off-Warnungen erlebt haben). Zudem kann sie Reihenfolge-Effekte aufdecken, z. B. dass sich das vorherige Erleben von Systemgrenzen auf nachfolgende Reaktionen auswirkt. Weiterhin kann die Methode den Einfluss eines spezifischen Systemverhaltens (z. B. die Güte der Spurführung) auf Verhalten und Erleben des Fahrers abbilden.

Im Gegensatz zur Checkliste können die unmittelbaren Auswirkungen auf das Fahrverhalten und die Fahrsicherheit erfasst werden, indem Use Cases mit Systemgrenzen oder Systemausfällen getestet werden. Die Umsetzung solcher Use Cases lässt sich in der Fahrsimulation oder auf einer Teststrecke jedoch deutlich leichter und ohne eine Gefährdung des Nutzers realisieren als im Realverkehr.

Für die Bewertung der Kontrollierbarkeit an Systemgrenzen gibt es bereits Verfahren, wie beispielsweise für L3-Systeme das sog. TOC-Rating (siehe NAJJOKS et al., 2018a). Es handelt sich dabei um eine expertenbasierte Bewertung der Kontrollierbarkeit, in dem anhand vorgegebener Beobachtungskategorien das Fahrerverhalten aus Videosequenzen auf einer 10-stufigen Skala bewertet wird. Insofern ähnelt es dem hier vorgestellten Verfahren

der S.A.D.E.-App, die u. a. an Systemgrenzen das Fahrerverhalten bewertet und Aspekte der Kontrollierbarkeit, wie die Kategorien der Gefährdung oder das Auftreten von Kollisionen berücksichtigt. Wesentliche Unterschiede sind allerdings, dass die Beobachtung des Fahrverhaltens im TOC-Rating wesentlich detaillierter erfolgen soll, was aufgrund der nachträgliche Kodierung und des wiederholten Betrachtens des Videos zwar leichter möglich ist, aber mit deutlich höherem Schulungsaufwand für die Rater verbunden ist. Zudem ist die Bewertung rein auf Übernahmesituationen beschränkt, Aspekte der Usability werden nicht betrachtet. Im Gegensatz dazu wurde die S.A.D.E.-Anwendung mit dem ausdrücklichen Ziel einer umfassenden Bewertung entwickelt, d. h. neben der Erfassung der Fahrsicherheit auch weitere Aspekte, wie der Systembedienbarkeit (mit ihren ebenfalls möglichen Beeinträchtigungen der Fahrleistung) und der Mode Awareness (in L2 vor allem das Bewusstsein über die Verantwortlichkeit für die Überwachungsaufgabe). Dies wurde über die szenarienbasierte Bewertung in der gesamten Bandbreite der möglichen Use Cases Rechnung getragen. Aufgrund der bereits diskutierten Probleme mit den subjektiven Bewertungen der Nutzer, was die Gestaltungsaspekte der MMS angeht, wird empfohlen, diese in der Nutzerstudie nachrangig zu betrachten und deren Bewertung eher anhand der Expertenbeurteilung vorzunehmen.

Anhand der WIVW-Simulatorstudie können die beiden Methoden Expertenbewertung und Einsatz der Tablet-App in Bezug auf ihre Bewertungsergebnisse miteinander verglichen werden. Die Checkliste erbrachte eindeutige Unterschiede zwischen den MMS-Varianten, während sich diese Unterschiede im Verhalten und in der subjektiven Bewertung nicht so stark widerspiegeln. Dies kann verschiedene Gründe haben: eine Möglichkeit ist, dass Fahrer bestimmte Aspekte in ihrem Verhalten kompensieren, z. B. eine schlechte Lesbarkeit von Anzeigen. Die daraus resultierende erhöhte Beanspruchung und evtl. Frustration würde sich aber evtl. in Aspekten wie der Systemakzeptanz auswirken, die hier nur ansatzweise erfasst wurde. Des Weiteren nehmen Fahrer bestimmte Aspekte u. U. gar nicht wahr, was sich aber in den hier gestalteten Szenarien nicht negativ ausgewirkt hat (z. B. weil ein Fahrstreifenwechsel problemlos ausgeführt werden konnte). U. U. würde sich aber eine fehlende Rückmeldung der Unterstützung beim Fahrstreifenwechsel in komplexeren Szenarien durchaus auswirken (z. B.



wenn ein Fahrstreifenwechsel aufgrund umgebenden Verkehrs erschwert ist). Zum Teil ergaben sich auch zur Checkliste widersprüchliche Ergebnisse, z. B. in der Reaktion auf das Pannenfahrzeug, die vermutlich durch Unterschiede im Systemverhalten resultierten und sich auf Aspekte auswirken, die hier nicht erfasst wurden, z. B. Systemvertrauen.

Abschließend ist deshalb zu empfehlen, für eine vollumfängliche Bewertung die beiden Methodenbausteine „Checkliste“ und „Nutzerstudien“ in Kombination zu verwenden. Beide Verfahren liefern wertvolle Erkenntnisse über das System. Eine Methode scheint die andere nicht vollständig ersetzen zu können. Für eine Gesamtbewertung des Systems wird daher aktuell die Empfehlung gegeben, die Ergebnisse beider Methoden zu nutzen, und dabei konservativ vorzugehen: Das jeweils schlechtere Ergebnis ist in die Gesamtbewertung einzubeziehen, wenn beide Methoden ein Ergebnis zu einem Aspekt liefern bzw. sind alle Aspekte einzubeziehen, die jeweils durch die andere Methode nicht abgedeckt werden können.

Die bereits beschriebenen Methoden der Nutzerstudie und der Checkliste, mit denen nach Meinung des Autorinnen-Teams die wesentlichen Aspekte der MMI abgedeckt und geprüft werden können, können um weitere Methodenbausteine ergänzt werden, die sich sinnvoll in die bestehenden Methoden integrieren lassen. Zum einen erscheint die Erweiterung um eine Methode zur Erfassung eines grundlegenden Systemverständnisses von Nutzern sinnvoll, die über die in-Situ Beobachtung des Verhaltens und der Bewertung des Systems in konkreten Situationen hinausgeht. Mit dieser sollte abge-

bildet werden können, wie gut das sich im Umgang mit einem System herausgebildete mentale Modell des Fahrers vom System ist. Hierfür wäre die geeignete Methode ein strukturiertes, standardisiertes Interview, das unmittelbar an die Nutzerstudie angeschlossen werden könnte. Es ist weiterhin denkbar, die Methode um Bausteine wie etwa Naturalistic Driving Studies zu erweitern, wenn auch Aspekte wie langfristiges Vertrauen oder Fehlgebrauch untersucht werden sollen, die aktuell nicht Teil der Methode sind. Tabelle 4-1 gibt einen Überblick über geeignete Methoden, um verschiedene Human-Factors-Aspekte der Mensch-Maschine-Interaktion zu identifizieren.

