VDA Verband der Automobilindustrie

FAT Forschungsvereinigung Automobiltechnik

FAT-Schriftenreihe 369

Level 2 hands-off–Recommendations and guidance

Level 2 hands-off
–
Recommendations and guidance

Report 216300

fka GmbH

Automated Driving

Final Report

Level 2 hands-off

–

Recommendations and guidance

Project Number

216300

Customer

Verband der Automobilindustrie e.V. (VDA)
z.Hd. Philipp Niermann
Behrenstr. 35
10117 Berlin
Deutschland

Contractor:	**Project Manager:**
fka GmbH	Dr. phil. Johanna Josten Dipl.-Ing. Philipp Seewald
Subcontractors:	
Institute for Automotive Engineering (ika), RWTH Aachen University	Univ.-Prof. Dr.-Ing. Lutz Eckstein
Chair of Ergonomics, Technical University of Munich	Prof. Dr. phil. Klaus Bengler
fka SV Inc.	

Aachen, January 2023

Preamble

This report on the project Level 2 hands-off (L2H-off), commissioned by VDA to fka GmbH with the three subcontractors Institute for Automotive Engineering (ika) of RWTH Aachen University, Chair of Ergonomics (LfE) of the Technical University of Munich and fka Silicon Valley Inc., has been composed with the intention to allow VDA and its members as well as other involved parties to recall assumptions and restrictions on the project's scope as agreed upon during joint discussions and to document the work towards VDA. This report does not have the intention to guide non-involved parties through each decision made, since presentations and discussions during the project phase are not included. It rather complements the milestone-based documentation of the work conducted.

Notwithstanding that one contracting party took the lead for each work package within the project, most work packages were conducted in close cooperation between the (sub)contractors. Many decisions on, e.g., data logging and analysis formats or research questions were agreed upon univocally for all work packages. Named as the authors of a chapter are those who provided the documentation for the corresponding work package.

Executive Summary

Documentation by J. Josten (fka GmbH)

While feet off pedals is a common behavior during the use of longitudinal driver assistance such as Adaptive Cruise Control (ACC) as well as during L2 use (L2H-on functions), hands-free supervision during use of lateral driver assistance is currently not admissible in Germany. First series-production vehicles with functions allowing for feet- and hands-free use of Level 2 assistance (L2H-off functions) are, however, available in other markets (e.g., USA). These functions comprise a glance-based driver monitoring system complementing the hands-on detection (i.e., L2 functions with eyes-on requirement). The question arises which technical and ergonomic requirements would need to be fulfilled to avoid or compensate any adverse foreseeable driver behavior in combination with hands-free L2 supervision.

The project Level 2 hands-off (L2H-off), assigned by VDA, was conducted between August 2021 and November 2022 by fka GmbH and three subcontractors, the Institute for Automotive Engineering (ika) of RWTH Aachen University, the Chair of Ergonomics (LfE) of the Technical University of Munich and fka SV Inc., based in Santa Clara, CA. The project's objective was to generate a reliable set of data, information and knowledge by application of different methods, aiming to assess five potential challenges related to interaction behavior during L2 hands-free driving as well as to derive recommendations on how these challenges shall, should or may be compensated. The five challenges and questions (CQ) in focus of the work conducted were:

- **CQ1: Hands-off = mind-off?**
  There are concerns that a lesser driver involvement in the driving task, due to the hands-free use of L2 assistance, will reduce the driver's involvement in the driving task, especially in terms of visual attention.
- **CQ2: Risks through prolonged transition times**
  There are concerns that the process of returning the hand(s) to the steering wheel (i.e., a physical disadvantage of hands-free monitoring) as well as longer reaction times in general lead to an increased risk of accidents.
- **CQ3: Foreseeable misuse**
  There are concerns that the use of L2H-off functions will lead to increased foreseeable misuse, particularly with respect to an increased engagement in non-driving related tasks.
- **CQ4: Mode confusion**
  There are concerns that with the introduction of L2H-off functions, drivers are no longer aware of their tasks or role as well as of the function's limitations.
- **CQ5: Safety level**
  There is uncertainty as to what level of safety (in terms of contributions to road safety) can be achieved by introducing L2H-off.

As a first step, existing knowledge was aggregated as a basis on which to build data collections and analyses according to the scope of the project. This included a theoretical as well as international field data based and expert on-road overview on existing series-production and prototypical L2 functions as well as their defined operational design domains (ODD). Furthermore, existing input to the five challenges was aggregated as an outcome of an extensive literature search. This knowledge basis allowed for the transformation of the state of the art into design hypotheses to be tested in own data collections in the field (e.g., field operational test in the Munich area), including series- and close-to-production L2H-off functions, as well as in four driving simulator studies (in Germany and the USA) and in an online survey targeting experienced users of L2 functions. These data collections allowed the comparison of users' behavior in interaction with (prototypical) design solutions when monitoring users' gaze behavior (L2H-off functions) as opposed to monitoring users' hand position (L2H-on functions).

A first transformation of the collected data targeted analyses of user behavior to derive a conclusion for each of the five challenges. Central to conclusions regarding L2H-off functions were changes in interaction behavior as observed for state of the art L2H-on functions:

- **CQ1: Hands-off = mind-off?**
  Visual attention to the road during L2H-off use, as an indicator of mind-off, appeared to be similar or improved in comparison to the use of L2H-on functions. Drivers further adapt their level of motoric control during L2 use.
- **CQ2: Risks through prolonged transition times**
  The option for hands-free driving does not translate into prolonged intervention times at functional limits. The physical disadvantage of hands-free driving can be compensated by supporting a sufficient involvement in the driving task by glance-based driver monitoring systems.
- **CQ3: Foreseeable misuse**
  The potential for misuse seems closely related to the DMS design. Given a glance-based driver monitoring system, data suggest that foreseeable misuse does not increase by the option to monitor hands-free alone.
- **CQ4: Mode confusion**
  Hands-free monitoring does not increase mode confusion in comparison to L2H-on functions when providing prior information on driver role and system functioning. Misconceptions of HMI signals can prevent successful driver interventions.
- **CQ5: Safety level**
  A similar interaction quality with L2H-off and L2H-on functions was found in terms of criticality metrics and perceived safety.

In a second transformation of results, conclusions and recommendations with regard to the design of L2H-off functions to address potential challenges were derived. Design guidance, including general best practice advice exceeding the research focus of this project, has been aggregated in a separate chapter of this report. Similar to the European Statement of Principles on human machine interface for in-vehicle information and communication systems (ESoP), the aggregated design principles provide a structured starting point for the design of L2H-off driver assistance systems and their necessary components. Design guidance addresses general design assumptions as well as the necessary components of L2H-off functions such as a driver monitoring and driver information and warning system, a risk mitigation function, an ODD monitoring functionality and a collision mitigation system. Key findings of the project have been presented in the 17th Taskforce ADAS in January 2023 and hopefully advance further regulatory discussions on Level 2 driver assistance.

Table of Contents

1 Introduction and Overview

Documentation by J. Josten (fka GmbH)

Following a common definition of Level 2 assistance (SAE J3016, 2018), steering and brake/acceleration support is provided by the function, while the driver remains responsible for the driving task at all times. Responsible means that the driver is always in control of the driving task, i.e., is required to supervise the function constantly and to actively intervene whenever necessary (i.e., to take direct control). While feet off pedals is a common behavior during the use of longitudinal driver assistance such as Adaptive Cruise Control (ACC) as well as during L2 use (L2H-on functions), hands-free supervision is currently not admissible during use of lateral driver assistance in Germany. First series-production vehicles with functions allowing for feet- and hands-free use of Level 2 assistance (L2H-off functions) are, however, available in other markets (e.g., USA). These functions comprise a glance-based driver monitoring system complementing the hands-on detection (i.e., L2 functions with eyes-on requirement).

The question arises which technical and ergonomic requirements would need to be fulfilled to avoid or compensate any adverse foreseeable driver behavior in combination with hands-free L2 supervision. The goal of this project was thus to generate a reliable set of data, information and knowledge by application of different methods, aiming to assess potential challenges related to L2 hands-free driving as well as to derive recommendations on how these shall, should or may be compensated. The five potential challenges and questions (CQs) that have been proposed by an expert group of VDA members based on VDA-internal and -external discussions as motivation for this project were:

- **CQ1: Hands-off = mind-off?**
  There are concerns that a lesser driver involvement in the driving task, due to the hands-free use of L2 assistance, will reduce the driver's involvement in the driving task, especially in terms of visual attention.
- **CQ2: Risks through prolonged transition times**
  There are concerns that the process of returning the hand(s) to the steering wheel (i.e., a physical disadvantage of hands-free monitoring) as well as longer reaction times in general lead to an increased risk of accidents.
- **CQ3: Foreseeable misuse**
  There are concerns that the use of L2H-off functions will lead to increased foreseeable misuse, particularly with respect to an increased engagement in non-driving related tasks.
- **CQ4: Mode confusion**
  There are concerns that with the introduction of L2H-off functions, drivers are no longer aware of their tasks or role as well as of the function's limitations.
- **CQ5: Safety level**
  There is uncertainty as to what level of safety (in terms of contributions to road safety) can be achieved by introducing L2H-off.

More specifically, this project considered the five potential challenges during normal L2 operation within a defined operational design domain (ODD) as well as during control transitions between assistance modes (driver- or system-initiated) with a special focus on the design of driver monitoring systems (DMS) for L2H-off functions. Certain aspects were not in the primary focus of experimental data collections or comparisons between L2 designs, but were nonetheless discussed and defined in this project in terms of minimum or best practice requirements. Examples are the design of driver information and warning systems (DIW) or the operational design domain (ODD) for the selection of relevant scenarios for the analysis of driver behavior. Furthermore, requirements for edge cases of interaction, such as medical emergencies or differences introduced by the choice of technological solution (influencing, e.g., failure types or the quality of driver state detection), were not in the focus of this project. Consequently, not all aspects that potentially need to be considered for a comprehensive design of L2H-off functions could be addressed by the project at hand.

1.1 About this project

The project was conducted between August 2021 and November 2022. For a comprehensive consideration of the challenges (see above), five subprojects were designed, each with a specific methodological focus (see Figure 1-1).

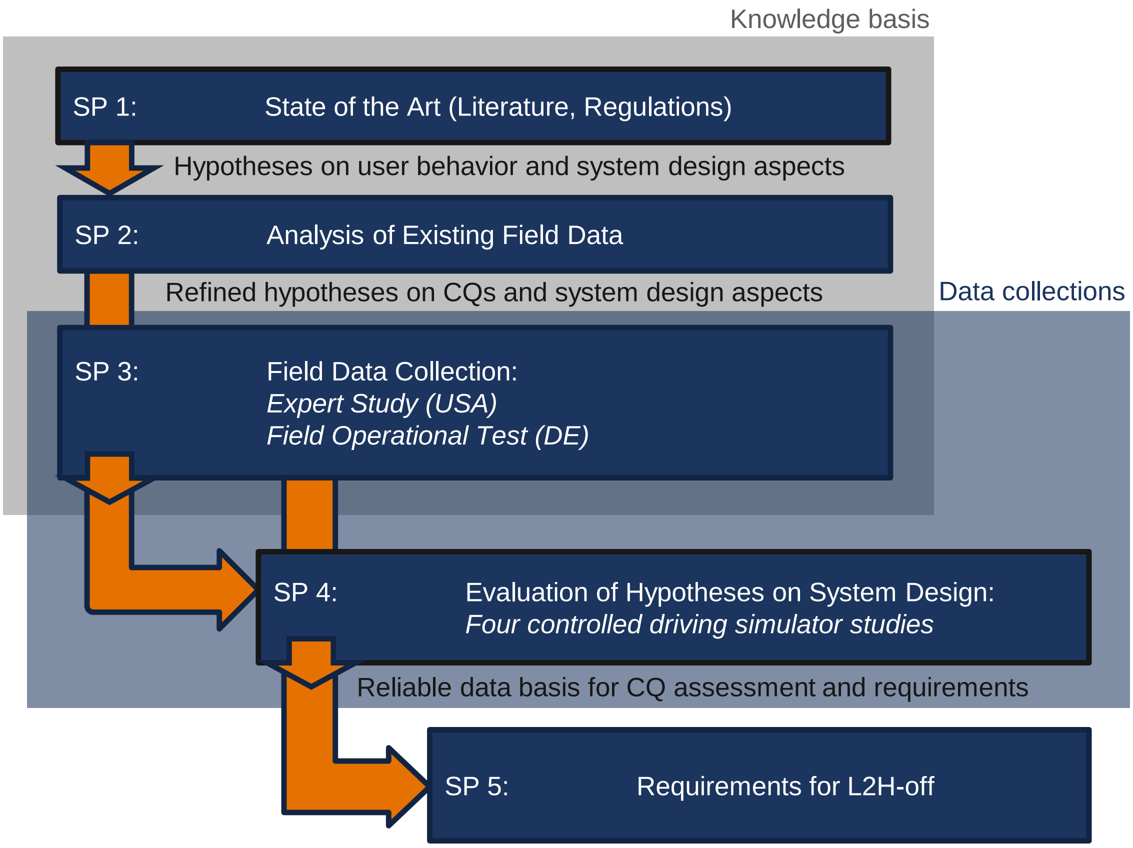


Figure 1-1: The project's five subprojects (SP), their purpose within this project and their interrelations (CQ = Challenge / question).

The first three subprojects build the basis on which design hypotheses for investigation in controlled experimental studies (SP4; Section 5) were formulated. This decision basis consisted of a literature overview (SP1, Section 2; scientific literature, normative references and other publications), an analysis of international field data sets provided for analysis in this project to review driver behavior in interaction with (existing) L2H-off and L2H-on functions in the field (SP2; Section 3), as well as first project-related data collections in the field (SP3; Section 4) such as expert and user assessments of existing (US) functions. A connecting link between studies comparing specific L2 designs in controlled environments and an analysis of user behavior as observed with state of the art L2 functions was the field operational test (FOT; Section 4.4). The FOT was conducted with series-production L2H-on and prototypical L2H-off functions in Germany.

Based on the overview of the state of the art and guided by the five questions behind this project, four controlled simulator studies were designed (SP4; Section 5), each investigating a specific design aspect or motivated by one or more of the five challenges (see Section 5.1 for an overview of relevant aspects for investigation). An overview of all data collections, based on the analysis of the state of the art, is given by Figure 1-2.

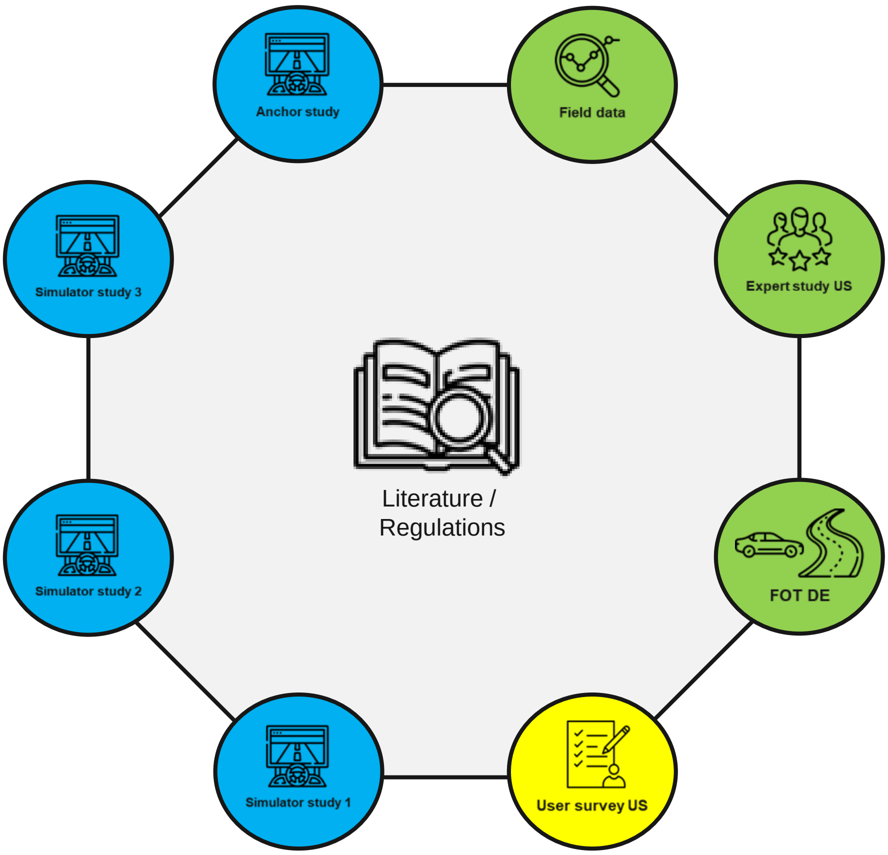


Figure 1-2: Overview on the data collections and data analyses conducted within this project, defined and based on an overview of the state of the art including literature (gray). Colors indicate the method applied: Green = field data, yellow = online survey, blue = controlled simulator studies. (*Figure provided by N. Grabbe, Lehrstuhl für Ergonomie, TU München, icon source: Flaticon.com*)

Study 1 (Section 5.2) conducted a general comparison between driver monitoring systems with a focus on hands-off wheel or eyes-off road detection. Study 2 (Section 5.3) focused on mode confusion as a result of different L2 designs. Study 3 (Section 5.4) investigated the timing of alerts based on the detection of non-compliant driver behavior. The fourth study (Anchor study; Section 5.5) investigated the generalizability of assumptions on driver behavior over different samples as well as the handling of planned and unplanned events.

Two transformation steps were central to achieving the project's goals: Firstly, the results and review of all analyses had to be aggregated and interpreted with regard to the five potential challenges for hands-free monitoring motivating the project (see above). In a second step, the conclusions drawn on driver behavior in interaction with L2 functions and the design options discussed needed to be transformed into technology-independent guidance on the design of L2H-off functions. The final SP5 summarizes the efforts described above by providing an overall assessment for each of the five challenges based on the state of the art, discussions lead within this project and, first and foremost, the results of all data collections (first transformation; Section 6.1). The second transformation of knowledge generated within this project was aimed at providing guidance on the design of L2H-off functions (Section 6.2). As described in more detail in the according chapter, the guidance provided has to be considered within the determining conditions set for the analysis of L2 use within this project.

To allow for an aggregation of data over different data collections and analyses in this project, metrics and procedures have been aligned where possible (e.g. questionnaires, metrics for visual attention; see Section 2.4 and Section 5.1). Another point of alignment concerned the reference functions or driving modes for comparison of user behavior when interacting with specific L2H-off functions. This includes the selection and design of relevant interaction scenarios within the investigated ODD, such as necessary transitions of control at functional limits with and without notification by the function. As prototypical L2H-off functions had to be defined for investigation of specific design aspects in the simulator studies (Section 5), basic requirements for the design of HMI, mode transitions and DMS criteria were established under consideration of the state of the art overview conducted in subprojects 1 to 3.

1.2 About this document

This document's main objective is rather completeness than conciseness as it comprises the documentation of all relevant works conducted within the L2H-off project in a chronological order of the subprojects and their work packages (see above). Documentation includes those work packages that focused on establishing a common understanding of wording or the selection of research foci (e.g., Section 2.4 and Section 2.5). As the different data collections and foci of this project were not addressed consecutively but partly in parallel, some assumptions and conclusions drawn at the end of the first subproject already include insights collected during subsequent SPs (e.g. from the expert assessment, see Section 4.2). Not all aspects that have been discussed at some point during the course of this project could be investigated within the four experimental studies in SP 4.

Each data collection or review focus within this project is documented within a separate chapter. For each data collection or analysis conducted, the specific focus of investigations, applied methods and results will be described and discussed. The focus of interpretations of results for each data collection/analysis is on the five challenges and questions presented as the motivation for this project. The overall goal is to aggregate findings of single investigations, methods and studies into one general conclusive picture for each challenge. Subsequently, design principles are derived, motivated by conclusions drawn from discussions, reviews and data analyses of this project.

2 State of the Art

The following chapters describe the procedures applied to generate a knowledge basis as a starting point for further investigations and discussions within the project (see Figure 2-1). Subproject 1 included a research of published (scientific) literature on L2 (hands-free) monitoring (Chapter 2.1). An overview on existing functions and regulatory aspects is provided in Chapter 2.2. Chapter 2.3 focusses on the analysis of potential risks inherent to the use of L2H-off functions as well as relevant scenarios for investigation. Chapter 2.4 provides a first overview on metrics to be investigated with regard to the five challenges and questions on driver behavior. Finally, Chapter 2.5 concludes the state of the art overview with the selection for potentially relevant design aspects.

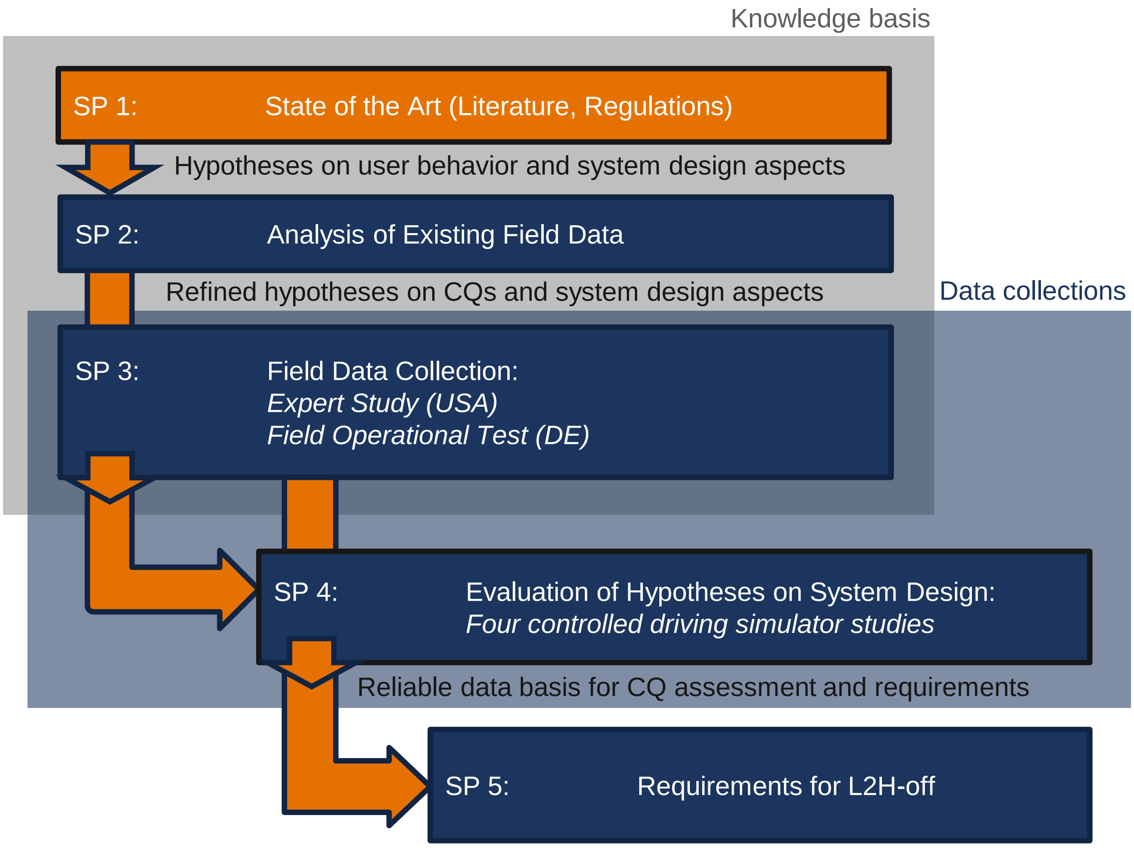


Figure 2-1: Overview on the five subprojects and the role of SP1 within the project.

2.1 Literature Overview

Documentation by T. Hecht, D. Albers, B. Biebl, A. Feierle, M. Hübner, N. Grabbe, K. Bengler (Lehrstuhl für Ergonomie, TU München)

Within this section, the aim was to create an overview of existing studies, accident and experience reports on driver behavior in Level 2 Hands-on (L2H-on) and Level 2 Hands-off (L2H-off). Furthermore, literature on driver monitoring systems and attention reminders has been reviewed. A structured review of existing literature was delivered in order to create a first set of hypotheses. This section started at the project beginning (09/21) and finished three months later. Its output served as a starting point for further work packages.

Relevant literature was scanned and systematically reviewed. Based on that, key findings and existing research gaps were derived. Input sources for the review were:

- FAT/WIVW database
- Literature from VDA and partners
- Google scholar search
- Experience & accident reports from forum discussions, news reports, and online videos

Before the research began, relevant key words per challenge and question were derived (see Table 2-1).

Table 2-1: Overview on Challenges and Questions (CQs) and associated keywords used for the literature review

Challenges and Questions	Associated Keywords
1: Hands-off = mind-off?	• Attention (visual and cognitive) • Mind wandering
2: Prolonged transition times	• Take-over performance • Take-over quality • Take-over readiness
3: Foreseeable misuse	• Non-driving related activity engagement • Visual and physical aversion • Overtrust/over-reliance • Intentional violations • "Tesla Orange" → bypassing driver monitoring mechanisms
4: Mode confusion	• Mode awareness • Situation awareness • Mental model
5: Safety level	• Global traffic safety level • Controllability • Potential safety gain for existing systems, e.g., through Driver Monitoring, improved HMI concepts

This literature review resulted in 58 articles and included journals, conference papers, doctoral theses, and reports. Out of the 58 articles, 22 highly relevant and comparable studies were identified. Studies were rated as comparable when both the scenario (highway or similar) and the automated driving function were within the focus of this review. Furthermore, 37 experience and accident reports were scanned and reported.

In the following, the results are structured in behavioral effects during regular operation and behavioral effects in transitions and critical situations. Both observed behavior and approaches to explain this behavior are reported.

2.1.1 Behavioral effects during regular operation

When comparing a reliable L2H-off system without a driver monitoring system to manual driving, several effects were found in previous studies: Lower attention ratios to the road ahead (Boos et al., 2020) and longer eyes-off path glances (e.g., on in-vehicle displays) (Kraft et al., 2018; Victor et al., 2018) were reported in L2H-off without a driver monitoring system. Furthermore, a higher likelihood of Non-Driving Related Activity (NDRA) engagement was indicated (Noble et al., 2021). Moreover, a study comparing a reliable L2H-off system without driver monitoring to Adaptive Cruise Control (ACC) was identified (Llaneras et al., 2013). The authors showed that L2H-off without driver monitoring leads to lower attention to the road ahead and longer eyes-off path glances, and a higher likelihood of NDRA engagement. Also, previous study results indicate a lower share of glances to the road ahead (Josten, 2021; Othersen, 2016) when driving with L2H-off without driver monitoring compared to with L2H-on. In general, the results indicate that the higher the level of automation, the lower the share of glances to the road ahead (Josten, 2021; Othersen, 2016).

Furthermore, the studies' results suggest that the longer people use a reliable L2H-off system without driver monitoring, the lower the share of glances to the road ahead (Boos et al., 2020; Llaneras et al., 2013; Reagan et al., 2021; Victor et al., 2018), the higher the likelihood of NDRA engagement (Boos et al., 2020; Llaneras et al., 2013; Reagan et al., 2021; Victor et al., 2018), and the higher the likelihood that both hands are taken off the steering wheel (Reagan et al., 2021).

In the literature, several approaches to explaining the observed behavioral effects were discussed. In Boos et al. (2020), subjects stated that NDRA engagement served as a countermeasure for sleepiness, that it was more attractive than the monitoring task, and that they highly relied in the L2 automation (i.e., over-reliance). Mode confusion was not a main reason for the observed behavior (Boos et al., 2020). Feldhütter et al. (2019) and Kim et al. (2021) claim a positive attitude towards automated driving to be a factor influencing users' likelihood of NDRA engagement. Furthermore, the hand position (hands on vs. hands off the steering wheel) was found to influence drivers' mental model (users perceives himself/herself as driver vs. as passenger) and in turn influence drivers' behavior (Cahour et al., 2021). In general, there is less literature explaining drivers' behavior than there is observing the behavior (see also Reagan et al. (2021)).

2.1.2 Behavioral effects in transitions and critical situations

In transitions with Requests to Intervene (RtIs), a reliable L2H-off system without driver monitoring was found to cause (slightly) increased take-over times (Cahour et al., 2021; Garbacik et al., 2021; Gold et al., 2013; Josten, 2021; Othersen, 2016) and slightly impaired take-over quality (Cahour et al., 2021; Garbacik et al., 2021; Ishida & Itoh, 2017) compared to L2H-on.

The motoric intervention can explain the delayed take-over times: it takes approximately 0.3 seconds to take the hands to the steering wheel (Damböck et al., 2013; Gold et al., 2013; Josten, 2021).

In critical situations without RtIs, a reliable L2H-off system with driver monitoring and attention reminders based on driver gaze behavior led to similar results in crash rate (Victor et al., 2018), surprise reaction timepoint, and driver steering timepoint (but more braking maneuvers) compared to L2H-on (Pipkorn et al., 2021). Gustavsson et al. (2018), Pipkorn et al. (2021), and Victor et al. (2018) found that driver reactions in system malfunction scenario does not change with hands-on condition, but with trust level: Crashers were aware of the threat when it emerged but felt safe due to the good driving performance of the vehicle automation; all crashers reported high trust in the vehicle to handle the conflict situation. This phenomenon was called automation expectation mismatch by Victor et al. (2018).

Tivesten et al. (2019) further investigated behavioral patterns that were associated with increased crash risks in malfunction scenarios and concluded that low levels of visual attention to the road ahead, high percent road center, and long visual reaction times to attention reminders can help predict the outcome of a malfunction scenario.

2.1.3 System design approaches

In addition to insights into driver behavior during regular operations and in transitional and critical situations, the literature review also revealed countermeasures when potentially dangerous driver behavior was identified. In four different studies tackling this field, attention reminders were found to be multimodal, typically featuring several warning cascades, and were based on eye-tracking (Blanco et al., 2015; Kurpiers et al., 2019; Llaneras et al., 2017; Victor et al., 2018). The eyes-off road time appeared to be a common input metric for driver monitoring systems. Different times were used for the first stage of the warning cascade: While Kurpiers et al. (2019) used 4 s eyes-off road time, Blanco et al. (2015) compared 2 s and 7 s, and Llaneras et al. (2017) started the first warning after 6 s. Victor et al. (2018) used a combination of multiple eye-tracking metrics to trigger their three-step cascade.

All studies revealed positive effects of the investigated three stage attention requests: improved monitoring behavior including less >4 s eyes-off-path glances (Blanco et al., 2015; Kurpiers et al., 2019; Llaneras et al., 2017; Victor et al., 2018), a positive influence on trust level and mode awareness (Kurpiers et al., 2019), and improved reactions to silent system malfunctions (Llaneras et al., 2017). However, Blanco et al. (2015) revealed that attention reminders based on 2 s visual inattention lead to habituation effects and the participants start to ignore the prompts. Victor et al. (2018) emphasize that visual attention alone does not guarantee a reliable driver reaction. Moreover, reviewed experience reports suggest that driver state monitoring systems (both camera and steering wheel based) can be bypassed (Alvarez, 2021; Bindley & Elliott, 2021).

Furthermore, the question “What actions should be taken if driver behavior does not cease?” was investigated. No studies were found that specifically address different degrees of system degradation and their effects on, e.g., acceptance and traffic safety. Therefore, reference is made to Llaneras et al. (2017): Warning stage three issues a “speech-based message with additional consequences, including disengaging the system, which causes the vehicle to coast, and locking-out system reactivation (drivers must now recycle the ignition to engage the partially automated feature). (Llaneras et al., 2017, p. 3)” Furthermore, reference is made to the analysis of existing systems in Section 2.2 and to studies on minimal risk maneuvers in the context of higher levels of automation (e.g., Karakaya and Bengler (2021)).

2.1.4 Overview/Summary

This section provides short summaries of the above-described literature separated by CQs.

2.1.4.1 CQ1: Hands-off = minds-off?

A reliable L2H-off system without a driver monitoring system and attention reminders leads to increased visual distraction compared to H-on, ACC, and manual driving. No insights on mind wandering, e.g., using mind wandering questionnaires or visual cues, were gathered from the literature review. Participants were found to cause an accident with their eyes on the conflict object, thus visual attention was rated insufficient to guarantee successful driver action. The term automation expectation mismatch was used in a series of publications to describe this phenomenon.

2.1.4.2 CQ2: Risk through prolonged transition times

A reliable L2H-off system without a driver monitoring system and attention reminders causes (slightly) prolonged take-over times and degraded take-over quality in scenarios with RtIs compared to L2H-on. Necessary motoric movements, decoupling from the steering wheel, and over-reliance were mentioned as reasons for this effect.