Weiterhin bietet das Verfahren kein Entscheidungskriterium, anhand dessen die Bewertung der Kontrollierbarkeit von Systemfehlern und Systemgrenzen solcher Systeme durch den Fahrer erfolgen kann. Um dies zu prüfen, müssen die Fahrer in sicherheitskritische Situationen gebracht werden, um daraufhin zu prüfen, dass ein angemessener Prozentsatz der Stichprobe solche Situationen beherrschen kann. Im Ansatz kann diese Fragestellung bereits in Nutzerstudien und mittels des hier vorgestellten S.A.D.E.-Tools beantwortet werden. Die Methode enthält zumindest in der Simulationsumgebung derartige Prüfzenarien (z. B. Scharfe Kurve, Hindernis) und entsprechende Bewertungskriterien, wie das Auftreten von Gefährdungen, Kollisionen, Spurverlassens-Ereignissen können durch die App kodiert werden. Allerdings müssten solche Kontrollierbarkeitsfragestellungen im realen Fahrzeug durchgeführt werden, um wirklich valide Ergebnisse erzielen zu können, da das konkrete, reale Systemverhalten hierbei eine entscheidende Rol-

Human-Faktors Aspekte der Mensch-Maschine-Interaktion	Geeignete Methode
Angemessene Gestaltung der MMS in Hinblick auf geltende Richtlinien (bzgl. Anzeigen, Bedienlogik und Bedienelemente, sowie Benutzerhandbuch)	Expertenbewertung mittels Checkliste
Erfassung des Verhalten und Erlebens von Fahrern als Ergebnis der Gestaltung der MMS in der konkreten Interaktion mit dem System (Usability-Aspekte)	Nutzerstudien in Simulation oder im Realverkehr
Erfassung des Verhalten und Erlebens von Fahrern als Reaktion auf besondere Systemeigenschaften/Systemverhalten in der konkreten Interaktion mit dem System	Nutzerstudien in Simulation oder im Realverkehr
Bewertung des ausgebildeten mentalen Modells, grundlegendes Systemverständnis der Nutzer	Strukturiertes Interview
Bewertung der Kontrollierbarkeit des Systems an Systemgrenzen und Systemfehlern	Nutzerstudien auf Teststrecke mit Realisierung von sicherheitskritischen Situationen
Langfristige Aspekte, wie Entwicklung von Systemvertrauen, Akzeptanz, Nutzungsintentionen, Verhaltensanpassungen	NDS-Studien

Tab. 4-1: Überblick über geeignete Methoden im Rahmen einer ganzheitlichen Bewertung der Mensch-Maschine-Interaktion

le spielt. Daher ist die Durchführung solcher Prüfungen eher auf Teststrecken zu empfehlen. Dort kann auch die Beherrschbarkeit anderer Systemfehler, wie fehlerhafte Lenkeingriffe durch das System, gezielt erzeugt und geprüft werden.

Interessant sind zudem die langfristigen Verhaltensanpassungen von Fahrern an teilautomatisierte Systeme, die wiederum eine andere Methodik notwendig machen würden, nämlich die unbeeinflusste Beobachtung des Fahrerverhaltens über einen längeren Zeitraum, wie es mittels „Naturalistic Driving Studies“ (sog. NDS-Studien) möglich wäre. Am Ende müsste für eine umfassende Bewertung eine Kombination aus verschiedenen Methoden und Bewertungskriterien eingesetzt werden. Erste Ansätze wurden mit der Entwicklung der hier dargestellten Methode der Nutzerstudie und der Expertenbewertung gemacht. Weitere stehen noch aus. Auch die Frage, wie bedeutsam die einzelnen Human-Factors-Aspekte für eine Gesamtbewertung erachtet werden sollten und entsprechend in Relation zueinander sowie in Relation zu technischen Aspekten gewichtet werden sollten, ist aktuell noch unbeantwortet. Hierfür sind weiterführende Überlegungen notwendig.

## Literatur

- ABRAHAM, H.; MCANULTY, H.; MEHLER, B. & REIMER, B.: Case study of today's automotive dealerships: Introduction and delivery of advanced driver assistance systems. In: *Transportation research record*, 2660(1), 7-14, 2017
- ABRAHAM, H.; SEPPELT, B.; MEHLER, B. & REIMER, B.: What's in a name: Vehicle technology branding & consumer expectations for automation. In: *Proceedings of the 9<sup>th</sup> International Conference on Automotive Interfaces and Interactive Vehicular Applications* (pp. 226-234), New York: ACM, 2017
- Adaptive Consortium: Final functional human factors recommendations (D3.3). Project Report. 2017
- Alliance of Automobile Manufacturers: Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems, 2006
- BAINBRIDGE, L.: Ironies of Automation. In: *Automatica*, 19(6), 775-779, 1983
- BANKS, V. A.; ERIKSSON, A.; O'DONOGHUE, J. & STANTON, N. A.: Is partially automated driving a bad idea? Observations from an on-road study. In: *Applied Ergonomics* 68, 138-145, 2018
- BANKS, V. A.; PLANT, K. L. & STANTON, N. A.: Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7<sup>th</sup> May 2016. In: *Safety Science*, 108, 278-285, 2018
- BEGGIATO, M. & KREMS, J. F.: The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 18, 47-57, 2013
- BEGGIATO, M.; HARTWICH, F.; SCHLEINITZ, K.; KREMS, J.; OTHERSEN I. & PETERMANN-STOCK, I.: What would drivers like to know during automated driving? Information needs at different levels of automation. In: *7. Tagung Fahrerassistenzsysteme*, 2015
- BIONDI, F. N.; LOHANI, M.; HOPMAN, R.; MILLS, S.; COOPER, J. M. & STRAYER, D. L.: 80 MPH and out-of-the-loop: Effects of real-world semi-automated driving on driver workload and arousal. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 1878-1882, 2018
- BLÖMACHER, K.; NÖCKER, G. & HUFF, M.: The role of system description for conditionally automated vehicles. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 54, 159-170, 2018
- BUBB, H.; BENGLER, K.; GRÜNEN, R. E. & VOLLRATH, M.: *Automobilergonomie*. Wiesbaden: Springer Verlag (ATZ/MTZ-Fachbuch), 2015
- CAMPBELL, J. L.; CARNEY, C. & KANTOWITZ, B. H.: Human factors design guidelines for advanced traveler information systems (ATIS) and commercial vehicle operations (CVO). United States. Federal Highway Administration, 1998
- CARNEY, R. N. & LEVIN, J. R.: Pictorial illustrations still improve students' learning from text. In: *Educational psychology review* 14(1), 5-26, 2002