2.1.4.3 CQ3: Foreseeable misuse

A reliable L2H-off system without a driver monitoring system and attention reminders increases the likelihood of NDRA engagement and negatively affects monitoring behavior (prolonged glances away from the road ahead, lower attention ratio to road ahead). A positive attitude towards assisted/automated driving, over-reliance, boredom, underload, and sleepiness were mentioned as reasons for the observed behavior.

2.1.4.4 CQ4: Mode confusion

Mode confusion as assessed in previous studies was found to be of low importance for misuse; studies rather found drivers being fully aware of their duties. Thus, their behavior can be rated as intentional violations. However, the issue of automation expectation mismatch can be seen as mode confusion to some extent. Nonetheless, results from previous studies on this issue are limited.

2.1.4.5 CQ5: Safety level

The impact of driver behavior and driver monitoring on overall safety levels was not the focus of previous studies and therefore cannot be assessed in this literature review.

2.1.5 References

Alvarez, S. (2021*). Tesla Autopilot is not the only system that can be tricked into operating with no driver.* Teslarati. https://www.teslarati.com/tesla-autopilot-vs-gm-super-cruise-test-video/

Bindley, K., & Elliott, R. (2021). *Tesla Drivers Test Autopilot's Limits, Attracting Audiences—and Safety Concerns*. Wall Street Journal. https://www.wsj.com/articles/tesla-drivers-test-autopilots-limits-attracting-audiencesand-safety-concerns-11621503008

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., Fitch, G. M., Russell, S. M., Green, C. A., Cullinane, B., & Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (DOT HS 812 182). https://trid.trb.org/view/1367054

Boos, A., Feldhütter, A., Schwiebacher, J., & Bengler, K. (2020). Mode Errors and Intentional Violations in Visual Monitoring of Level 2. In *23rd IEEE International Conference on Intelligent Transportation Systems*, Virtual Conference.

Cahour, B., Jean-Francois, F., & Arnaud, K. (2021). Driving a Partially Automated Car with the Hands On or Off the Steering Wheel: Users' Subjective Experiences. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Springer eBook Collection: Vol. 221. Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021): Volume III: Sector Based Ergonomics* (1st ed., Vol. 221, pp. 583–592). Springer International Publishing; Imprint Springer. doi:10.1007/978-3-030-74608-7_71

Damböck, D., Weissgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance. In 2013 *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*: The Hague, Netherlands, 6 - 9 October 2013 (pp. 1656–1661). IEEE. doi:10.1109/ITSC.2013.6728467

Feldhütter, A., Härtwig, N., Kurpiers, C., Hernandez, J. M., & Bengler, K.(2019). Effect on Mode Awareness When Changing from Conditionally to Partially Automated Driving. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)* (Vol. 823, pp. 314–324). Springer International Publishing. doi:10.1007/978-3-319-96074-6_34

Garbacik, N., Mastory, C., Nguyen, H., Yadav, S., Llaneras, R., & McCall, R. (2021). Lateral Controllability for Automated Driving (SAE Level 2 and Level 3 Automated Driving Systems). In *SAE Technical Paper Series,.* SAE International 400 Commonwealth Drive, Warrendale, PA, United States. doi:10.4271/2021-01-0864

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* doi:10.1177/1541931213571433

Gustavsson, P., Victor, T. W., Johansson, J., Tivesten, E., Johansson, R., & Aust, L. (2018). What were they thinking? Subjective experiences associated with automation expectation mismatch. In *Proceedings of the 6th Driver Distraction and Inattention conference*, Gothenburg, Sweden.

Ishida, W., & Itoh, M. (2017). Effects of vehicle behavior for situation awareness when using a driving automation system. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*.

Josten, J. (2021). *User and system within the context of use – how users cope with partial automation* [PhD Thesis, RWTH Aachen University]. Retrieved from https://publications.rwth-aachen.de/record/817292

Karakaya, B., & Bengler, K. (2021). Investigation of Driver Behavior During Minimal Risk Maneuvers of Automated Vehicles. In N. L. Black, W. P. Neumann, & I. Noy (Eds*.), Springer eBook Collection: Vol. 221. Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021): Volume III: Sector Based Ergonomics* (1st ed., Vol. 221, pp. 691–700). Springer International Publishing; Imprint Springer. doi:10.1007/978-3-030-74608-7_84

Kim, H., Song, M., & Doerzaph, Z. (2021). Is Driving Automation Used as Intended? Real-World Use of Partially Automated Driving Systems and their Safety Consequences. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(1), 30–37. doi:10.1177/03611981211027150

Kraft, A.-K., Naujoks, F., Wörle, J., & Neukum, A. (2018). The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. Transportation Research Part F: Traffic Psychology and Behaviour, 52, 40–50. doi:10.1016/j.trf.2017.11.012

Kurpiers, C., Lechner, D., & Raisch, F. (2019). The influence of a gaze direction based attention request to maintain mode awareness. In *Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles*, Eindhoven.

Llaneras, R. E., Cannon, B. R., & Green, C. A. (2017). Strategies to Assist Drivers in Remaining Attentive While Under Partially Automated Driving: Verification of Human–Machine Interface Concepts. *Transportation Research Record: Journal of the Transportation Research Board*, 2663(1), 20–26. doi:10.3141/2663-03

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human Factors Issues Associated with Limited Ability Autonomous Driving Systems: Drivers' Allocation of Visual Attention to the Forward Roadway. *In Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design 2013* (pp. 92–98). University of Iowa. doi:10.17077/drivingassessment.1472

Noble, A. M., Miles, M., Perez, M. A., Guo, F., & Klauer, S. G. (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. *Accident Analysis & Prevention*, 151, 105959. doi:10.1016/j.aap.2020.105959

Othersen, I. (2016). *Vom Fahrer zum Denker und Teilzeitlenker: Einflussfaktoren und Gestaltungsmerkmale nutzerorientierter Interaktionskonzepte für die Überwachungsaufgabe des Fahrers im teilautomatisierten Modus* [Dissertation, Technische Universität Braunschweig]. GBV Gemeinsamer Bibliotheksverbund.

Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter? *Transportation Research Part F: Traffic Psychology and Behaviour*, 76, 14–25. doi:10.1016/j.trf.2020.10.001

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation Research Part F: Traffic Psychology and Behaviour*, 82, 400–411. doi:10.1016/j.trf.2021.09.010

Tivesten, E., Victor, T. W., Gustavsson, P., Johansson, J., & Aust, M. L. (2019). Out-of-the-loop crash prediction: the automation expectation mismatch (AEM) algorithm. *IET Intelligent Transport Systems*, 13(8), 1231–1240.

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, 60(8), 1095–1116. doi:10.1177/0018720818788164

2.2 Analysis of L2H-off Functions and Regulatory Aspects

Documentation by F. Reimer (fka GmbH), T. Oetermann (Institut für Kraftfahrzeuge, RWTH Aachen University)

In this section, current L2H-off functions and regulatory aspects are analyzed. The differences between L2H-on and L2H-off are identified. First, the methodology is defined (2.2.1.1). The differences are evaluated regarding the Operational Design Domain (2.2.1.2), the Human-Machine-Interface (2.2.1.3) and the Driver Monitoring System (2.2.1.4). Finally, a generic L2H-off function is described (2.2.1.5).

2.2.1 Technical considerations

2.2.1.1 Comparison of existing L2H-on and L2H-off systems

The methodology for the comparison of available L2H-off and L2H-on systems is shown in Figure 2-1. For both – L2H-off and L2H-on system – the following steps are carried out separately: First, existing systems are reviewed and summarized. Common features, e.g., of L2H-on systems, are identified and documented. For each system type, a generic function is characterized. The operational design domain (ODD), the human machine interface (HMI), and the driver monitoring system (DMS) are described. Further aspects are considered as well but are not focused within the work package.

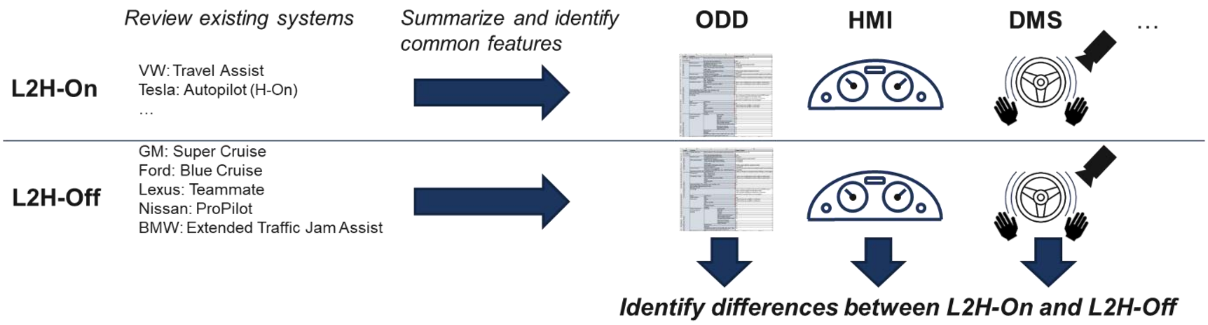


Figure 2-1: Methodology of the work package described in this chapter.

After analyzing L2H-on and L2H-off systems separately, they are compared. In the following Sections the similarities and differences regarding the ODD (Section 2.2.1.2), the HMI (Section 2.2.1.3), and the DMS (Section 2.2.1.4) are identified. Furthermore, the function scope and its description are briefly touched (Section 2.2.1.5), as it is transferred and discussed in more detailed in context of the risk analysis (see Section 2.3). The analysis focuses mainly on L2H-off systems which are available in the USA.

2.2.1.2 Operational Design Domain (ODD)

Please note: In discussion with the VDA members at the beginning of the project, it has been decided that the project is only focusing on highway scenarios. Urban environment and parking scenarios have been excluded explicitly.

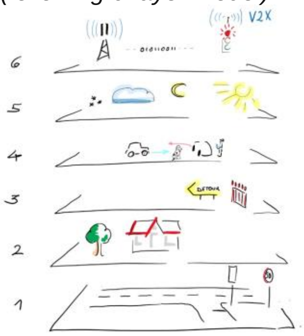
The ODD is characterized according to the documents and guidelines listed in Table 2-1.

Table 2-1: ODD Guidelines

Document	Author
SAE J3016	SAE (2021)
6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment	Scholtes et al. (2021)
AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon	Automated Vehicle Safety Consortium (2020)
Acclimatizing the Operational Design Domain for Autonomous Driving Systems	Sun et al (2022)
DOT HS 812 623: A Framework for Automated Driving System Testable Cases and Scenarios	NHTSA (2018)

The attributes considered in the ODD description have been classified based on the 6-layer model. A summary of the key aspects is shown in Figure 2-2. In general, the described ODD shall not be understood as requirements which a L2H-off system has to fulfill. The ODD summarizes limitations of currently available L2H-off systems, according to their manufacturers.

Systematic ODD description for existing L2 systems
(following 6-layer model)



Layer	L2H-On	L2H-Off
Digital Information	No specific requirements	• **HD map (incl. regular updates)** • **GPS permanently available**
Environmental Conditions	No extreme weather conditions: e.g. snow, sun glare, …	No extreme weather conditions: e.g. snow, sun glare, …
Dynamic Objects	Only motorized vehicles, no oncoming or crossing traffic, no pedestrians or animals	L2H-On criteria (+ **Currently: towing not yet included (update in progress)**)
Temporary Modifications of L1 and L2	No modifications allowed	No modifications allowed
Roadside Structures	No specific requirements	• **Obstructions excluded** • **Tunnels excluded**
Road Network and Traffic Guidance Objects	• Expressways or well-developed roads • Good lane markings	• **Limited set of divided highways (HD-maps required)** • Good lane markings (or lead vehicle)

Figure 2-2: Systematic ODD description (Scholtes et al, 2021) and comparison

In contrast to generic L2H-on systems, most L2H-off systems require digital information such as HD-map, which is updated regularly (layer 6 in Figure 2-2). L2H-off appears to be only available, if the GPS signal quality is sufficient. Some L2H-off systems don’t adhere to these limitations.

Regarding environmental conditions (layer 5), both L2H-off and L2H-on systems typically are not available at extreme weather conditions such as sun glare, snow, heavy rain, or slippery road surface. With regard to dynamic objects (layer 4), both system types have similar ODDs: Only motorized vehicles traveling in the same direction are covered. Interaction with pedestrians, animals, or oncoming targets are not supported. Both L2H-off and L2H-on system do not cover temporary modification (layer 3) of level 1 & 2, such as construction zones. Regarding the roadside structure (layer 2), such as surrounding buildings, no limitations apply for L2H-on systems. Some L2H-off systems are limited by tunnels or obstructions to achieve a sufficient GPS signal quality (see layer 6). On road network and traffic guidance objects level (layer 1), L2H-on systems are in general limited to expressways and well-developed roads. L2H-off systems are further limited to divided highways only. Most systems are even limited to a set of divided highways. Both L2H-off and L2H-on system usually require well visible lane markings.

As part of the ODD description, the narrative is formulated as followed:

“The system is designed to operate only on a defined set of divided highways in Germany (“Bundesautobahnen”). No restrictions regarding the speed limit apply. Merge lanes, on and off ramps are excluded from the ODD. The system is capable of operating during night and day excluding situation with sun glare on the vehicle front. It can operate in fair weather, excluding extreme weather conditions such as heavy rain or snow, and if the road surface is not covered by snow or slippery. It recognizes speed limit signs inside this ODD, other traffic signs and control devices are excluded. Work zones and other temporary changes in the road network are excluded from the ODD. The system recognizes clearly visible lane markings. Sections with faded markings with unreliable consistency are excluded from the ODD.”

2.2.1.3 Human Machine Interface

The following section presents the findings of the analysis of the Human Machine Interface (HMI) of existing L2H-off systems on the market. A particular focus is set on the differences in the context of HMI between L2H-on and L2H-off systems. Furthermore, a general description of a potential HMI for an L2H-off system is presented based on the similarities of the systems analyzed.

Differences between L2H-on an L2H-off Systems

Depending on the design of the L2H-on system, there are few significant differences between L2H-on and L2H-off systems in terms of HMI. The information display, the control elements, and the escalation cascade in case of unsuitable behavior of the driver are comparable to the corresponding realizations of L2H-off systems. However, minor differences are also identifiable and are briefly described below.

Major differences occur in the context of differentiating between the hands-on and hands-off modes of an L2H-off system while the system is activated. In principle, systems facilitate a clear differentiation between the various driving modes and the associated tasks and responsibilities of the driver. The representation of the driving mode is achieved on the one hand by displaying the hands-off or the hands-on icon during the assisted driving. In addition, the system status is communicated, for example, via changing color schemes (e.g., blue or turquoise for driving in hands-off mode). Generally, assisted driving in hands-off mode is accompanied by an adjustment of the IC's assistance display. Some manufacturers (e.g. Mercedes-Benz, GM) provide an additional, dedicated display element to communicate the system status. This is used exclusively to communicate the system status (system active/inactive; hands-on/hands-off) during use of the L2H-off system. In this context, this dedicated display element also participates in the escalation cascade of the Direct Control Request (DCR).

Another common feature of two L2H-off systems (Lexus, Nissan) consists of the display of upcoming events in relation to the hands-off driving mode. The display is not a generally implemented feature of the L2H-off systems, but since it might contribute to a clear classification between the hands-on and hands-off systems, it is nevertheless mentioned here with limited validity. The list of upcoming events includes takeover requests to the user that can be reliably predicted by the system and the reasons for the necessary takeover (e.g., sharp curve radii, highway entrances or exits, highway Sections that are not specified, etc.).

General Description of an HMI for a L2H-off system

Based on the preceding analysis of the considered L2H-off systems with respect to the system-side HMI, a summary of the common features of the systems is given below. Based on the commonalities, the following chapter will thus provide a generic HMI as an average of the currently available as well as announced systems. For a detailed description of the respective commonalities, please refer to the previous chapter, as this Section merely provides a clear summary.

Relevant components of L2H-off systems

Input

Hardkeys

Most of the manufacturers of systems under consideration use hardkeys as dedicated control elements to activate and parameterize the systems. Depending on the design of the system (extension of existing Advanced Driver Assistance Systems (ADAS) functions or new, stand-alone assistance system), the existing interaction logic is extended and includes the existing dedicated hardkeys or, alternatively, a new hardkey may be introduced to fulfill this goal. In either case, the hardkey integrates into the existing interaction landscape with respect to the vehicle's ADAS functions. Placement close to the instrument cluster in the multifunction steering wheel is therefore primarily observed.

Softkeys

Softkeys are generally not favorable for use within an L2H-off functionality. The only exception found is the steering column lever of the direction indicator. While this is only used to activate the direction indicator during manual driving (whereby assistance functions such as blind spot warning can react to this trigger), the steering column lever is used to perceive the user's lane change request in systems that have a lane change assistant.

Output

Visual

Since the visual information channel in the vehicle is one of the most important channels for transmitting central information, warnings, and instructions, all the main display surfaces of the vehicle are involved in the presentation of information with regard to the L2H-off systems. Depending on the design of the vehicle concept, the IC (instrument cluster), HUD (head-up display), and CID (central information display) are involved in the presentation of central information and warnings from the L2H-off system.

Manufacturers of the systems under consideration present the current system state at any point in time. Typically, simple, understandable, and quickly comprehensible pictograms are used for this purpose, e.g. hands-on/hands-off pictograms. In addition, the majority of manufacturers link a special driving assistance display to the use of L2H-off systems. In this context, the displays visualize the objects or road users detected in the environment. Special color schemes are also tied to the system status (L2H-off system on/off; hands-off mode on/off). The colors blue and turquoise are increasingly used for hands-off driving functionalities (IC backlighting, color illumination of pictograms, color scheme for textual output, etc.). Warnings are also generally transmitted in color coding. Less critical warnings are usually indicated in orange or yellow, while critical warnings (e.g., DCR) are highlighted in red.

Audio

In addition to visual displays, a multimodal approach is used by a large number of manufacturers for general system communication, especially for critical warnings. In concrete terms, a large proportion of manufacturers incorporate an auditory output. Depending on the system structure and system configuration, the acoustic output is designed in the form of signal tones or alternatively as voice output. The design of the acoustic output is variable depending on the manufacturer. In principle, the output is adjusted in frequency and volume depending on the criticality of the information.

Kinesthetic/Haptic

Other communication modalities that are used for multimodal warning approaches are haptic or kinesthetic outputs. There are already some implementations in systems that are currently market-available or have been announced (e.g. brake impulse, vibrating seat cushion, seat belt pre-tensioner). Typically, these warnings are used as part of the escalation cascade for the DCR.

Activation of the L2H-off functionality

With regard to the activation of the systems, as described above, a large proportion of the systems considered implement dedicated controls by means of hard keys. Since the activation of these systems is usually depending on the fulfillment of external boundary conditions (which are described in the ODD), the user must be informed whether the boundary conditions are fulfilled at the current time or not. Exemplary implementations of the considered systems use e.g., illuminated control elements, the display of a pictogram, or a text output. Some of the systems also use an acoustic display for this purpose. If an error occurs during the activation of the system, the user is informed about the failed activation and the consequences by a large part of the considered systems. Generally, a textual information and, optionally, a color-visual feedback is given. In addition, the system status is communicated at each point in time. Some systems also provide recommendations for action in the form of textual information (e.g., adjustment of the current driving parameters, reactivation of the system, etc.). If the system is successfully activated, the user receives information about when he or she may take the hands off the steering wheel. This is usually implemented in the form of the pictograms described. Alternatively, acoustic or textual outputs are also used.

User-intended adaptations of semi-automated driving

In terms of user-intended adaptation of various parameters of the systems, manufacturers use similar interaction logics from existing solutions of ADAS functions. Established parameterization logics for distance and speed from ACC systems are used. The distance regulation can usually be varied with a dedicated hardkey in a multi-stage scale (long following distance – short following distance). The target–speed is usually set to the current driving speed with a dedicated hardkey and can then be increased or decreased in set steps with further hardkeys (+/-).

DCR, inattentiveness warning, and escalation cascade

The considered systems react comparably in case of detected inattention or a necessary takeover by the driver. Commonly, a multi-stage escalation cascade is implemented for this purpose. There are differences in the number and exact design of the stages depending on the manufacturer or system. The greatest possible commonality is described below. There are also differences in whether a detected driver inattention directly triggers a takeover request or triggers an additional escalation stage before the actual DCR escalation.

Inattention warning

The inattention warning represents the first stage of the escalation cascade and is triggered by the detected driver inattention (e.g., eyes of road, no reaction to system message). The systems that implement this stage of the escalation cascade usually use less urgent signals compared to the other escalation stages. Warnings (textual/pictography) in a yellow or orange color scheme are typically used. Some manufacturers also use a multimodal communication approach at this stage and reinforce the visual attention request by acoustic and/or haptic/kinesthetic information.

1st stage of DCR escalation

The first escalation level of the actual DCR is already more urgent compared to the inattention warning. A red color scheme of the corresponding displays, warnings and icons is usually used. The pictography, in the form of the hands-on icon, indicates the need to switch to hands-on mode. In addition, textual information about the DCR is displayed. The warnings and ambient light pulses at a high frequency to indicate the urgency of the warning. The request is already reinforced by an acoustic output, which is represented in the rule by a pulsating signal tone.

Haptic and/or kinesthetic stimuli are used by single manufacturers at this stage for additional reinforcement.

2nd stage of DCR escalation

In the second escalation level, the urgency increases. The frequency of the pulsing displays as well as the warning tones increases. In this stage, some manufacturers switch from an auditory output in the form of a signal tone to an output as a verbal takeover request. If the system can display a haptic/kinesthetic warning, this is used by the various systems at this point at the latest to emphasize the urgency of the takeover request. As described above, the haptic/kinesthetic output is represented by the systems either as brake impulse, vibrating seat cushion or seat belt pre-tensioner.

3rd stage of DCR escalation

The last stage of the escalation scheme is the emergency braking of the system. First, the hazard warning lights are switched on and careful deceleration to a standstill is initiated. Then, an automatic emergency call is sent. The user is informed about the reasons for the system-initiated emergency stop (textual information in the IC/CID/HUD). In addition, some of the systems examined recommend action to the user if he or she is able to resume manual control of the vehicle. The recommended action may also be presented as textual information and usually includes the information that the user should resume manual vehicle guidance (longitudinal and lateral guidance).

2.2.1.4 Driver Monitoring System (DMS)

In current L2H-on systems, a detection is implemented, whether the driver's hands are on the steering control according to UN ECE R79. This detection is implemented using the principle of measuring the manual steering effort (e.g., via torque sensors at the steering column) or inductive sensors in the steering control.

L2H-off system use an additional camera-based driver monitoring system. A driver-facing camera is positioned in the cockpit e.g., on the steering column, in the central display, next to the rear mirror or in the instrument cluster. For situations with low illumination, infrared lights are added. According to manufacturers, the camera measures the gaze direction and the head position. Some manufacturers also measure the eyelid closure.

2.2.1.5 Generic Function and Item Definition

As part of this chapter, the generic L2H-off function is described and an Item Definition is formulated, which is used in the following section (such as in the Hazard and Risk Analysis).

The boundary diagram, Figure 2-3, visualizes the interaction of the L2H-off System with other components.

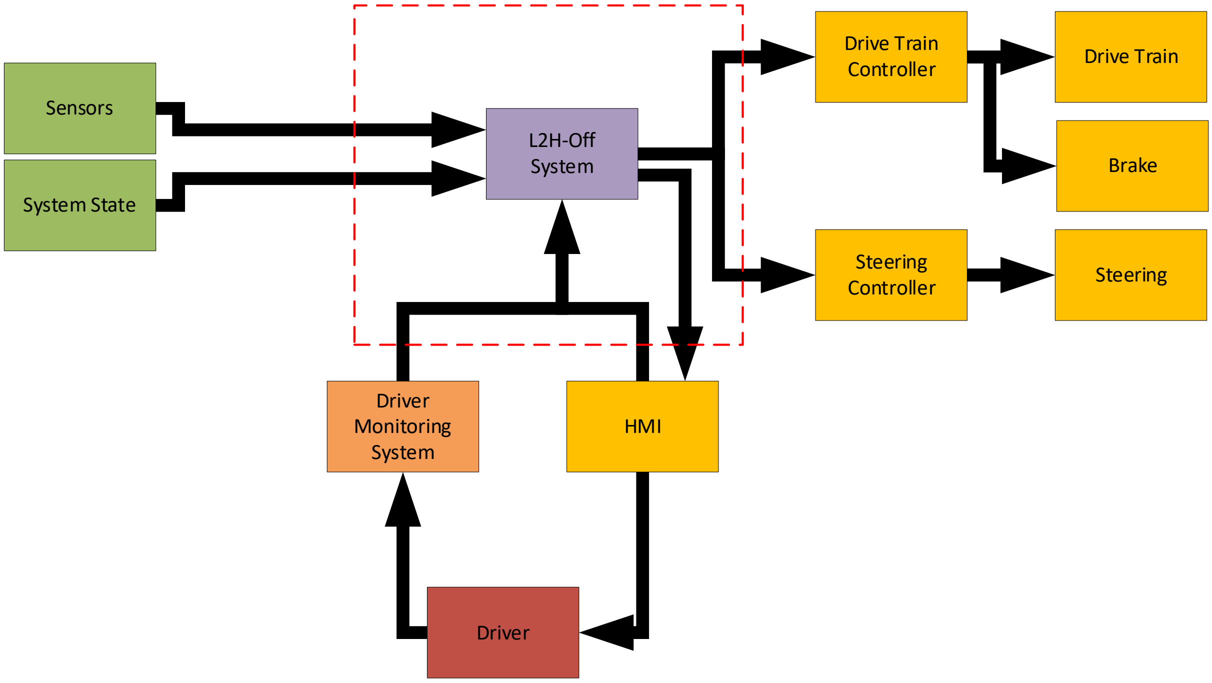


Figure 2-3: Boundary diagram: generic L2H-off system

As inputs, the function requires information regarding its environment, which is provided by the sensor modules. Typically, this information is represented by a list of objects including its position, type, velocity, and further attributes. Additionally, the state of the vehicle is required. This includes the states of involved components such as brake or steering as well as other safety relevant components. The state of the driver is provided by the driver monitoring system and his/her inputs are provided as input by the HMI (e.g., primary and secondary control inputs). Based on these inputs the L2H-off system calculates the needed actions to control the vehicle: It feeds longitudinal control commands to the Drive Train Controller as well as the brakes and lateral control commands to the Steering Control. The L2H-off system also provides information to the driver via HMI. This includes the current system state as well as actions required by the driver.

The generic function and its system states are defined based on the state of the art and especially based on the ISO 21717 (International Organization for Standardization. (2018)).

2.2.2 Relevant standards and regulations: ISO 21717, UN ECE R79 & UN ECE R157

In the context of L2H-off systems, three documents have been identified as especially relevant: ISO 21717, UN ECE R79, and UN ECE R157. The following sections give an overview over the relevant points addressed by the standards and regulations.

2.2.2.1 ISO 21717

ISO 21717 (International Organization for Standardization (2018)) standardizes Partially Automated In-Lane Driving System (PADS), which shall support the driver in keeping the vehicle within the lane, a maximum vehicle speed and a minimal distance to the leading vehicle. The document does not specify if the driver is required to hold the steering control. Therefore, it is valid for L2H-on and L2H-off systems. ISO 21717 describes a possible generic design of such a system and refers to ISO 15622 regarding the longitudinal vehicle control. The following requirements are formulated for the driver monitoring: “PADS shall have means to recognize whether the driver is in principle able to take over the vehicle control and if the driver can supervise the system behavior. How this is done in detail is up to the Original Equipment Manufacturer (OEM). Examples for such means are a hands-off detection, a driver recognition camera, or other suitable means.” (ISO 21717:2018).

Neither minimum risk maneuvers nor emergency maneuvers, which shall be performed by the PADS, are formulated or required.

2.2.2.2 UN ECE R79

UN ECE R79 is regulating steering systems in general. Within the scope of this project, automatically commanded steering functions (ACSF) of category B1 are considered as relevant. These systems shall keep the vehicle within its lane by controlling the lateral movement. ACSF of category B1 do not include support for longitudinal control. Currently, this regulation is relevant for state-of-the-art lane keeping systems and in combination with an ACC for L2H-on systems.

While the system is active, an optical warning shall be prompted, if the driver is not holding the steering wheel for more than 15 s. After no longer than 30 s, an additional acoustic warning shall be activated, and the warning shall be changed to a red symbol. At the latest 30 s after the start of the acoustic warning, the lateral control shall be activated, and the driver shall be informed.

The UN ECE R79 does not describe or require minimum risk maneuvers or emergency maneuvers.

2.2.2.3 UN ECE R157

The UN ECE R157 No, U. R. (2021) has been introduced in 2021 to address Automated Lane Keeping Systems (ALKS), which are carried out as L3 systems. The operational speed is limited to 60 km/h and the system may only be used on divided highways. In the context of the UN ECE R157, ALKS consist of lateral and longitudinal control of the vehicle: “The activated system shall detect the distance to the next vehicle in front as defined in paragraph 7.1.1. and shall adapt the vehicle speed in order to avoid collision”. “The activated system shall keep the vehicle inside its lane of travel and ensure that the vehicle does not cross any lane marking (outer edge of the front tire to outer edge of the lane marking). The system shall aim to keep the vehicle in a stable lateral position inside the lane of travel to avoid confusing other road users.” (UN ECE R157)

The regulation introduces specific maneuvers in case of an emergency (EM) and a minimum risk maneuver (MRM). The system behavior is specified within the document. The MRM shall be performed if the driver is unavailable for a defined amount of time.

Furthermore, the driver monitoring is described with regard to different driver states (e.g., attentive or unavailable). To evaluate the driver states, criteria focusing on the driver's head and eye position and movement are proposed.

In addition, UN ECE R157 requires a dedicated optical driving mode display, which shall be "an easily perceptible indication in the peripheral field of vision and located near the direct line of driver's sight".

2.2.3 Discussion of current L2H-off systems

The identified main differences between current L2H-on and L2H-off systems, namely

- the extension of the driver monitoring by using cameras,
- a dedicated driving mode display,
- an attentiveness alert, reminding the driver once inattention is detected,
- an adaptation of the takeover request cascade and
- limitations of the road network: limited set of divided highways, availability of GPS and maps (optionally HD-maps)

are discussed in this work package referring to the five challenges and questions (CQs). Figure 2-4 gives an overview of the interaction between identified differences and the CQs. The matrix shows which technical difference (horizontal axis) addresses which CQ (vertical axis).

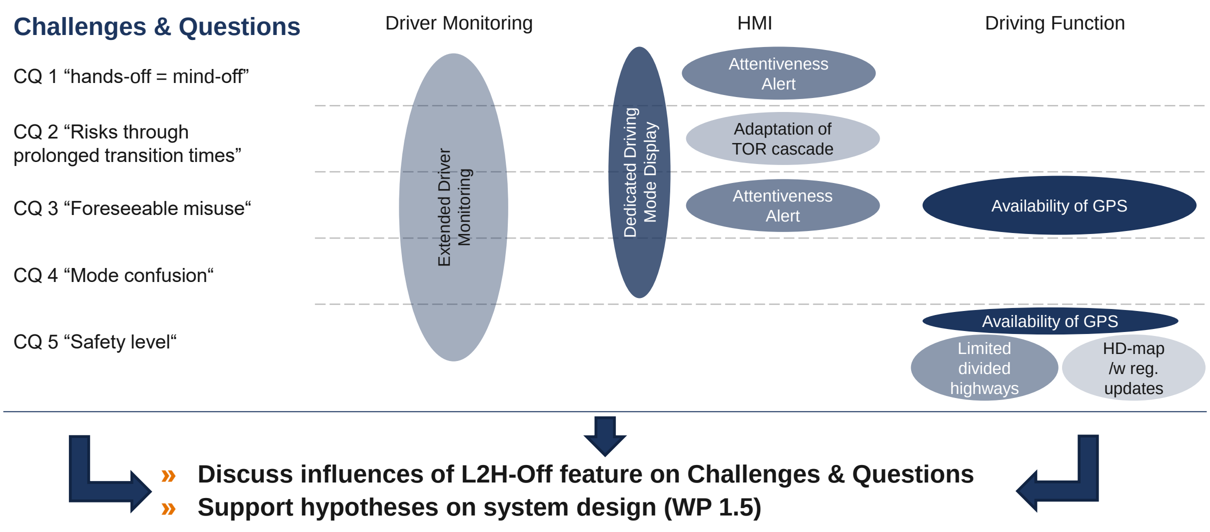


Figure 2-4: CQ: Discussion of current L2H-off systems

The figure shall be understood as broad overview over the discussion which took place within the project. The main purpose of the discussion is to create a backlog which can be used in following work packages related to the definition of design hypotheses such as Section 1.5 or Section 5.1, e.g., to identify technical solutions which could address a specific CQ.