- CASNER, S. M. & HUTCHINS, E. L.: What Do We Tell the Drivers? Toward Minimum Driver Training Standards for Partially Automated Cars. In: *Journal of Cognitive Engineering and Decision Making*, 13(2), 55-66, 2019
- De WINTER, J. C.; STANTON, N. A.; PRICE, J. S. & MISTRY, H.: The effects of driving with different levels of unreliable automation on self-reported workload and secondary task performance. In: *International Journal of Vehicle Design*, 70(4), 297-324, 2016
- Department of Defense: Design criteria standard: Human engineering. In: *Military Standard 1472F*, 2012
- DESMOND, P. A.; HANCOCK, P. A. & MONETTE, J. L.: Fatigue and automation-induced impairments in simulated driving performance. In: *Transportation Research Record*, 1628(1), 8-14, 1998
- DICKIE, D. A. & BOYLE, L. N.: Drivers' understanding of adaptive cruise control limitations. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 53, No. 23, pp. 1806-1810). Sage CA: Los Angeles, CA: Sage Publications, 2009
- DIN 15008: Straßenfahrzeuge – Ergonomische Aspekte von Fahrerinformations- und Assistenzsystemen – Anforderungen und Bewertungsmethoden der visuellen Informationsdarstellung im Fahrzeug. Berlin: Beuth Verlag, 2003
- DIN EN 82079-1: Erstellen von Gebrauchsanleitungen – Gliederung, Inhalt und Darstellung – Teil 1: Allgemeine Grundsätze und ausführliche Anforderungen. Berlin, Beuth Verlag, 2018
- Economic Commission for Europe: Application of information Technology to Road Safety – External intervention in speed and the Vienna Convention on Road Traffic of 8 November 1968 transmitted by Germany, TRANS/WP.1/2001/15 vom 25.1.2001
- ENDSLEY, M. R.: The application of human factors to the development of expert systems for advanced cockpits. In: *Proceedings of the Human Factors Society Annual Meeting* (Vol. 31, No. 12, pp. 1388-1392). Sage CA: Los Angeles, CA: SAGE Publications, 1987
- ENDSLEY, M. R.: Design and evaluation for situation awareness enhancement, In: *Proceedings of the Human Factors Society*, 32, 97-101, 1988
- Europäische Kommission: Empfehlung der Kommission vom 26. Mai 2008 über sichere und effiziente bordeigene Informations- und Kommunikationssysteme: Neufassung des Europäischen Grundsatzkatalogs zur Mensch-Maschine-Schnittstelle. In: *Amtsblatt der Europäischen Union* 51, 2008
- Europäische Union: Regelung Nr. 26 der Wirtschaftskommission der Vereinten Nationen für Europa (UN/ECE) – Einheitliche Bedingungen für die Genehmigung der Kraftfahrzeuge hinsichtlich ihrer vorstehenden Außenkanten. *Amtsblatt der Europäischen Union*, 37. Jg., S. 14, 2010
- FAVARÒ, F.; EURICH, S. & NADER, N.: Autonomous vehicles' disengagements: trends, triggers, and regulatory limitations. In: *Accident Analysis & Prevention*, 110, 136-148, 2018
- FELDHÜTTER, A.; HÄRTWIG, N.; KURPIERS, C.; MEJIA HERNANDEZ, J. & BENGLER, K.: Effect on Mode Awareness When Changing from Conditionally to Partially Automated Driving. In: *Congress of the International Ergonomics Association*: Springer, Cham, 2018
- FORSTER, Y.; HERGETH, S.; NAUJOKS, F.; BEGGIATO, M.; KREMS, J. F. & KEINATH, A. (a): Learning to use automation: Behavioral changes in interaction with automated driving systems. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 62, 599-614, 2019
- FORSTER, Y.; HERGETH, S.; NAUJOKS, F.; KREMS, J. & KEINATH, A. (b): User Education in Automated Driving: Owner's Manual and Interactive Tutorial Support Mental Model Formation and Human-Automation Interaction. In: *Information* 10(4), 143, 2019
- GASPAR, J. & CARNEY, C.: The Effect of Partial Automation on Driver Attention: A Naturalistic Driving Study. In: *Human Factors*, 61(8), 1261-1276, 2019
- GASSER, T. M.; ARZT, C.; AYOUBI, M.; BARTELS, A.; BÜRKLE, L. & EIER, J. et al.: Rechtsfolgen zunehmender Fahrzeugautomatisierung. In: *Berichte der Bundesanstalt für Straßenwesen. Unterreihe Fahrzeugtechnik* (83), 2012

- GOLD, C.; KÖRBER, M.; HOHENBERGER, C.; LECHNER, D. & BENGLER, K.: Trust in automation – Before and after the experience of take-over scenarios in a highly automated vehicle. In: *Procedia Manufacturing* 3, 3025-3032, 2015
- GREEN, P.; LEVISON, W.; PAELKE, G. & SERAFIN C.: *Suggested Human Factors Design Guidelines for Driver Information Systems*, 1994
- GREENLEE, E. T.; DELUCIA, P. R. & NEWTON, D. C.: Driver Vigilance in Automated Vehicles: Effects of Demands on Hazard Detection Performance. In: *Human Factors*, 61(3), 474-487, 2018
- HERGETH, S.; LORENZ, L. & KREMS, J. F.: Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. In: *Human Factors* 59(3), 457-470, 2017
- HINDER, E.: Effects of verbal and pictorial context cues on free recall and clustering of text themes. *Advances in Psychology*, Bd. 8: Elsevier, 279-289, 1982
- ISO 9241–11. Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts. International Organization for Standardization. Geneva, Switzerland: ISO; 2018
- Japan Automobile Manufacturers Association: *Guideline for in-vehicle display systems, Version 3.0*. In: Japan Automobile Manufacturers Association, 2014
- JEFFRIES, R.; DESURVIRE, H.: Usability testing vs. heuristic evaluation: Was there a contest?. In: *ACM SIGCHI Bulletin* 24(4): 39-41, 1992
- JEFFRIES, R.; MILLER, JR.; WHARTON, C.; UYEDA, K.: User interface evaluation in the real world: A comparison of four techniques. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 119-124). New York: ACM, 1991
- KAUSSNER, Y.: Assessment of driver fitness: An alcohol calibration study in a high-fidelity simulation. Paper presented at the Fit to Drive 7<sup>th</sup> International Traffic Expert Congress, Berlin, 25.04.-26.04.2013
- KAUSSNER, Y.; HOFFMANN, S.; FISCHER, F. & KENNTNER-MABIALA, R.: SPDE\_DFA: S.A.F.E: Anleitung zur Durchführung von Fahrverhaltensbeobachtungen mit der Standardized Application for Fitness to Drive Evaluations (S.A.F.E.) am Tablet-PC. Würzburg, 2013
- KIRCHER, K.; LARSSON, A. & HULTGREN, J. A.: Tactical Driving Behavior With Different Levels of Automation. In: *IEEE Transactions on Intelligent Transportation Systems*, 1(15), 158-167, 2014
- KLEEN, A. T.: Beherrschbarkeit von teilautomatisierten Eingriffen in die Fahrzeugführung. (Doctoral dissertation). Technische Universität, Braunschweig, 2014
- KÖRBER, M.; BASELER, E. & BENGLER, K.: Introduction matters: Manipulating trust in automation and reliance in automated driving. In: *Applied Ergonomics* 66, 18-31, 2018
- KRAFT, A. K.; NAUJOKS, F.; WÖRLE, J. & NEUKUM, A.: The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. In: *Transportation research part F: Traffic psychology and behaviour*, 52, 40-50, 2018
- LARGE, D. R.; BANKS, V. A.; BURNETT, G.; BAVERSTOCK, S. & SKRYPCHUK, L.: Exploring the Behaviour of Distracted Drivers during Different Levels of Automation in Driving. In: *Proceedings of the 5<sup>th</sup> International Conference on Driver Distraction and Inattention (DDI2017)* (pp. 20-22), 2017
- LARSSON, A. F. L.: Driver usage and understanding of adaptive cruise control. In: *Applied Ergonomics*, 43(3), 501-506, 2012
- LEVIE, W. H. & LENTZ, R.: Effects of text illustrations: A review of research. *ECTAJTRD* 30(4), 195-232, 1982
- LLANERAS, R. E.; SALINGER, J. & GREEN, C. A.: Human Factors Issues Associated with Limited Ability Autonomous Driving Systems: Drivers' Allocation of Visual Attention to the Forward Roadway. In: *Proceedings of the 7<sup>th</sup> International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, June 17-20, 2013, Bolton Landing, New York. Iowa City, IA: Public Policy Center, University of Iowa, 92-98, 2013