2.2.4 Transfer to other chapters in this document

As mentioned in the previous sections, the main goal of the work package and the sub project is to build a foundation representing the current state of the art for the upcoming work packages. Within this section, an overview of the interfaces of this work package with work packages of the project, as described in other chapters of this document, is shown (see Figure 2-5).

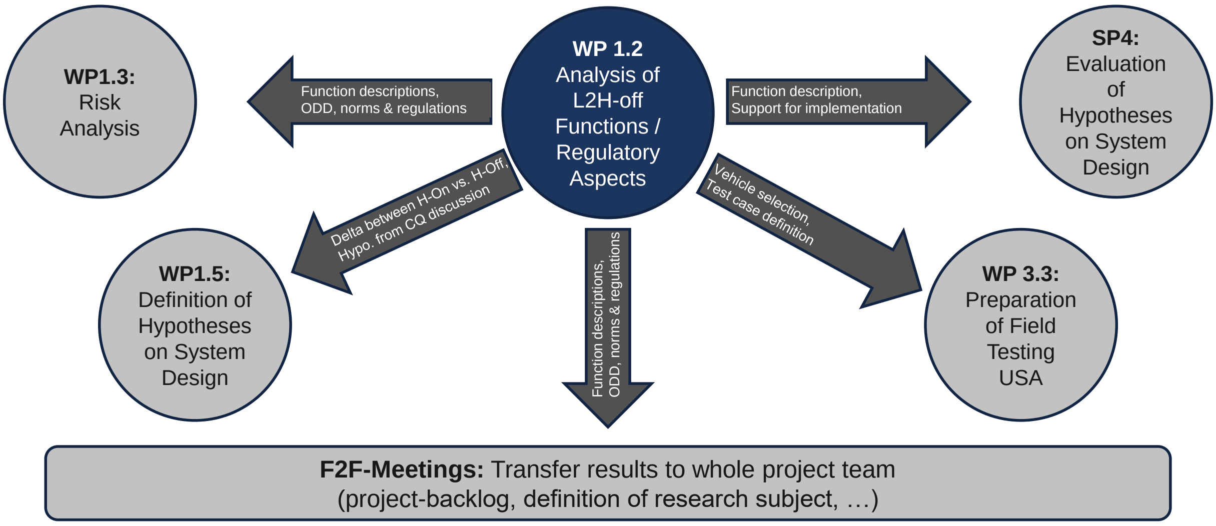


Figure 2-5: Interfaces of this chapter to other work packages and subprojects (WP1.3 – see Chapter 2.3; WP1.5 – see Chapter 2.5; WP3.3 – see Chapters 4.1 and 4.2; SP4 – see Section 5).

2.2.5 References

Automated Vehicle Safety Consortium (2020). AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon. (AVSC00002202004). *SAE Industry Technologies Consortia.* https://www.sae.org/standards/content/avsc00002202004/

International Organization for Standardization. (2018). *Intelligent transport systems — Partially Automated In-Lane Driving Systems (PADS) — Performance requirements and test procedures* (ISO Standard No. 21717:2018). Retrieved from https://www.iso.org/standard/71471.html

National Highway Traffic Safety Administration (2018). A Framework for Automated Driving System Testable Cases and Scenarios. (DOT HS 812 623) *NHTSA* https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13882-automateddrivingsystems_092618_v1a_tag.pdf

SAE (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, (J3016_202104). *SAE. https://www.sae.org/standards/content/j3016_202104/*

Scholtes, M. et al. (2021). 6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment. *IEEE Access, 9*, 59131-59147. doi:10.1109/ACCESS.2021.3072739.

Sun, C. et al. (2022). Acclimatizing the Operational Design Domain for Autonomous Driving Systems. *IEEE Intelligent Transportation Systems Magazine*, *14(2)*, 10-24. doi: 10.1109/MITS.2021.3070651.

United Nations Economic Commission for Europe (2021). *Automated Lane Keeping Systems (ALKS).* UN Regulation No. 157, pp. 75-137. Retrieved from https://unece.org/transport/documents/2021/03/standards/un-regulation-no-157-automated-lane-keeping-systems-alks

2.3 Risk Analysis / Relevant Scenarios

Documentation by F. Reimer (fka GmbH)

Within this section, a risk analysis for the generic L2H-off function is carried out. The generic L2H-off system has been defined in the previous chapter (Section 2.2). The item definition is refined and enhanced during the risk analysis. The relevant scenarios for the risk analysis are specified and selected based on VDA 702 (VDA Verband der Automobilindustrie e.V., 2015a).

The focus of this section is the comparison of L2H-off and L2H-on systems and identifying the impact of the differences on the risk analysis. Therefore, the risk analysis is not carried out strictly according to ISO 26262-3 (International Organization for Standardization, 2018a). Based on this comparison the need for action with regard to requirements for additional components is derived.

2.3.1 Scenarios

Within this section, relevant scenarios for the risk analysis are defined and selected. It is not the goal to cover all possible scenarios in which the generic L2H-off system may be active, but the focus is set to scenarios in which differences between L2H-off and L2H-on systems are expected with regard to the risk analysis. The scenarios which are considered within this section are derived from the VDA 702 catalogue (VDA Verband der Automobilindustrie e.V., 2015a). It lists scenarios occurring during road use. Each situation is paired with its exposure parameter E.

The scenario description focuses on the following attributes: velocity, longitudinal and lateral dynamics, traffic and location. Attributes like number of passengers, load and ignition state are not considered. Furthermore, all scenarios which are outside of the L2H-off ODD are gathered as “Outside ODD”.

As a result, the scenario attributes as displayed in Table 2-1 have been selected for the risk analysis.

Table 2-1: Attributes for relevant scenarios for risk analysis

Attribute	Value
Velocity	Driving at 12-130 km/h
Longitudinal Dynamics	Driving with normal longitudinal acceleration (<2 m/s²) & with normal longitudinal deceleration (>-4 m/s²)
Lateral Dynamics	Driving on straight road Driving on road with max. allowed curvature (according to “Richtlinie für Anlage von Autobahnen” (Arbeitsgruppe Straßenentwurf, 2008)
Traffic	Free driving Following lead vehicle with normal distance Approaching slower vehicle
Location	Highway (“Bundesautobahn”)

The complete scenario list (Table 2-1) consists of all possible combinations of scenario attributes.

2.3.2 Item Definition

As described in Section 2.2, an item definition for the generic L2H-off system has been formulated. Within this work package, the item definition has been refined and enhanced in the process of performing the risk analysis, e.g. by adding safety relevant assumptions.

As shown in Figure 2-1, the relevant ISO standards as well as the ODD description defined in Section 2.2 are considered in the item definition.

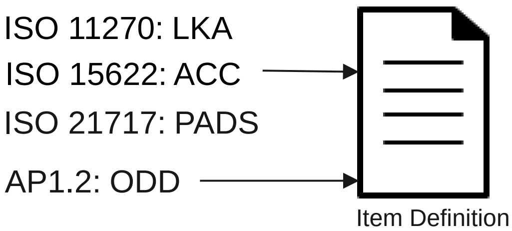


Figure 2-1: Inputs to Item Definition (International Organization for Standardization, 2014; International Organization for Standardization, 2018b; International Organization for Standardization, 2018c)

The functional components of the L2H-off system are defined. The system is deconstructed into longitudinal and lateral control as well as into the interaction with the driver.

Based on ISO 21717 the function states are defined, see Figure 2-2. The state "Lateral Active" is divided into two sub states: "Hands-On Active" and "Hands-Off Active". All other states are kept identical to ISO 21717 (International Organization for Standardization, 2018c).

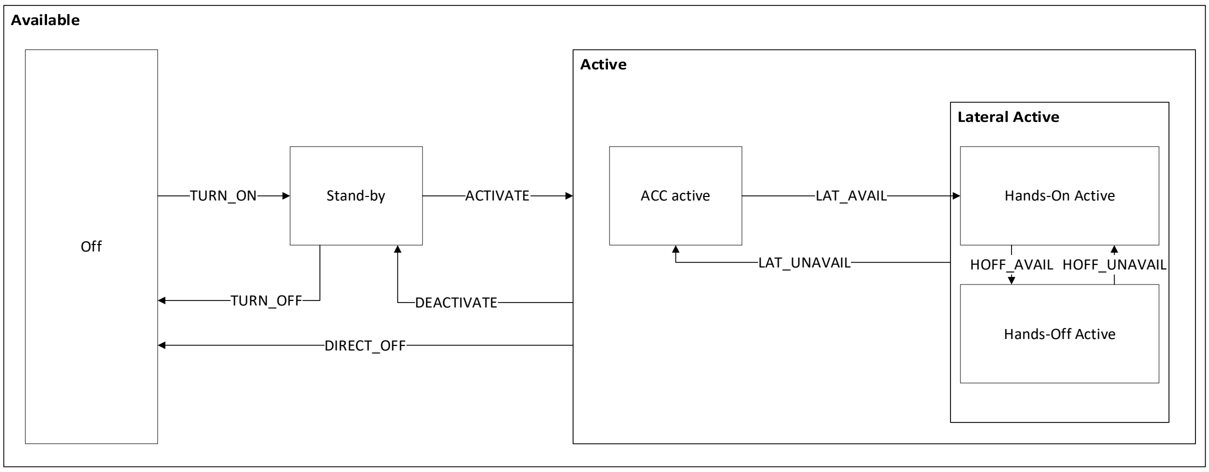


Figure 2-2: L2H-off States and Transitions

The operational limits are defined according to ISO 21717 (International Organization for Standardization, 2018c), e.g. lateral acceleration or curvature. In addition, the velocity is limited to 0-130 km/h. This has been discussed and decided by the VDA members.

2.3.3 Hazard and Risk Analysis

The Hazard and Risk analysis (HARA) has been carried out following ISO 26262-3 (International Organization for Standardization, 2018a). However, changes to the process are made to decrease the complexity of the process. The main goal is the identification of differences in the HARA between L2H-off and L2H-on systems. Aspects, which are identical for both system types, are not the focus of this section.

Figure 2-3 shows the steps which are needed in general to perform the HARA.

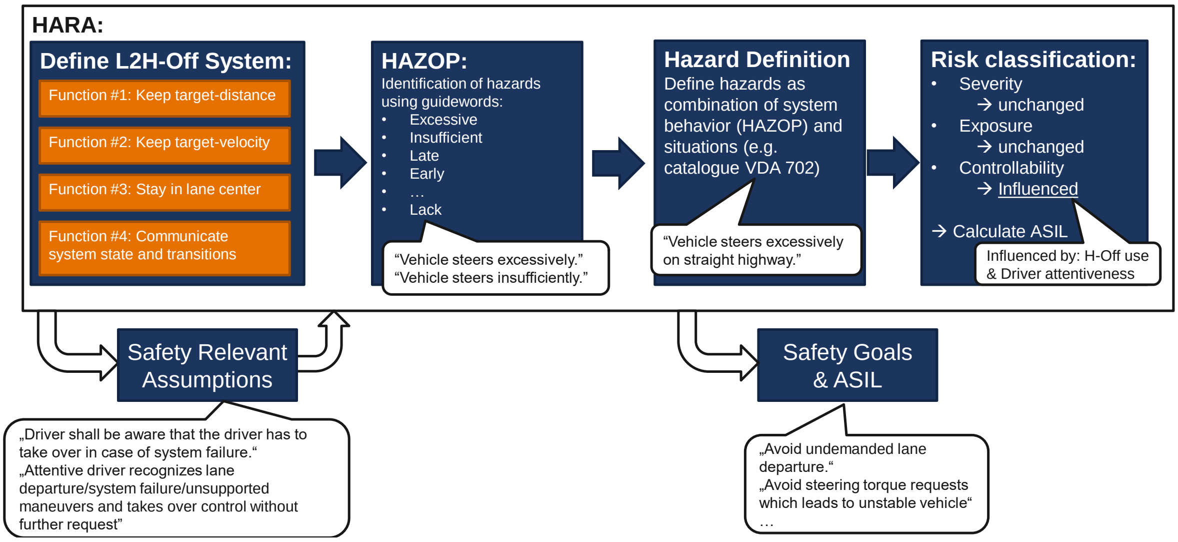


Figure 2-3: Risk Analysis within this project

In the first step, the system is broken down to functions. In this section, the L2H-off system consists of the following four functions:

- "Function #1: Keep target-distance"
- "Function #2: Keep target-velocity"
- "Function #3: Stay in lane center"
- "Function #4: Communicate system state and transitions"

It shall be noted that the deconstruction of the L2H-off system is also valid for a generic L2H-on system.

As a first simplification of the HARA, function #1 and #2 are considered as not relevant for the comparison of L2H-off and L2H-on. Function wise, the longitudinal control of the vehicle is equivalent to an ACC system for both L2H-off and L2H-on. For both systems, the driver is allowed to remove his or her feet from both the acceleration and braking pedal. The driver's responsibility to supervise the vehicle longitudinal control stays unchanged. It is assumed, that changes in the driver's focus and reaction times are also included in the comparison of the lateral control (function #3).

In the second step, a hazard and operability study (HAZOP) is carried out for the remaining functions.

Based on this step, the hazards are defined as a combination of the system behavior and the selected scenarios.

The risk classification is only carried out partially: The severity and exposure are expected to stay unchanged, as the scenarios are valid for both L2H-on and L2H-off. However, the controllability is expected to change based on the system design. As this is very specific to the final implementation and has to be reevaluated by each OEM, the controllability and calculation of the ASIL shall not be in focus of this analysis. In discussion with VDA experts, it has been noted, that the formulation of the safety goals is very specific to each OEM. Additionally, VDA experts agreed, that the safety goals stay unchanged between L2H-off and L2H-on systems.

As a result of the analysis, the following requirements are formulated based on the safety relevant assumptions and based on the safety goals. These requirements describe the difference between L2H-off and L2H-on systems:

- The system displays the driver’s task to the driver;
- The system detects whether the driver is in principle able to take over the vehicle control and if the driver supervises the system behavior;
- The system detects whether the driver holds the steering wheel;
- The system directs the driver’s attention to the driving task.

These requirements are transferred to following work packages, where hypotheses on system design are gathered and formulated.

2.3.4 References

Arbeitsgruppe Straßenentwurf (2008). RAA – Richtlinien für die Anlage von Autobahnen Ausgabe 2008. *FGSV Verlag*

International Organization for Standardization. (2018a). *Road vehicles - Functional safety - Part 3: Concept phase* (ISO 26262-3:2018). Retrieved from https://www.iso.org/standard/68385.html

International Organization for Standardization. (2014). *Intelligent transport systems — Lane keeping assistance systems (LKAS) — Performance requirements and test procedures* (ISO 11270:2014). Retrieved from https://www.iso.org/standard/50347.html

International Organization for Standardization. (2018b). *Intelligent transport systems — Adaptive cruise control systems — Performance requirements and test procedures* (ISO 15622:2018). Retrieved from https://www.iso.org/standard/71515.html

International Organization for Standardization. (2018c). *Intelligent transport systems — Partially Automated In-Lane Driving Systems (PADS) — Performance requirements and* *test procedures* (ISO 21717:2018). Retrieved from https://www.iso.org/standardhttps://www.iso.org/standard/71471.html

Verband der Automobilindustrie e.V. (2015). VDA 702: Situationskatalog E-Parameter nach ISO 26262-3. *Dokumentation Kraftfahrwesen: Berlin, Germany*.

2.3.5 Appendix

Table 2-2: Relevant Scenarios

ID	Velocity	Longitudinal Dynamics	Lateral Dynamics	Traffic	Location
#1	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Free driving	Highway
#2	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Free driving	Highway, construction zone without structural separation ("baulicher Trennung")
#3	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Following lead vehicle with normal distance	Highway
#4	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Following lead vehicle with normal distance	Highway, construction zone without structural separation ("baulicher Trennung")
#5	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Approaching slower vehicle	Highway
#6	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on road with max. allowed curvature (according to "Richtlinie für Anlage von Autobahnen" (Arbeitsgruppe Straßenentwurf, 2008))	Approaching slower vehicle	Highway, construction zone without structural separation ("baulicher Trennung")
#7	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving on straight road	Free driving	Highway
#8	12-130 km/h	-2 to 2 m/s² (normal driving)	Driving outside ODD	Driving outside ODD	Driving outside ODD

2.4 Definition of Relevant Metrics

Documentation by J. Hiller (Institut für Kraftfahrzeuge, RWTH Aachen University)

Whilst the applied methods differ in detail for the various data sources (field data, simulator data, expert surveys, field operational test (FOT) data, …), the metrics used for each of these data sources should be harmonized. This offers the chance of streamlining the outcome.

The aim is therefore the operationalization of the interaction quality with L2H-off systems using relevant metrics. These metrics should cover the user behavior and interaction as well as how well the driving task is fulfilled. As additional metrics, measures relating to the objective and subjective traffic safety are to be considered.

In the following, the terms “objective data” and “subjective data” are used. In the context of this project, objective data refers to data recorded within the vehicle or from the vehicle itself using recording equipment installed within the vehicle. Subjective data in the context of this project on the other hand refers mainly to data collected by questionnaires and interviews before, during or after the experiments.

2.4.1 Goals of the Definition of Relevant Metrics

For the definition of all relevant metrics, the varying needs of the different evaluations within the project need to be considered. Additionally, all five “Challenges and Questions” (cf. Section 2.1) should be considered during the evaluation and therefore for the definition of the metrics.

In detail, the following tasks are to be considered:

- Definition of relevant metrics for the evaluation of user behavior including the interaction with the system and/or vehicle (gaze detection, hand detection)
- Definition of relevant metrics on guidance and stabilization level which describe the quality of the fulfillment of the driving task
- Derivation of a suitable measurement concept with regard to the necessary measurement equipment for addressing the challenges and questions (e.g. driver monitoring) as well as for the additional acquisition of subjective data.

Since each of the experiment has a different focus, it was and is not possible to find one set of metrics that can be used for all experiments. Instead, focus was put on harmonizing the metrics in their definitions and aligning them within the project but also with other previously conducted studies found in literature.

2.4.2 Interactions, Inputs and Outputs

To account for all metrics, the interaction with the other work packages within the project needs to be analyzed. Not only does this concern the literature research and metrics used in similar studies, but also the interaction with the work packages related to the experiments conducted within the project. All interactions can be seen in Figure 2-1 and will be detailed in the following.

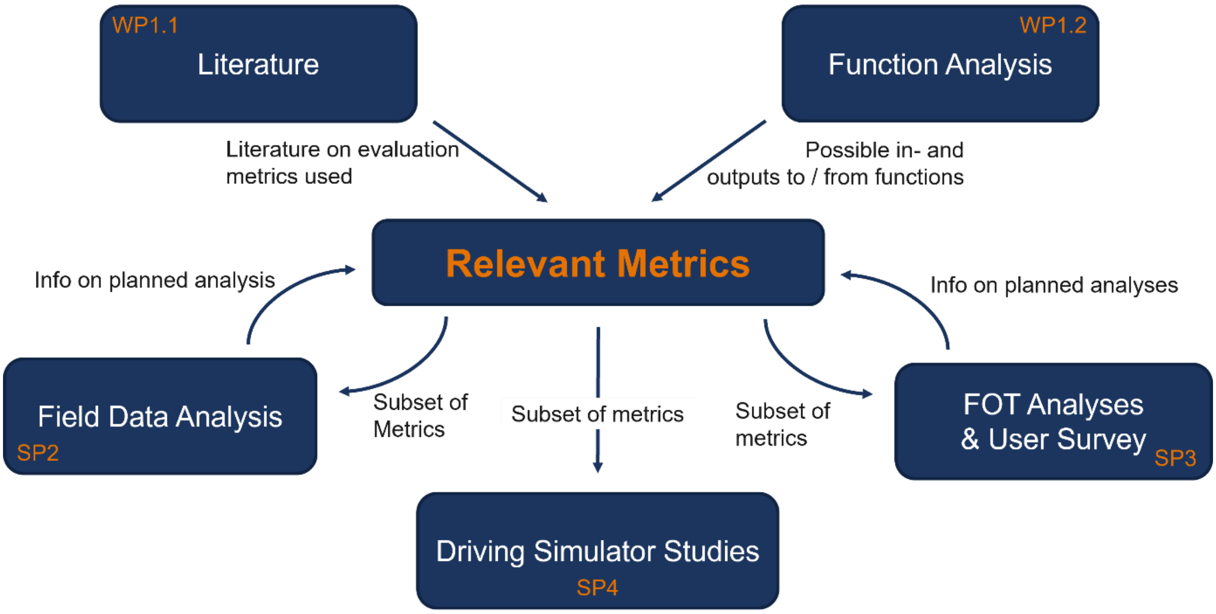


Figure 2-1: Interactions with inputs and outputs necessary for the definition of the relevant metrics

As stated in Section 2.1, there are not yet many studies related to L2H-off systems with driver monitoring. However, over the past years, there have been multiple studies regarding SAE L2H-on systems as well as Level 3 automation (Section 2.1). Although there are differences between L2H-on or Level 3 systems and L2H-off systems, certain aspects or evaluations are the same and can be used as a basis for the evaluations done within this project. As an input from the literature review, an overview of used evaluation metrics is therefore used.

For the definition of the relevant metrics, an understanding of the analyzed functions is of importance. The input from the work package analyzing the functions (cf. Section 2.2) fills this gap. The input is in the form of necessary driver input or action to activate the function, what he or she is allowed to do while using the function and what modality is used for displaying messages to the driver and also issuing warnings.

Although the metrics are also based upon literature review and the analysis of the functions, the performed experiments within the project are a separate category for the relevant metrics as they serve not only as input to the relevant metrics but also benefit from the defined metrics. In the following, the inputs and output to the experiments planned in SP2, SP3 and SP4 are described.

For the analysis of field data (SP2), multiple existing datasets related to the assessment of L2H-off systems are considered (cf. Section 3.1). As these datasets already exist, the analyses that can be performed on them are to some extent already predefined or are at least limited in possible metrics. Regarding the harmonization of the results, these data can be used to confirm or at least contribute to the discussion of the results that are generated in the experiments conducted in the field (SP3) or in the simulator (SP4). For the set of relevant metrics this results in a limitation of metrics and their exact design.

As mentioned above, multiple analyses are planned in the field (SP3), namely an expert study and a user survey in the US as well as a FOT in Germany. In contrast to SP2, these analyses are not predefined and therefore the metrics can be chosen more freely. Of course, the design of the function as mentioned before has an influence on the metrics as well as the research questions or hypotheses that are to be answered in these analyses. Here, the harmonization of the metrics can also be used to align these. On the side of measurement equipment, same to similar equipment for the recording of vehicle and environment data is used in Germany and the US, which leads to an alignment between the two data sources. For the surveys conducted in these field experiments as well as in the survey among user in the US, an alignment is achieved by a working group of fka, ika and LfE. It should be noted here that the questionnaires administered can be aligned to some extent, but that there will always be differences due to the slightly different focus of each experiment. This is known to the analysts and is accounted for in the evaluation.

For the driving simulator studies (SP4), direct input for the definition of the relevant metrics is not planned because they are conducted after the other experiments. Their used set of metrics for the analyses and their surveys is therefore based on the set of relevant metrics defined within this work package and the analyses performed in SP2 and SP3.

2.4.3 Relevant Metrics

As an output of this work package, a set of relevant metrics was not defined in detail, as each performed analysis had slightly different requirements on these. However, four categories of metrics are defined, and their harmonization is triggered.

The four categories of metrics are:

- Driver monitoring,
- Function monitoring,
- Driving dynamics and
- Subjective metrics.

The category of driver monitoring contains several subcategories relating to tracking the behavior of the driver. Hand detection has the aim of detecting the position of the hand of the driver from a simple on or off the steering control up to the detection of where each hand is on the wheel. Similar measures are defined for head tracking and body posture. Especially head tracking is detailed further in the subcategory of gaze and pupil detection, which on the one hand aims at finding out where the driver is currently looking at but also contains measures possibly related to fatigue and trust. In the subcategory related to the driver’s activities, metrics for the non-driving related activities (NDRA) as well as function related tasks such as take-over performance can be found.

For the second category of function monitoring, the cycle of activation, deactivation and warnings of the single functions was closely analyzed. The status of the function is the first subcategory identified here, with details on the different levels of the system operation states as well as subsystems included here. This is complemented by the subcategory related to the HMI with details on the information displayed to the driver as well as the modality. Although warnings are often displayed in the HMI, they play an important role in the analysis and are therefore categorized as own subcategory including details on the warning cascade, the timing and modality of the warning as well as the visualization. And lastly, the usage of the function by the user is a metric relevant to the analyses performed.

With the third category of driving dynamics, details on the vehicle and its environment are considered. For the vehicle itself, this is split into metrics on the driver´s input and metrics on vehicle dynamics. Driver's input comprises metrics such as steering or pedal actuation. Regarding the latter, classical vehicle dynamics metrics such as acceleration, speed and yaw rate are located. With regard to traffic, dynamic objects located around the ego vehicle are of interest. For these, measures such as distance, (relative velocity) and safety measures such as TTC and THW are considered. To better analyze the relevance of dynamic objects, information on the lanes is also included. Additionally, the data is supplemented by information drawn from maps such as the speed limit as well as information on the weather including precipitation and temperature.

For the last category of subjective metrics, the aim is the harmonization of the items used within the questionnaires in the expert survey, the user study, the FOT and the driving simulator studies. The aim here is not to harmonize the complete subjective metrics but to harmonize the scales used for the different attributes, as well as the method how data on different attributes are collected. This leads to a basis that can be used and adapted for each of the questionnaires. The subcategories defined for the subjective metrics are sample criteria and assessment criteria.

Within the subcategory of sample criteria, attributes such as age, gender and driving experience are collected, but also attributes relevant for this project such as experience with Level 2 assistance, trust in automation or technology acceptance. For the subcategory of assessment criteria, several types are defined related to the assessment of an event, the HMI, the driver monitoring system together with the state of the driver and the system behavior. Each of the attributes within these subcategories are detailed as far as possible without restricting them from the usage in the different analyses. This means defining the scales, defining the answering options or simply stating that this is a free text answer.

2.5 Review of First Hypotheses on Function Design

Documentation by D. Becker (Institut für Kraftfahrzeuge, RWTH Aachen University)

To conclude the overview on the state of the art, the findings from literature, regulations and analyses of state of the art functions (L2H-on and L2H-off) were reviewed, concentrated, and then translated into a first draft of hypotheses on the system design of L2H-off functions. The according work package within this project was a starting point for further project-internal discussions on how to address the five challenges and questions motivating this project (cf. Section 1). The state of the art was reviewed and discussed under consideration of two lead questions:

1. Driver behavior: Which driver behavior shall be achieved by a suitable system design?
2. System design: Which system design aspects are needed to achieve such a driver behavior?

The motivation for the first guidance question is that the desired driver behavior should guarantee a safe delegation of longitudinal and lateral vehicle control in the context of Level 2 automation. The overview on literature (see Section 2.1 and Section 2.2) and existing L2 functions (e.g. the expert assessment in the USA, see Section 4.2) inspired the following assumptions regarding a desirable driver behavior:

- The driver is physically able to execute the driving task at all times.
- The driver observes the environment and traffic situation in a sufficient manner at all times.
- The driver knows at any time that he/she is responsible at all times.
- When the system presents an attentiveness warning, the driver observes the environment again.
- When a direct control request is issued by the system, an attentive driver intervenes (i.e., executes direct control).

After establishing which user behavior shall be achieved by an L2H-off system, the second guiding question (i.e.: *Which system design aspects are needed to achieve such driver behavior?*) was discussed to help select a collection of possibly relevant aspects for system design investigations. To this end, aspects that came up during the review of the state of the art were categorized (cf. Figure 2-1) to make it easier to identify design aspects that will need to be investigated in the further course of the project.

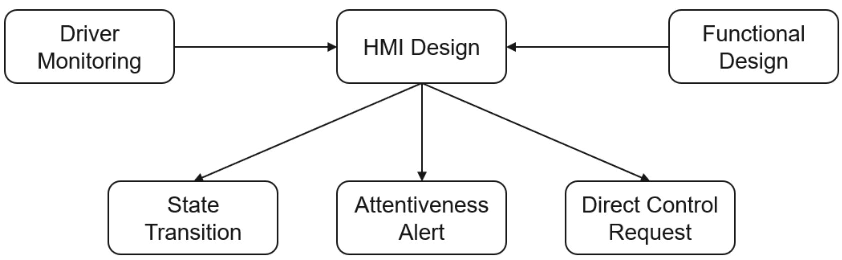


Figure 2-1: Categories of relevant system design aspects

To influence driver behavior, three main aspects need to be considered: “Driver Monitoring”, “HMI Design”, and “Functional Design”. On the one hand, the driver has to be monitored to detect or ensure suitable behavior when the L2 function is active (e.g., CQ1: hands-off=mind-off and CQ3: foreseeable misuse, see Section 1). If the behavior is not as expected, an HMI is needed to inform the driver about necessary behavioral adaptations or consequences. On the other hand, the system communicates its state to the HMI (cf. CQ4: mode confusion). Since the driver has to monitor the vehicles behavior at all times and to intervene whenever necessary, the HMI is considered an important aspect in L2 functions. Thus, the extracted system design aspects for the HMI are further divided into three categories. First, “State Transitions” define how the HMI should be designed to communicate in which mode the driver assistance function is or will be. The other two categories, “Attentiveness Alert” and “Direct Control Request”, comprise aspects that describe the interaction with the driver when he or she is inattentive (based on input of the driver monitoring system; e.g., frequency/timing or modality of alerts) or needs to take over the control of the vehicle (function-issued direct control request; e.g., modality or timing).

The state-of-the-art analysis of these aspects has raised several questions that have to be considered when deriving hypotheses on system design of a L2H-off system. Regarding the functional design, possible questions are: Is the H-off option immediately active after a driver activation or is there a transition phase via L2H-on? When deactivating the L2H-off system, is the longitudinal ACC function still active, or does the function directly switch to L0 (no assistance active)? Other aspects are the concrete physical parameters the system should be implementing such as maximum lateral acceleration or the maximum speed that can be selected. As mentioned above, clear communication to the driver via an HMI system is important to ensure that the driver knows at all times in which mode the function is or which settings rule functional behavior. Further questions arose on, e.g., the timing and amount of inattentiveness alerts if the driver does not monitor the environment sufficiently or in which way the HMI communicates such alerts in different situations of use (e.g., visual, acoustic, haptic alerts). Different approaches for warning cascades were identified by the overview on the state of the art that could be a focus of further investigation. Additionally, it must be decided which criteria the DMS uses to classify driver attention or when and how detected inattention is communicated and regulated by the function.

The considerations described above are the basis for Section 5.1, in which concrete system design hypotheses for the four controlled simulator studies are derived and selected. Following the five challenges and questions behind this project, the focus will be mainly on the interaction of the driver between the L2 system with respect to different driver monitoring systems, HMI concepts, and different system degradation implementations.

3 Analysis of Existing Field Data

Documentation by J. Hiller (Institut für Kraftfahrzeuge, RWTH Aachen University)

L2H-off systems are not completely new to the market and as shown in Section 2.2, vehicles with these L2 systems are available in other countries. The aim of this work package was therefore to tap into those markets and make use of data available from them for the analysis of L2H-off systems. The aim was to substantiate the current hypotheses on system design (see Figure 3-1). Additionally, the field data was to serve as input towards the work done with SP3 (FOT) and SP4 (simulator studies).