- LORENZ, L. & HERGETH, S.: Einfluss der Nebenaufgabe auf die Überwachungsleistung beim teilautomatisierten Fahren. In VDI-Berichte 2264. Paper presented at Fahrer, Fahrerunterstützung und Bedienbarkeit, VDI-Tagung Der Fahrer im 21. Jahrhundert (pp. 159-172). Düsseldorf: VDI-Verlag, 2015
- LOUW, T.; KUO, J.; ROMANO, R.; RADHAKRISHNAN, V.; LENNÉ, M. G. & MERAT, N.: Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. In: *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 870-882, 2019
- MEHLENBACHER, B.; WOGALTER, M. S. & LAUGHERY, K. R. (HG.): On the reading of product owner's manuals: Perceptions and product complexity. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*: SAGE Publications Sage CA: Los Angeles, CA (6), 2002
- MENOLD, N. & BOGNER, K.: Gestaltung von Ratingskalen in Fragebögen. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (SDM Survey Guidelines), 2015
- NASA Space flight human-system standard. Volume 2: Human Factors, habitability and environmental health. NASA Space Flight human-system Standard, 2015
- National Highway Traffic Safety Administration: Federal automated vehicles policy: Accelerating the next revolution in roadway safety: US Department of Transportation, 2016
- National Highway Traffic Safety Administration: Federal Automated Vehicles Policy 2.0. Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT), 2017
- National Transportation Safety Board: Collision Report: Collision Between a Car Operating With Automated Vehicle Control Systems and a Tractor-Semitrailer Truck Near Williston, Florida May 7, 2016 TSB/HAR-17/02 PB2017-102600 Notation 56955 Adopted September 12, 2017
- NAUJOKS, F.; WIEDEMANN, K.; SCHÖMIG, N.; JAROSCH, O. & GOLD, CH. (a): Expert-based controllability assessment of control transitions from automated to manual driving. In: *MethodsX* 5, 579-592, 2018
- NAUJOKS, F.; HERGETH, S.; WIEDEMANN, K.; SCHÖMIG, N. & KEINATH, A. (b): Use cases for assessing, testing, and validating the human machine interface of automated driving systems. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 1873-1877). Sage CA: Los Angeles, CA: SAGE Publications, 2018
- NAUJOKS, F.; HERGETH, S.; WIEDEMANN, K.; SCHÖMIG, N.; FORSTER, Y. & KEINATH, A. (a): Test procedure for evaluating the human-machine interface of vehicles with automated driving systems. In: *Traffic Injury Prevention*, 20(1), 146-151, 2019
- NAUJOKS, F.; HERGETH, S.; WIEDEMANN, K.; SCHÖMIG, N. & KEINATH, A. (b): Development and application of an expert assessment method for evaluating the usability of SAE Level 3 ADS HMIs. *ESV Conference 2019*; Eindhoven, Netherlands, 2019
- NAUJOKS, F.; WIEDEMANN, K.; SCHÖMIG, N.; HERGETH, S. & KEINATH, A. (c): Towards guidelines and verification methods for automated vehicle HMIs. In: *Transportation Research Part F: Traffic Psychology and Behaviour* 60, S. 121-136, 2019
- NAUJOKS, F.; MAI, C. & NEUKUM, A.: The effect of urgency of take-over requests during highly automated driving under distraction conditions. In: *Advances in Human Aspects of Transportation*, 7 (Part I), 431-438, 2014
- NAUJOKS, F.; PURUCKER, C. & NEUKUM, A.: Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. In: *Transportation research part F: traffic psychology and Behaviour*, 38, 67-82, 2016
- NAUJOKS, F.; PURUCKER, C.; WIEDEMANN, K.; NEUKUM, A.; WOLTER, S. & STEIGER, R.: Driving performance at lateral system limits during partially automated driving. In: *Accident Analysis & Prevention* 108, 147-162, 2017
- NEES, M. A.: Drivers' Perceptions of Functionality Implied by Terms Used to Describe Automation in Vehicles. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 1893-1897). Sage CA: Los Angeles, CA: SAGE Publications, 2018

- NEUKUM, A. & KRÜGER, H.-P.: Fahrerreaktionen bei Lenksystemstörungen – Untersuchungsmethodik und Bewertungskriterien. In: VDI-Gesellschaft Fahrzeug- und Verkehrstechnik (Hrsg.), Reifen-Fahrwerk-Fahrbahn, VDI-Berichte Nr. 1791. Düsseldorf: VDI-Verlag, 2003
- NEUKUM, A.; LÜBBEKE, T.; KRÜGER, H.-P.; MAYSER, C. & STEINLE, J.: ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In: M. MAURER & C. STILLER (Hrsg.), 5. Workshop Fahrerassistenzsysteme – FAS2008 (S. 141-150). Karlsruhe: Freundeskreis Mess- und Regelungstechnik e. V., 2008
- NIELSEN, J. & MOLICH, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York: ACM; 249-256, 1990
- NIELSEN, J.: Usability inspection methods. In: CHI 1994: Conference companion on human factors in computing systems, 413-414, 1994
- PRICE, M. A.; VENKATRAMAN, V.; GIBSON, M.; LEE, J. & MUTLU, B.: Psychophysics of trust in vehicle control algorithms (No. 2016-01-0144). SAE Technical Paper, 2016
- PURUCKER, C.; NAUJOKS, F.; PRILL, A. & NEUKUM, A.: Evaluating distraction of in-vehicle information systems while driving by predicting total eyes-off-road times with keystroke level modeling. In: Applied Ergonomics 58, 543-554, 2017
- Response Consortium: Code of Practice for the Design and Evaluation of ADAS. RESPONSE 3: A PReVENT Project, 2006
- ROSS, T.; MIDTLAND, K.; FUCHS, M.; PAUZIÉ, A.; ENGERT, A.; DUNCAN, B. et al.: HARDIE Design Guidelines Handbook: Human factors guidelines for information presentation by ATT systems. DRIVE II Project V2008, Commission of the European Communities, 1996
- SAE on-road automated vehicle standards committee: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (J3016\_201401). SAE International, Warrendale, USA, 2013
- SAE on-road automated vehicle standards committee: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (J3016\_201609). SAE International, Warrendale, USA, 2016
- SAE on-road automated vehicle standards committee: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (J3016\_201806). SAE International, Warrendale, USA, 2018
- SCHALLER, T.; SCHIEHLEN, J. & GRADENEGGER, B. (2008). Stauassistentz – Unterstützung des Fahrers in der Quer- und Längsführung: Systementwicklung und Kundenakzeptanz. In: 3. Tagung Aktive Sicherheit durch Fahrerassistenz
- SCHWARZ, C.; GASPAR, J. & BROWN, T.: The effect of reliability on drivers' trust and behavior in conditional automation. In: Cognition, Technology & Work, 21(1), 41-54, 2019
- STANTON, N. A.; YOUNG, M. S.; WALKER, G. H.; TURNER, H. & RANDLE, S.: Automating the driver's control tasks. In: International Journal of Cognitive Ergonomics, 5(3), 221-236, 2001
- STAPEL, J.; MULLAKKAL-BABU, F. A. & HAPPEE, R.: Driver behavior and workload in an on-road automated vehicle. In: Road Safety & Simulation International Conference (pp. 17-19), 2017
- STEVENS, A. & CYNK, S.: Checklist for the assessment of in-vehicle information systems. 2011
- STEVENS, A.; QUIMBY, A.; BOARD, A.; KERSLOOT, T. & BURNS, P.: Design guidelines for safety of in-vehicle information systems. In: TRL Limited, 2002
- UN ECE R79: Uniform provisions concerning the approval of vehicles with regard to steering equipment, 2017
- Van HUYSDUYNEN, H. H.; TERKEN, J. & EGGEN, B.: Why Disable the Autopilot?. In: Proceedings of the 10<sup>th</sup> International Conference on Automotive User Interfaces and Interactive Vehicular Applications 247-257, 2018
- VICTOR, T. W.; TIVESTEN, E.; GUSTAVSSON, P.; JOHANSSON, J.; SANGBERG, F. & LJUNG AUST, M.: Automation Expectation Mismatch: Incorrect Prediction despite Eyes on Threat and Hands on Wheel. In: Human Factors, 60(8), 1095-1116, 2018

VIKTOROVÁ, L. & ŠUCHA, M.: Learning about advanced driver assistance systems – The case of ACC and FCW in a sample of Czech drivers. In: *Transportation Research Part F: Traffic Psychology and Behaviour*, 65, 576-583, 2018

WIEDEMANN, K.; SCHÖMIG, N.; NAUJOKS, F.; HERGETH, S. & KEINATH, A.: Expert evaluation of automated driving HMI – does a checklist-based method work? Poster presented at the Annual Meeting of the Human Factors and Ergonomics Society Europe Chapter (HFES), Berlin, Germany, 08.10.-10.10.2018

WRIGHT, T. J.; SAMUEL, S.; BOROWSKY, A.; ZILBERSTEIN, S. & FISHER, D. L.: Experienced drivers are quicker to achieve situation awareness than inexperienced drivers in situations of transfer of control within a Level 3 autonomous environment. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 270-273). Sage CA: Los Angeles, CA: SAGE Publications, 2016

WULF, F.; ZEEB, K.; RIMINI-DÖRING, M.; ARNON, M. & GAUTERIN, F.: Effects of human-machine interaction mechanisms on situation awareness in partly automated driving. In: *16<sup>th</sup> International IEEE Conference on Intelligent Transportation Systems* (pp. 2012-2019). IEEE, 2013

ZEEB, K.; BUCHNER, A. & SCHRAUF, M.: What determines the take-over time? An integrated model approach of driver take-over after automated driving. In: *Accident Analysis & Prevention*, 78, 212-221, 2015

ZHOU, H. P.; ITOH, M. & KITAZAKI, S.: Effect of instructing system limitations on the intervening behavior of drivers in partial driving automation. In: *Cognition, Technology & Work*, 1-14, 2019

## Bilder

Bild 1-1: Automatisierungsstufen nach SAE (2016) – Rot markiert ist Level 2, die sog. Teilautomation (partial driving automation)

Bild 1-2: Definition von Prüfkriterien, die zur Bewertung einer erfolgreichen und sicheren Interaktion mit einem teilautomatisierten System herangezogen werden

Bild 2-1: Auszug aus der Checkliste

Bild 2-2: Vergleich der visuellen MMS-Varianten in verschiedenen Systemzuständen (Quelle: WIVW GmbH)

Bild 2-3: Mensch-Maschine-Schnittstelle des Tesla Model 3

Bild 3-1: 7-stufige bipolare Ratingskala zur Beantwortung der Fragen

Bild 3-2: SBFV-Skala zur Bewertung der Situationskritikalität (nach NEUKUM et al., 2008)

Bild 3-3: Ratingskala zur Vergabe eines globalen Versuchsleiter-Urteils zur Bewältigung einer Situation

Bild 3-4: Tablet-Benutzeroberfläche mit Beobachtungs- und Befragungsvariablen (Quelle: WIVW GmbH)

Bild 3-5: Optimale Bildschirm-Anordnung zur Mitschau für den Versuchsleiter während des Versuchs im Simulator (Quelle: WIVW GmbH)

Bild 3-6: VL-Rating zur Bewältigung des Szenarios pro Szenario im Vergleich der MMS-Varianten. Dargestellt sind Mittelwerte und Standardabweichungen. (\*) markiert tendenziell signifikante Unterschiede ( $p < .10$ ), \* markiert signifikante Unterschiede ( $p < .05$ )

Bild 3-7: Visuelle Anzeige des Systemzustands im Kombi-Instrument (links: Screenshot; rechts: implementiert im BAST-Simulator)

Bild 3-8: Bedienlogik und verwendete Lenkradtastenzuordnung im BAST-Simulator

Bild 3-9: Tests im Realverkehr mit Tesla Model 3 (Quelle: WIVW GmbH)

## Tabellen

- Tab. 1-1: Verantwortlichkeiten des Fahrers und der Automation in L2-Systemen gemäß SAE (übersetzt aus SAE, 2018, Seite 21)
- Tab. 1-2: Übersicht aller theoretisch möglichen Use Cases in L2-Systemen
- Tab. 1-3: Übersicht über zu prüfenden Use Cases in der MMS-Bewertung mittels Expertenbewertung und Nutzerstudie
- Tab. 2-1: Übersicht über die Kapitel der Checkliste
- Tab. 2-2: Unterschiedliche Gestaltung der MMS-Varianten basierend auf den Checklisten-Items in der für die Studie verwendeten Version und die tatsächliche Bewertung durch 2 Rater in der vorangegangenen Expertenbewertung (Grün: fulfilled, Gelb: satisfactory, Rot: not-acceptable)
- Tab. 2-3: Gemeinsames Urteil aller Anforderungen auf 3-stufiger Bewertungsskala
- Tab. 3-1: Übersicht über die Empfehlungen zur Vergabe des Versuchsleiter-Ratings basierend auf den pro Szenario beobachteten Fehlern
- Tab. 3-3: Strecke der Prüffahrt mit den einzelnen Szenarien und Streckenlängen
- Tab. 3-4: Versuchsplan der WIVW-Simulatorstudie
- Tab. 3-5: Stichprobencharakteristika der beiden MMS-Gruppen
- Tab. 3-6: Ablauf Prüfstrecke für Studie „Spurhaltung“
- Tab. 4-1: Überblick über geeignete Methoden im Rahmen einer ganzheitlichen Bewertung der Mensch-Maschine-Interaktion



## Anhang

### 1 Kurzinformationen zum Testablauf und Benutzung der Checkliste

#### Brief instructions on how to use the checklist

##### Structure of the checklist:

- 14 requirements for the user manual and human-machine interface, each of which is to be assessed on the basis of different items.
- The structure of the checklist corresponds to the procedure of the application: First the user manual is examined, then the human-machine interface.