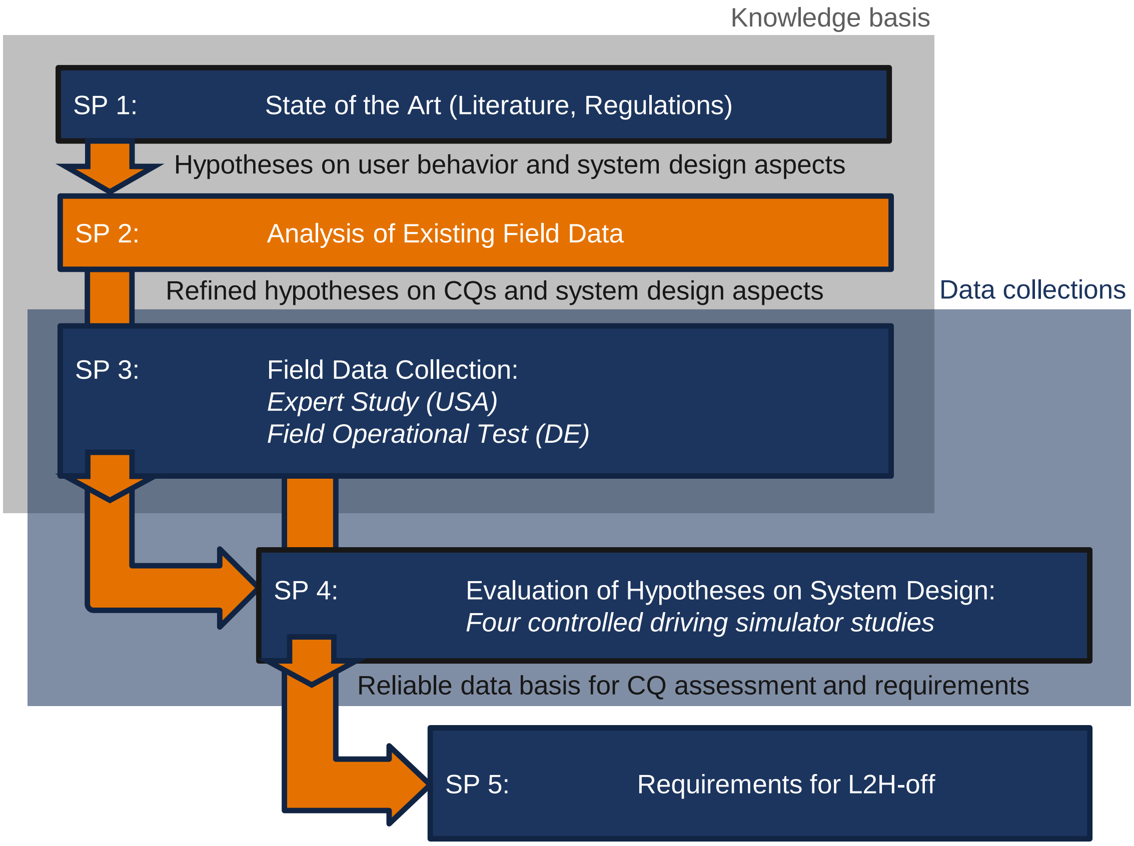


Figure 3-1: Overview on the five subprojects and the role of SP 2 within the project.

Since this subproject was conducted in parallel to the definition of hypotheses in SP1, preliminary hypotheses and research questions were used to kick off the communication and discussions with VDA members on available data. Specifications and requirements on data and sources were discussed in the data acquisition phase.

As a result, the scope of SP2 was extended in a way to also accommodate previously unpublished studies as an input to the generation of hypotheses. Additionally, since two independent data sources were promised, the sufficient provision of field data was declared and confirmed by all project parties.

Due to delays in the provision phase of the data, the focus of the subproject was shifted towards the confirmation of results from the FOTs and simulator studies. Thus, the input to the design of the FOT was cancelled and that to the simulator studies was reduced.

The following section describes the field data examined within this subproject and the concept applied for the evaluation. Afterwards, the analysis of the data and the subsequent hypotheses substantiation is reported in detail.

3.1 Evaluation Concept

Documentation by J. Hiller (Institut für Kraftfahrzeuge, RWTH Aachen University)

For the data analysis, existing field data of VDA members has been made available for this project. Due to company-specific restrictions in terms of confidentiality, the available data differs significantly in terms of type and resolution. While a subset of information is available in both datasets, other signals are only available in one of the two, in different resolutions or with different focuses.

The differences in data has an impact on the evaluation concept and the evaluations performed and reported. To substantially confirm effects seen in the data, only analyses are performed which can be covered by both datasets. This is further enforced by the fact that the data provided for the analysis are confidential and provided for the project under a non-disclosure agreement (NDA). The latter prohibits reporting any effects which can only be observed in one of the two datasets in order to not expose a single data provider or dataset. In the following the evaluation based upon this principle is described. A short introduction to analyses available on one of either dataset are described without going into deeper details. The results cannot be reported due to the NDAs, but they were used within the design of the simulator studies as well as for the evaluation of results in the FOT and simulator studies.

In order to harmonize the analysis, a dataset is created which harmonizes the signal levels and values as far as possible. Within this harmonized dataset, three function levels are defined:

- L2H-off: L2 with hands-off permission
- L2H-on: L2 with hands-on requirement
- Manual: Driving with or without ACC activated

These function levels are available in both datasets. Note that L2H-off also includes driving episodes with hands on the steering wheel (if not explicitly stated otherwise) and L2H-on also includes driving episodes with hands off the steering wheel within the limits set by the system.

The amount of data available in the before mentioned function levels can be seen in Table 3-1. As the data is available over a widespread range of speeds, the availability is only given in time.

Table 3-1: The time spent in each function level in the data provided in the available datasets.

Condition	Time
L2H-off	~ 320 h
L2H-on	~ 160 h
Manual	~ 130 h

The signals available for analysis in both datasets are detailed in Table 3-1. Although the datasets include between over 20 and over 100 signals, the limitations mentioned above reduced the signals for the analyses to these eight signals.

In addition to the signals specified in Table 3-2, further signals for the analyses are calculated. The first such signal targets harmonized attention areas. In general, an attention area specified an area of interest (AoI) to the driver. However, there are differences in the exact specification of those areas between the two datasets. In alignment with the simulator studies (cf. Section 5), the harmonized areas are defined as windshield, instrument cluster (cluster, steering wheel), center stack (radio, navigation, climate control, etc.) and other where the latter includes everything not included in the first three. Using this approach, the areas become comparable between the two datasets.

Table 3-2: Signals available in the datasets for the evaluations.

Signal	Description
Index	Index of the signals
Velocity	The velocity of the vehicle
Attention area	Signal(s) describing where the driver is looking at each timestep
System status	The status of the system with the varying levels of availability and function
Hand status	Hands on and off on the steering wheel
Velocity lead object	The velocity of the object before the vehicle identified as lead object
Distance lead object	The distance to the object in front of the vehicle identified as lead object
Accelerator pedal	Actuation of the accelerator pedal in percent or as absolute value
Steering wheel movement	Movement of the steering wheel by the driver and the system

For the analysis of the criticality of situations, the well-known metrics of time headway (THW) and time-to-collision (TTC) are calculated from the signals provided on the lead object.

Each dataset consists of multiple sessions. Each session is a trip most likely performed by a unique driver. There are no limitations as to the duration of such a session and the sessions are therefore in the range of several minutes to several hours. A detailed overview of session durations and the total duration of data available for the analyses is not possible, as exact timestamps are not available for all datasets.

For the analyses, the data is prepared as a *DataFrame*. This is the internal representation of data in the Python package *pandas*, but in general it can be seen as a table where the signals are represented as columns and the rows correspond to their development over time. A visualization can be seen in Figure 3-1. Within the figure, the total index of the dataframe (“index”) is shown along the other signals. These include the id of the session (“id”), the speed of the vehicle (“v”), the area of interest of the driver (“AOI”), the status of the hands-on detection (“Hands”) and the system status (“System”).

Within this large table, all sessions of the dataset are listed. In Figure 3-1, the change between such a trip is visualized with a change in color. For reference within the dataframe, each session is equipped with a unique identifier.

index	Id	v	AOI	Hands	System	…
…	…	…	…	…	…	…
66	424242	113.8	4	1	2	..
67	424242	113.8	4	1	2	…
68	424242	113.9	4	1	3	…
69	424242	114.0	5	1	3	…
…	…	…	….	…	…	…
…	…	…	…	…	…	…
42	171337	98.3	2	0	0	..
…	…	…	….	…	…	…
			⋮			
…	…	…	…	…	…	…
96	163429	57.4	3	1	4	..
97	163429	57.5	3	1	4	…
…	…	…	….	…	…	…

Figure 3-1: Visualization of the prepared data used within the analyses, the change in color corresponds to a change in the session

One such session is visualized in Figure 3-2. As can be seen, one session can have multiple time frames with different system levels. In this case, there is an alternation between manual driving and L2H-off. Visualized in the lower part of the figure are two other signals that are also of interest for the analysis. As an example, consider the driven speed categorized into bins (i.e. segments in a certain speed range, e.g. from 100 – 110 km/h) and the before mentioned attention areas. For an analysis of the green attention area of the yellow driven speed under the system level L2H-off, this would lead to a small subset of this particular session that would be analyzed.

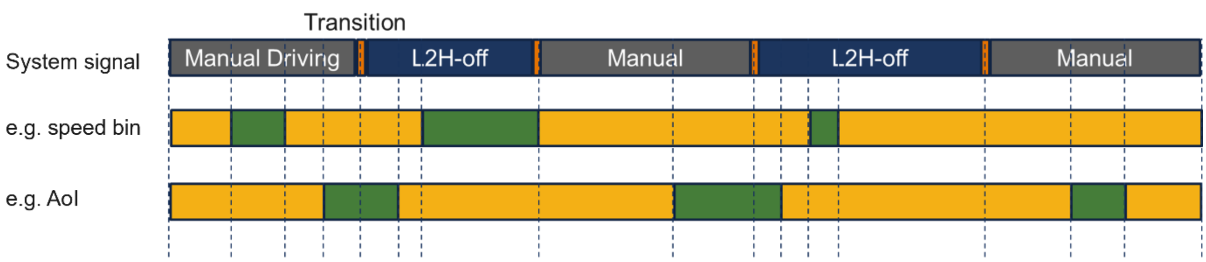


Figure 3-2: Extract of a session with multiple parts with different system status. Provided below in green and yellow are further signals and possible bin values.

This concept of using different conditions and applying it to sessions is used across both datasets and the results are aggregated to generate harmonized results.

For the representation of the results, histograms or the development of single bins over another bin are used. In order to compare the different results, the change between two conditions x and y is expressed as:

$$change[\%] = \frac{\bar{x} - \bar{y}}{|\bar{x}|}$$

Further analyses are provided as descriptive analysis of the data.

3.2 Data Analysis and Substantiation of Hypotheses on System Design

Documentation by J. Hiller (Institut für Kraftfahrzeuge, RWTH Aachen University)

With the datasets provided for the analysis of existing field data (cf. Section 3.1), the analyses that can be performed are mainly limited to an evaluation of the attention areas of the driver and the hands position. As there is no time reference within the harmonized dataset, the analyses are limited to an aggregated comparison. In the following, hypotheses and research questions regarding the attention areas (cf. Section 3.2.1), the hands position (cf. Section 3.2.2) and further analyses performed are presented and discussed (cf. Section 3.2.3).

3.2.1 Attention Areas

As specified in Section 3.1, the harmonized attention areas consist of windshield, instrument cluster (IC), center stack (CS) and other. The first question is whether there are any differences in attention between the different levels of the harmonized function levels (cf. Section 3.1). This is formulated as:

How does L2H-off affect the attention of the driver?

The result of the comparison between L2H-off and L2H-on as well as manual driving is shown in Figure 3-1.

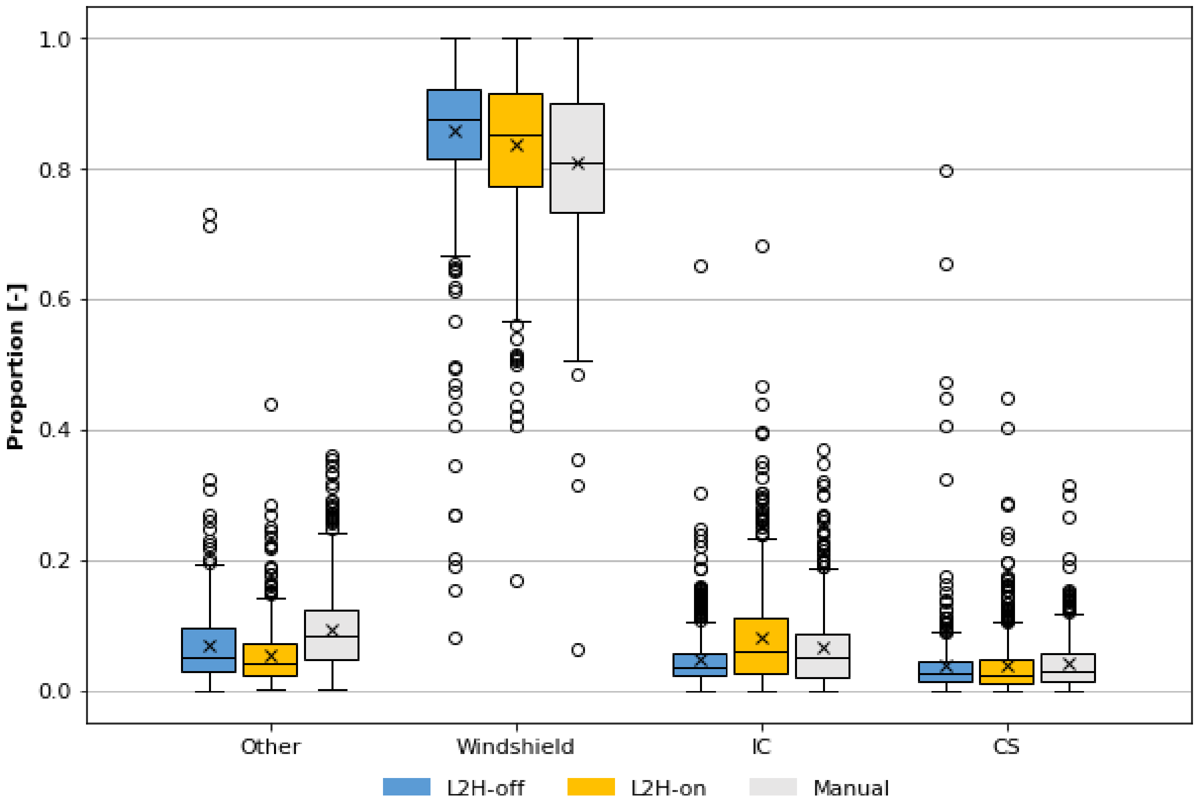


Figure 3-1: The harmonized attention areas compared over the harmonized function levels.

The use of the L2H-off system is associated with a slightly higher attention of the drivers towards the windshield and less attention towards IC and CS. These differences are present in comparison to L2H-on as well as manual driving. For the other attention areas, there is no notable difference between L2H-off and L2H-on, but both attention ratios are lower than for manual driving.

As a further analysis, it is of interest, how the system influences the driver in challenging situations. A challenging driving situation is defined by its criticality. Therefore, the following question can be derived:

How does the criticality of the driving situation influence the attention area of the driver?

As metrics for the criticality, the TTC and the THW are used. In general, it can be said that for both metrics, the number of situations which are critical ($THW \leq 1\ s, TTC \leq 1.75\ s$) is limited (Metz et al., 2019). This is especially true for L2H-off and L2H-on. However, manual driving also doesn't show many situations, but more than during use of the two systems.

For the analysis, the gazes towards the windshield are analyzed, i.e. it is investigated how the attentiveness of the driver changes with the criticality. The results for the TTC can be seen in Figure 3-2.

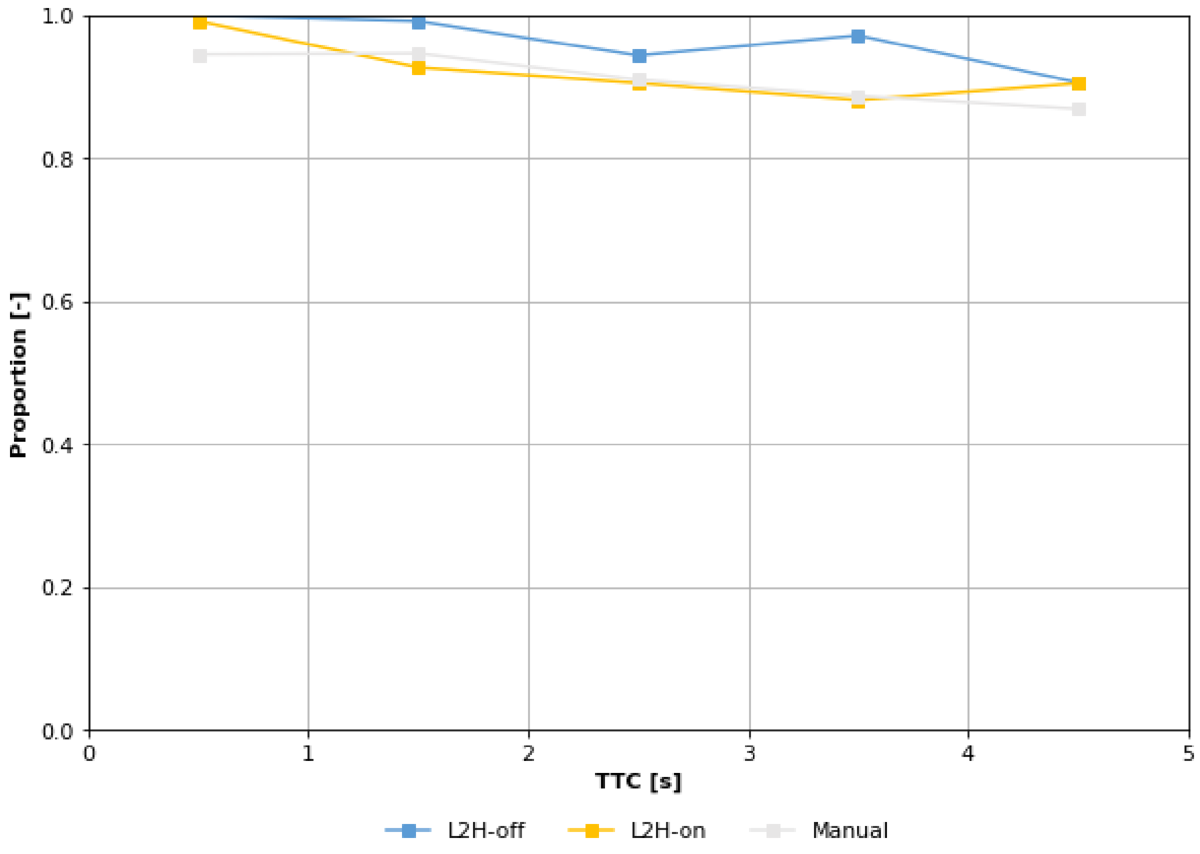


Figure 3-2: The proportion of gazes towards the windshield over the TTC.