##### Application of the checklist:

- Two rater evaluate the system
- **Reading of the user manual**
  - First the user manual is evaluated
  - The user manual is read individually by each rater
- **Experiencing the human-machine interface while driving**
  - The human-machine interface is then experienced by one rater on the basis of specific use cases while driving (activation, deactivation, driving with active system, driver-initiated lane change, system limits; see checklist). Order of use cases in checklist not binding. Re-experiencing the use cases possible at any time.
  - If the test is performed on a real road
    - Road type: motorway or similar road type.
    - Speed: within the range of the system (preferably between 100-120 km/h)
    - Drive mainly on slower lane.
    - Time: Outside the hours of rush hour traffic
  - Driver change if the first rater has completed all use cases.
  - During the ride, the rater refrains as far as possible from expressing opinions in order not to influence the other rater.
- **Rating at item level**
  - After reading the user manual and experiencing the human-machine interface, the raters fill in the checklist items separately.
  - Response format: "yes" / "no" / "not applicable"
  - Zu jedem Item können Kommentare geschrieben werden, die eine Begründung für die Bewertung liefern können
  - For each item comments can be made, which can provide a justification for the evaluation
  - Note 1: The evaluation of the user manual can also be carried out directly after reading it, i.e. before the drive.
  - Note 2: If the test is done in the simulator or on the test track, the item evaluation can be done immediately after a use case.
- **Evaluation of the 14 requirements**
  - Joint discussion and agreement of the raters on rating at the item level
  - After agreement on item level: three-stage judgment on requirement level according to the following rule:
    - „fulfilled“, if all subitems were answered with "yes"

- „satisfactory“ , if max. one sub-item was answered with "no"
  - „not acceptable“, if more than one sub-item was answered with "no"
- **Submission of the final rating for system assessment**
  - Through a joint discussion, the raters decide on one of the four categories for system assessment
  - The assessment takes place at the discretion of the raters.
    - „very good – no improvements necessary“
    - „acceptable with minor options for improvement“
    - „acceptable with major options for improvement“
    - “not acceptable”

## 2 Checklist

### Checklist for the evaluation of human-machine-interface and user manual of Level 2 automated vehicles

---

Date: \_\_\_\_\_

Time: \_\_\_\_\_

Rater: \_\_\_\_\_

Vehicle/System: \_\_\_\_\_

Lane change assistance included:  yes  no

Type of user manual used for evaluation:  printed version  system-integrated version

Final rating:

	very good – no improvements necessary
	acceptable with minor options for improvement
	acceptable with major options for improvement
	not acceptable

Category: User Manual					
Item	Requirements	Fulfilled [yes; no; n/a]			
1	<b>Requirements for the description of system functionality, operational design domain, system limitations and driver responsibility.</b>				
	1.1	System is described and named as an assistance or comfort system which executes parts of the driving task, i.e. both lateral and longitudinal control.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	1.2	Manual describes the conditions under which the system is designed to be used (e.g. roadway, traffic, environmental limitations etc.).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	1.3	Manual describes system limits, and/or reasons for a limited system functionality and their consequences	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	1.4	Manual informs the driver that he/she is responsible for permanently supervising the system in order to intervene in case of system limits or malfunctions.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	1.5	Naming of the function doesn't enhance inappropriate expectations.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
<table border="1" style="margin: auto;"> <tr> <td>Fulfilled</td> <td>Satisfactory</td> <td>Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments:					
<hr/> <hr/>					
2	<b>Requirements for the description of system operation (activation, deactivation, switching between automation levels, lane change).</b>				
	2.1	The manual describes how the system can be activated and which preconditions have to be fulfilled for activation.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	2.2	The manual describes how the system can be deactivated or overruled.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	2.3	The manual describes how the driver can switch between different system modes (e.g. change from L2 to ACC only or vice versa).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	2.4	If implemented, the manual describes how a lane change can be performed, how the system will behave and which responsibilities the driver has.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
	2.5	The manual describes the necessary control elements to execute the above mentioned actions.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a		
<table border="1" style="margin: auto;"> <tr> <td>Fulfilled</td> <td>Satisfactory</td> <td>Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments:					
<hr/> <hr/>					

3	<b>Requirements for the description of different system modes (e.g. L2, ACC only, lateral control only) and its relevant indicators in the user manual.</b>				
	3.1 The manual describes all different system modes and its relevant indicators.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	3.2 Informative pictures are used for illustration.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	3.3 The manual names the relevant displays the driver has to look for (e.g. Cluster display and HUD).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments: <hr/> <hr/>					
4	<b>Requirements for the comprehension of the user manual.</b>				
	4.1 The manual uses simple sentences of moderate length.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	4.2 The manual uses a non-technical language. If technical terms are used, they are explained.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	4.3 Chapters are well structured and separated by headlines.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	4.4 Links to other relevant chapters (e.g. ACC) are made.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments: <hr/> <hr/>					
5	<b>Requirements for a quick way to retrieve the most important information is provided.</b>				
	5.1 A short manual (e.g. video) is available for fast access.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	5.2 If this short manual is available, all information necessary for system use (e.g. section 1 and 2) are included.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	5.3 If this short manual is available, the information is presented in an understandable way.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	5.4 If this short manual is available, references to relevant chapters in the long manual are made.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			

		Fulfilled	Satisfactory	Not-acceptable	
	Comments:				
<b>Category: Specific Use Cases</b>					
<b>6</b>	<b>Requirements for the activation process/activation of the system.</b>				
6.1	Activation of the system out of manual driving is easy to perform (i.e. doesn't require complex driver input e.g. by using a submenu).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
6.2	Activation of the system out of lower automation levels (e.g. ACC active) is easy to perform (i.e. doesn't require complex driver input).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
6.3	The indication of successful activation of the system or parts of the system is given simultaneously to the driver input.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
6.4	A potential delay between system operation and activation or an unsuccessful activation is indicated.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
		Fulfilled	Satisfactory	Not-acceptable	
	Comments:				
<b>7</b>	<b>Requirements for the deactivation process.</b>				
7.1	Deactivation of the full system is easy to perform (i.e. doesn't require complex driver input e.g. using of submenu).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
7.2	Deactivation of single parts of the system (e.g. ACC only) is easy to perform (i.e. doesn't require complex driver input e.g. using of submenu).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
7.3	The indication of successful deactivation of the system or parts of the system is given simultaneously to the driver input.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
		Fulfilled	Satisfactory	Not-acceptable	
	Comments:				
<b>8</b>	<b>Requirements for driving in L2</b>				

8.1	Active system status is permanently displayed.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
8.2	The vehicle limits distractive options for the driver, by limitation of entertainment functions (e.g. internet usage), when using the system.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<b>Instruction: Use system under insufficient preconditions</b>					
8.3	Degraded system functionality is clearly distinguishable from the full system functionality (e.g. lanes not detected).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<b>Instruction: Take eyes off road for about 10 s</b>					
8.4	The system requests the driver to direct visual attention back to the road.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<b>Instruction: Take hands off steering wheel for at least 60 s</b>					
8.5	A notification/warning is issued (hands-off warning).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
8.6	The hands-off warning requests the driver to take back their hands or hold the steering wheel.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
8.7	The design and urgency of the hands-off warning is according to UNR79 regarding the timing, acoustic output and visual elements (first optical warning latest after 15 s, first acoustic warning latest after 30 s, deactivation of system incl. acoustic warning latest after 60 s).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table> <p>Comments:</p> <hr/> <hr/>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
<b>9 Requirements for the assistance in lane change.</b>					
9.1	Activation of the lane change is easy to perform (i.e. doesn't require complex driver input).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
9.2	The indication of successful activation of the lane change is given simultaneously to the driver input.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
9.3	A potential delay between system operation and execution of assistance in lane change or unsuccessful activation is indicated.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
9.4	The indication of an abortion (system- or driver-initiated) or the non-feasibility of the lane change is given simultaneously.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table> <p>Comments:</p> <hr/>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			

10	<b>Requirements in case it is detected that the system or parts of it have stopped/or will stop functioning.</b>				
	10.1 A notification/warning is issued.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	10.2 The notification/warning is triggered long enough to be noticed by the driver.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	10.3 The notification/warning is adequate to the urgency of the situation (in use of colours, presence of additional sound (i.e. multimodality) and content of message).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	10.4 The warning doesn't startle the driver.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments:					
<hr/> <hr/>					
<b>Category: General HMI requirements</b>					
11	<b>Requirements for the display of system mode (e.g. ACC, lateral control, L2).</b>				
	11.1 System mode is indicated at all times.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	11.2 Indicators for different modes (e.g. ACC active, lateral control active; L2 active) are distinguishable from each other.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>			Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
Comments:					
<hr/> <hr/>					
12	<b>Requirements for legibility of visual information.</b>				
	12.1 Visual information is displayed close to driver's line of sight.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.2 Indicators relevant to the system are in spatial proximity of each other in each display.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.3 Colour of indicators contrasts adequately from background colour.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.4 Red/green and blue/yellow colour combinations are avoided.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			



	12.5 Text is easy to read (e.g. size, length).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	12.6 Flashing indicators are used only to attract attention when a warning message is issued.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>		Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
	Comments: _____ _____				
<b>13</b>	<b>Requirements for the comprehension of visual information.</b>				
	13.1 The colours confirm to conventions or stereotypes (for example green related to “system active”; red to “warning or danger”).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.2 The use of different colours is limited to as few as possible (not more than five).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.3 Text messages must be easy to understand (i.e. as short as possible, no technical language, language of user).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.4 Informative semantics is used for non-critical information, command style is used for safety critical information.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	13.5 Indicators are commonly accepted or standardized or are designed in a way that the underlying function can be assumed.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	<table border="1" style="margin: auto;"> <tr> <td style="padding: 5px;">Fulfilled</td> <td style="padding: 5px;">Satisfactory</td> <td style="padding: 5px;">Not-acceptable</td> </tr> </table>		Fulfilled	Satisfactory	Not-acceptable
Fulfilled	Satisfactory	Not-acceptable			
	Comments: _____ _____				
<b>14</b>	<b>Requirements for the control elements.</b>				
	14.1 Controls can easily be reached from normal driving position (e.g. multifunction steering wheel).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	14.2 Controls (e.g. buttons) related to the system are grouped together.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	14.3 Controls are labelled.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	14.4 The labels are visible from the normal driving position and easy to understand (e.g. Res $\hat{=}$ Resume).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	14.5 Operating logic or arrangement of control elements confirms to conventions or stereotypes (e.g. upper button $\hat{=}$ +; lower button $\hat{=}$ -).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			
	14.6 Controls or operational steps for reaching different system modes are distinguishable.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> n/a			

	<table border="1"><tr><td>Fulfilled</td><td>Satisfactory</td><td>Not-acceptable</td></tr></table>	Fulfilled	Satisfactory	Not-acceptable
	Fulfilled	Satisfactory	Not-acceptable	
	Comments:			
<hr/> <hr/>				

## Schriftenreihe

### Berichte der Bundesanstalt für Straßenwesen

#### Unterreihe „Fahrzeugtechnik“

#### 2014

- F 93: **Entwicklung eines Verfahrens zur Erfassung der Fahrerbeanspruchung beim Motorradfahren**  
Buld, Will, Kaussner, Krüger € 17,50
- F 94: **Biokraftstoffe – Fahrzeugtechnische Voraussetzungen und Emissionen**  
Pellmann, Schmidt, Eckhardt, Wagner € 19,50
- F 95: **Taxonomie von Fehlhandlungen bei der Fahrzeugführung**  
Oehme, Kolrep, Person, Byl € 16,50
- F 96: **Auswirkungen alternativer Antriebskonzepte auf die Fahrdynamik von Pkw**  
Schönemann, Henze € 15,50
- F 97: **Matrix von Lösungsvarianten Intelligenter Verkehrssysteme (IVS) im Straßenverkehr Matrix of alternative implementation approaches of Intelligent Transport Systems (ITS) in road traffic**  
Lotz, Herb, Schindhelm, Vierkötter  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.
- F 98: **Absicherungsstrategien für Fahrerassistenzsysteme mit Umfeldwahrnehmung**  
Weitzel, Winner, Peng, Geyer, Lotz Sefati € 16,50
- F 99: **Study on smoke production, development and toxicity in bus fires**  
Hofmann, Dülsen € 16,50

#### 2015

- F 100: **Verhaltensbezogene Kennwerte zeitkritischer Fahrmanöver**  
Powelleit, Muhrer, Vollrath, Henze, Liesner, Pawellek € 17,50
- F 101: **Altersabhängige Anpassung von Menschmodellen für die passive Fahrzeugsicherheit**  
Wagner, Segura, Mühlbauer, Fuchs, Peldschus, Freßmann € 19,00
- F 102: **6th International Conference on ESAR „Expert Symposium on Accident Research“**  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.
- F 103: **Technische Möglichkeiten für die Reduktion der CO<sub>2</sub>-Emissionen von Nutzfahrzeugen**  
Süßmann, Lienkamp  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.
- F 104: **Abbiege-Assistenzsystem für Lkw – Grundlagen eines Testverfahrens**  
Schreck, Seiniger € 14,50
- F 105: **Abgasverhalten von in Betrieb befindlichen Fahrzeugen und emissionsrelevanten Bauteilen – Feldüberwachung**  
Schmidt, Georges € 14,50
- F 105b: **Examination of pollutants emitted by vehicles in operation and of emission relevant components – In-service conformity**  
Schmidt, Johannsen  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

F 106: **Untersuchung des Abgasverhaltens von in Betrieb befindlichen Fahrzeugen und emissionsrelevanten Bauteilen – Austauschкатаlysatoren**  
Schmidt, Johannsen € 13,50

F 106b: **Examination of pollutants emitted by vehicles in operation and of emission relevant components – Replacement catalytic converters**  
Schmidt, Johannsen  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

F 107: **Sicherheitsaspekte beim Laden von Elektrofahrzeugen**  
Vogt, Link, Ritzinger, Ablingyte, Reindl € 16,50

F 108: **Interoperabilität zwischen öffentlichem Verkehrsmanagement und individuellen Navigationsdiensten – Maßnahmen zur Gewährleistung**  
von der Ruhren, Kirschfink, Ansorge, Reusswig, Riegelhuth, Karina-Wedrich, Schopf, Sparmann, Wöbbeking, Kannenberg € 17,50

F 109: **Ermittlung des Umfangs von Abweichungen bei Durchführung der Abgasuntersuchung zwischen Messung am Auspuff und Abfrage des On-Board-Diagnosesystems**  
Schröder, Steickert, Walther, Ranftl  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

F 110: **Wahrnehmung und Bewertung von Fahrzeugaußengeräuschen durch Fußgänger in verschiedenen Verkehrssituationen und unterschiedlichen Betriebszuständen**  
Altinsoy, Landgraf, Rosenkranz, Lachmann, Hagen, Schulze, Schlag  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

F 111: **Geräuschkinderung von Dünnschichtbelägen**  
Schulze, Kluth, Ruhnau, Hübelt  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

#### 2016

F 112: **Ersatz von Außenspiegeln durch Kamera-Monitor-Systeme bei Pkw und Lkw**  
Schmidt, Hoffmann, Krautscheid, Bierbach, Frey, Gail, Lotz-Keens € 17,50

F 112b: **Final Report Camera-Monitor-Systems as a Replacement for Exterior Mirrors in Cars and Trucks**  
Schmidt, Hoffmann, Krautscheid, Bierbach, Frey, Gail, Lotz-Keens  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

F 113: **Erweiterung der Software TREMOD um zukünftige Fahrzeugkonzepte, Antriebe und Kraftstoffe**  
Bergk, Heidt, Knörr, Keller € 15,50

F 114: **Barrierefreiheit bei Fernlinienbussen**  
Oehme, Berberich, Maier, Böhm € 17,50

F 115: **Statischer und dynamischer Fahrsimulator im Vergleich – Wahrnehmung von Abstand und Geschwindigkeit**  
Frey  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

## 2017

### F 116: Lang-Lkw – Auswirkung auf Fahrzeugsicherheit und Umwelt

Süßmann, Förg, Wenzelis  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 117: 7th International Conference on ESAR „Expert Symposium on Accident Research“ – Reports on the ESAR-Conference 2016 at Hannover Medical School

Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 118: Bedeutung kompensativer Fahrerstrategien im Kontext automatisierter Fahrfunktionen

Voß, Schwalm € 16,50

### F 119: Fahrzeugtechnische Eigenschaften von Lang-Lkw

Förg, Süßmann, Wenzelis, Schmeiler  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 120: Emissionen von über 30 Jahre alten Fahrzeugen

Steven, Schulte, Hammer, Lessmann, Pomsel  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 121: Laufleistungsabhängige Veränderungen der CO<sub>2</sub>-Emissionen von neuen Pkw

Pellmann, Schmidt  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

## 2018

### F 122: Revision der Emissionsmodellierung für leichte Nutzfahrzeuge – Bedarfsanalyse auf Basis einer Vorstudie

Auf der Maur, Strassburg, Knörr, Heidt, Wuethrich  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 123: Motorradschutzhelme – Identifizierung ihres Verbesserungspotenzials unter Berücksichtigung des Motorradunfallgeschehens

Pollak, Schueler, Bourdet, Deck, Willinger € 19,50

### F 124: Aufbau eines Qualitätsmanagementsystems für die Erfassung und Weiterverarbeitung von Daten für IVS-Dienste

Heinrich, Pollesch, Schober, Stamatakis, Grzebellus, Radike, Schneider, Stapelfeld, Huber  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 125: Untersuchung zu Elektrokleinstfahrzeugen

Bierbach, Adolph, Frey, Kollmus, Bartels, Hoffmann, Halbach € 19,50

## 2019

### F 126: Einfluss zunehmender Fahrzeugautomatisierung auf Fahrkompetenz und Fahrkompetenzerwerb

Weißgerber, Grattenthaler, Hoffmann € 15,50

### F 127: Erhöhung der Verkehrssicherheit älterer Kraftfahrer durch Verbesserung ihrer visuellen Aufmerksamkeit mittels „Sehfeldassistent“

Kupschick, Bürglen, Jürgensohn € 16,50

### F 128: Potenzieller gesellschaftlicher Nutzen durch zunehmende Fahrzeugautomatisierung

Rösener, Sauerbier, Zlocki, Eckstein, Hennecke, Kemper, Oeser  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 129: Anforderungen an die dynamische Leuchtweitenregelung zur Vermeidung der Blendung entgegenkommender Verkehrsteilnehmer

Kosmas, Kobbert, Khanh € 15,50

### F 130: Infrastrukturbedarf automatisierten Fahrens – Grundlagenprojekt

Dierkes, Friedrich, Heinrich, Hoffmann, Maurer, Reschka, Schendzielorz, Ungureanu, Vogt  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 131: Fahrerassistenz- und Fahrerinformationssysteme (FAS/FIS) – Personale Voraussetzungen ihres Erwerbs und Nutzung durch ältere Kraftfahrerinnen und -fahrer

Hargutt, Kenntner-Mabiala, Kaussner, Neukum  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

## 2020

### F 132: Handbuch Barrierefreiheit im Fernbuslinienverkehr

Boenke, Grossmann, Nass, Schäfer € 17,50

### F 133: Lkw-Notbremsassistenzsysteme

Seiniger, Heint, Bühne, Gail € 15,50

### F 134: Stationär-Geräusch von elektrisch angetriebenen Fahrzeugen

Altinsoy, Lachmann, Rosenkranz, Steinbach € 19,00

### F 135: Abweichungen von der akzeptierten Fahrleistungsschwelle in automatisierten Fahrsituationen

Voß, Schwalm € 18,00

## 2021

### F 136: Kamera-Monitor-Systeme als Fahrerinformationsquelle

Leitner, Oehme, de Silva, Blum, Berberich, Böhm  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 137: Konzept für die Erzeugung eines ISO-konformen UML-Modells und Generierung eines GML-Applikationsschemas für DATEX II zur Verbesserung der Interoperabilität

Lauber, Steiger, Kopka, Lapolla, Freudenstein, Kaltwasser  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 138: Grundlagen zur Kommunikation zwischen automatisierten Kraftfahrzeugen und Verkehrsteilnehmern

Schaarschmidt, Yen, Bosch, Zwickel, Schade, Petzold € 16,50

### F 139: Einfluss von Notbremssystemen auf die Entwicklung von Lkw-Auffahrunfällen auf Bundesautobahnen

Straßgütl, Sander € 14,50

### F 140: Reibwertprognose als Assistenzsystem

Leschik, Sieron, Gregull, Müller, Trapp, Brandenburg, Haalman, Terpstra  
Dieser Bericht liegt nur in digitaler Form vor und kann unter <https://bast.opus.hbz-nrw.de/> heruntergeladen werden.

### F 141: Methoden für die Bewertung der Mensch-Maschine-Interaktion beim teilautomatisierten Fahren

Schömig, Wiedemann, Julier, Neukum, Wiggerich, Hoffmann € 18,00

---

Fachverlag NW in der Carl Ed. Schünemann KG  
Zweite Schlachtpforte 7 · 28195 Bremen  
Tel.+(0)421/3 69 03-53 · Fax +(0)421/3 69 03-48

Alternativ können Sie alle lieferbaren Titel auch auf unserer Website finden und bestellen.

[www.schuenemann-verlag.de](http://www.schuenemann-verlag.de)

Alle Berichte, die nur in digitaler Form erscheinen, können wir auf Wunsch als »Book on Demand« für Sie herstellen.