The TTC is put into bins covering one second, e.g. from $1 - 2$ seconds. In the field data, L2H-off leads to the driver facing the windshield more in situations where the TTC is lower and the proportion slowly decreasing with higher TTC. The same is true for L2H-on, however, the proportion is lower compared to L2H-off. For manual driving compared to L2H-on, the overall proportion of gazes to the windshield in low TTC situations is also lower. Looking at Spearman's r for the three function levels L2H-off, L2H-on and manual, the values are -0.1420, -0.1355 and -0.2270 respectively. This confirms the slight decrease in the proportion with rising TTC that can be seen in Figure 3-2, but shows no clear correlation.

The number of occurrences in the single bins for the TTC is shown in Table 3-1. Even though L2H-off has approx. twice the driven duration in the dataset (cf. Section 3.1), the number of critical situations is lower compared to L2H-on and even more notable compared to manual driving.

Table 3-1: The occurrences of TTC values in the given bins for the attention area windshield and with the mentioned function levels.

Condition / THW	$[\mathbf{0}\,s\ \dots\ \mathbf{1}\,s]$	$[\mathbf{1}\,s\ \dots\ \mathbf{2}\,s]$	$[\mathbf{2}\,s \dots \mathbf{3}\,s)$	$[\mathbf{3}\,s \dots \mathbf{4}\,s)$	$[\mathbf{4}\,s \dots \mathbf{5}\,s)$
L2H-off	5	6	10	22	55
L2H-on	10	28	77	146	196
Manual	42	133	282	349	416

For the same analysis but with the THW as metric for criticality, the results can be seen in Table 3-1. For this metric, the differences between the different function levels are so small that they are not considered as difference. This means, that there are no differences conceivable in the field data between the different function levels.

Looking at the dependency of the attentiveness from the THW values there are merely slight dependencies that might be seen from the figure. Looking at Spearman's r of -0.0222 (L2H-off), -0.0477 (L2H-on) and -0.0504 (manual driving), it can be seen that there is no correlation between THW values and the gazes towards the windshield. However, for values above approx. 2 s (cf. Metz et al., 2019), the situation would no longer be considered critical, but it shows that there is no difference between critical and non-critical situations.

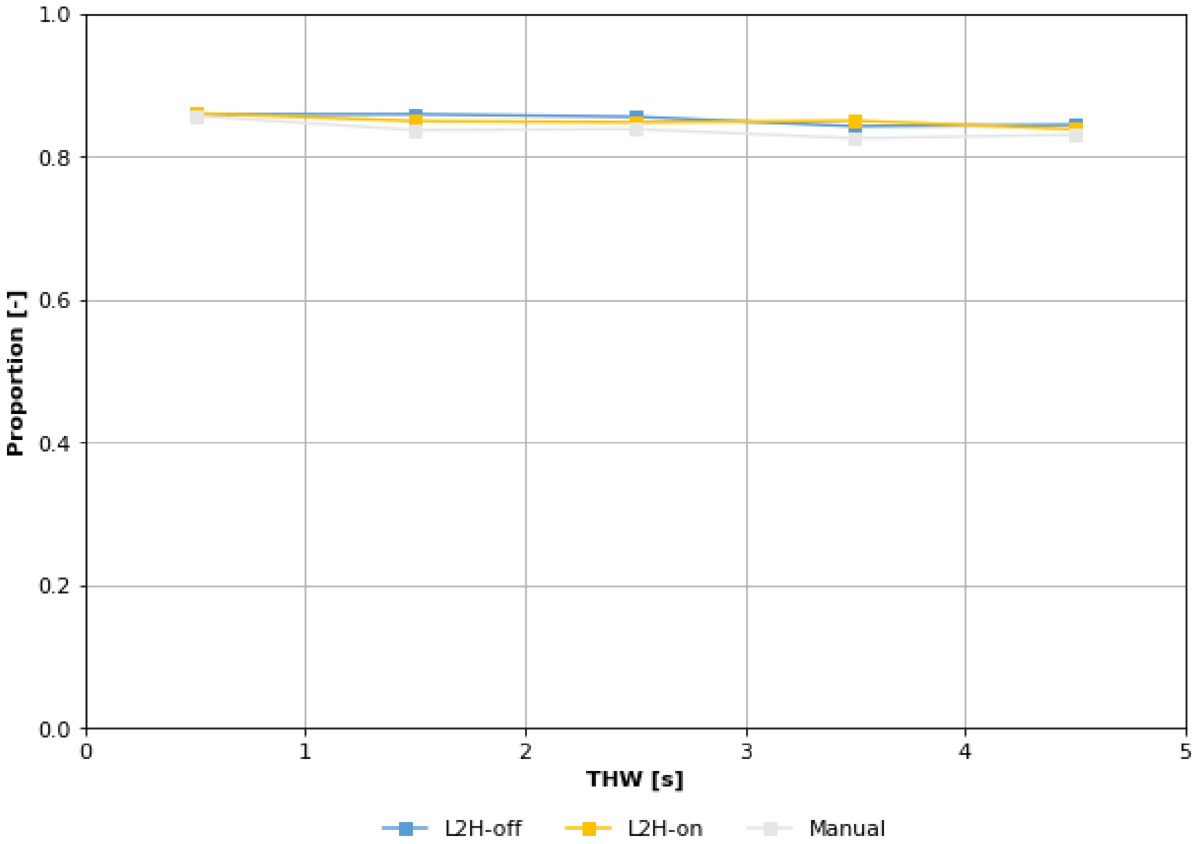


Figure 3-3: The proportion of gazes towards the windshield over THW.

The number of occurrences in the single bins for the THW is shown in Table 3-2. The relative difference is smaller compared to the TTC, but it can still be seen that there are less occurrences for L2H-off compared to L2H-on and manual driving.

Table 3-2: The occurrences of THW values in the given bins for the attention area windshield and with the mentioned function levels.

Condition / THW	$[0\,s\ ...\ 1\,s]$	$[1\,s\ ...\ 2\,s]$	$[2\,s\ ...\ 3\,s)$	$[3\,s\ ...\ 4\,s)$	$[4\,s\ ...\ 5\,s)$
L2H-off	383	523	526	502	457
L2H-on	447	607	596	582	531
Manual	560	641	624	605	578

Overall, L2H-off has at least no negative influence on the assumed attentiveness of the driver and in some cases increases the assumed attentiveness of the driver.

3.2.2 Hand Position

As stated earlier, the hand position is also a driver characteristic that can be analyzed with the available field data. For the analyses regarding the hand position, only the function levels L2H-off and L2H-on are used, as the hand position is only available for these two.

To answer the first question in this analysis, the ratio of the hand positions over the session duration is analyzed. The rationale behind this analysis is to investigate the influence of the system on the hands position. The related question is:

How does L2H-off affect the hand position of the driver?

The result of this analysis can be seen in Figure 3-4. L2H-off leads to the drivers taking their hands off whereas L2H-on has a larger share of hands-on driving. However, the drivers don’t necessarily take their hands off the wheel when they are allowed to which can be seen through the outliers with a hands-off share of close to zero. On the other hand, drivers also extensively make use of the possibility to take their hands off with L2H-on active without the function immediately cancelling and thereby sanctioning such behavior.

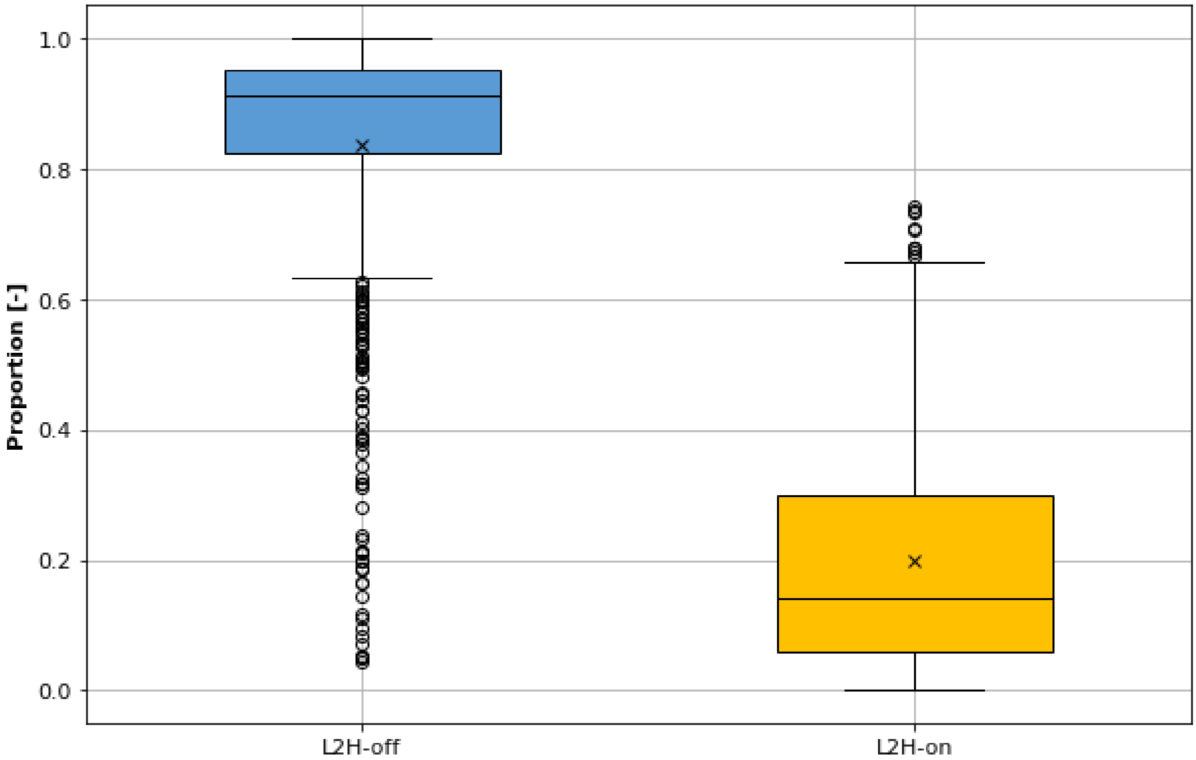


Figure 3-4: The ratio of hands-off per session for the two function levels L2H-off and L2H-on.

Looking at the sessions, where the drivers have their hands on during L2H-off, it is of interest why. Therefore, the question is analyzed if they are maybe more or less attentive in these cases:

Where is the drivers’ attention while having their hands on the steering wheel during L2H-off driving?

The result of this analysis can be found in Figure 3-5. It directly compares the attention shares for L2H-off driving with the additional condition of hands off and on the wheel. Although there are differences between the two, they are small, and it can therefore be stated that the hand position has no influence on the gaze behavior of the drivers.

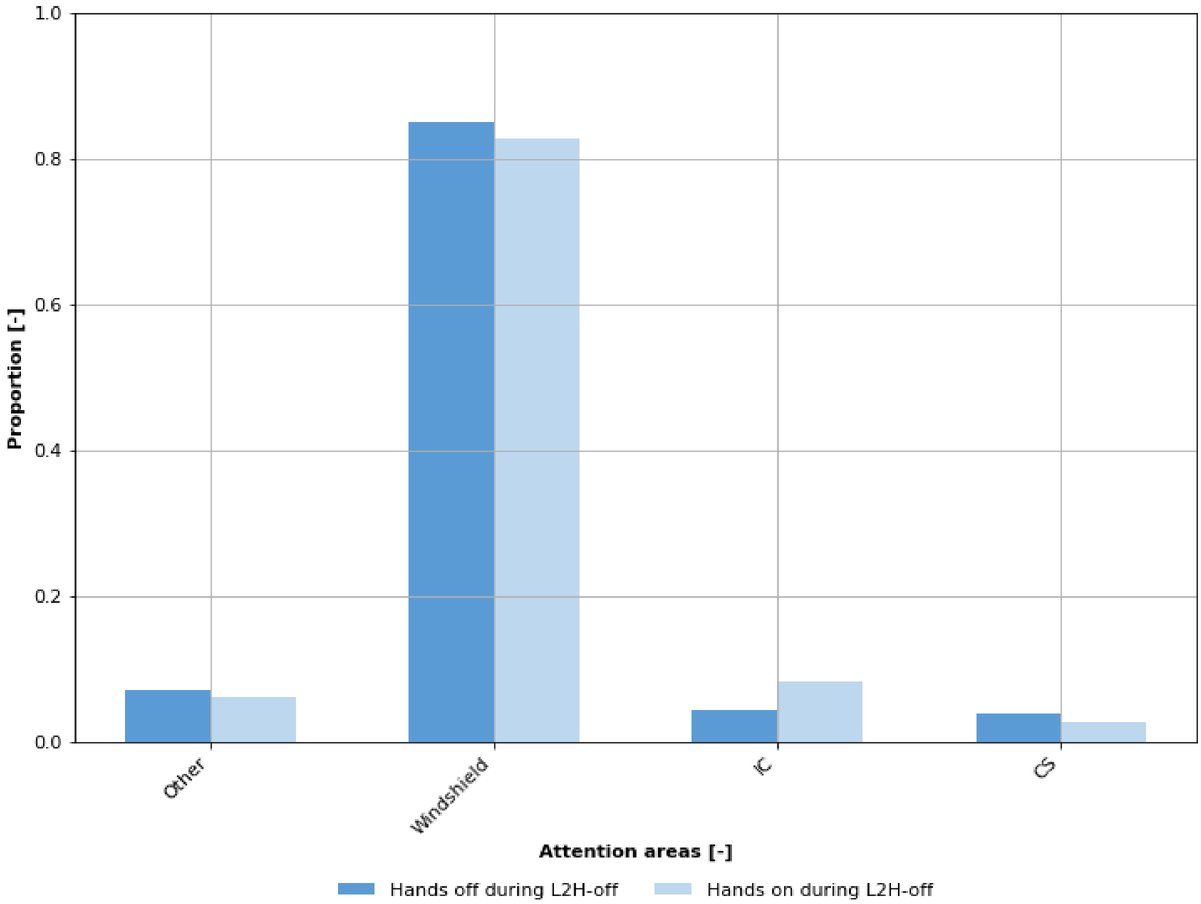


Figure 3-5: Drivers' attention during L2H-off driving in relation to the position of their hands.

As for the attention areas, the analysis of critical or challenging situations is of interest for the analysis of the hand positions. The TTC and the THW are analyzed. Similar to the attention areas, the question is:

How does the criticality of the driving situation influence the hands-off behavior of the driver for both L2H-off and L2H-on?

As for the attention areas, it can be stated that the number of situations that are critical is limited and therefore the results need to be treated with caution.

For the TTC, the proportion of hands-off driving can be seen in Figure 3-6. The few values for the TTC with L2H-off show no clear tendency for the proportion of hands-off with rising TTC values. This is also due to the low proportion in the bin [3 s … 4 s), where there were multiple occurrences of situations in which the driver had his hands on the wheel. For the function L2H-on there are more values and a tendency of drivers to rather take their hands off if the situation is less critical (i.e. higher TTC) is conceivable. This is backed by a Spearman's r of 0.2318 for L2H-on, but no correlation can be derived.

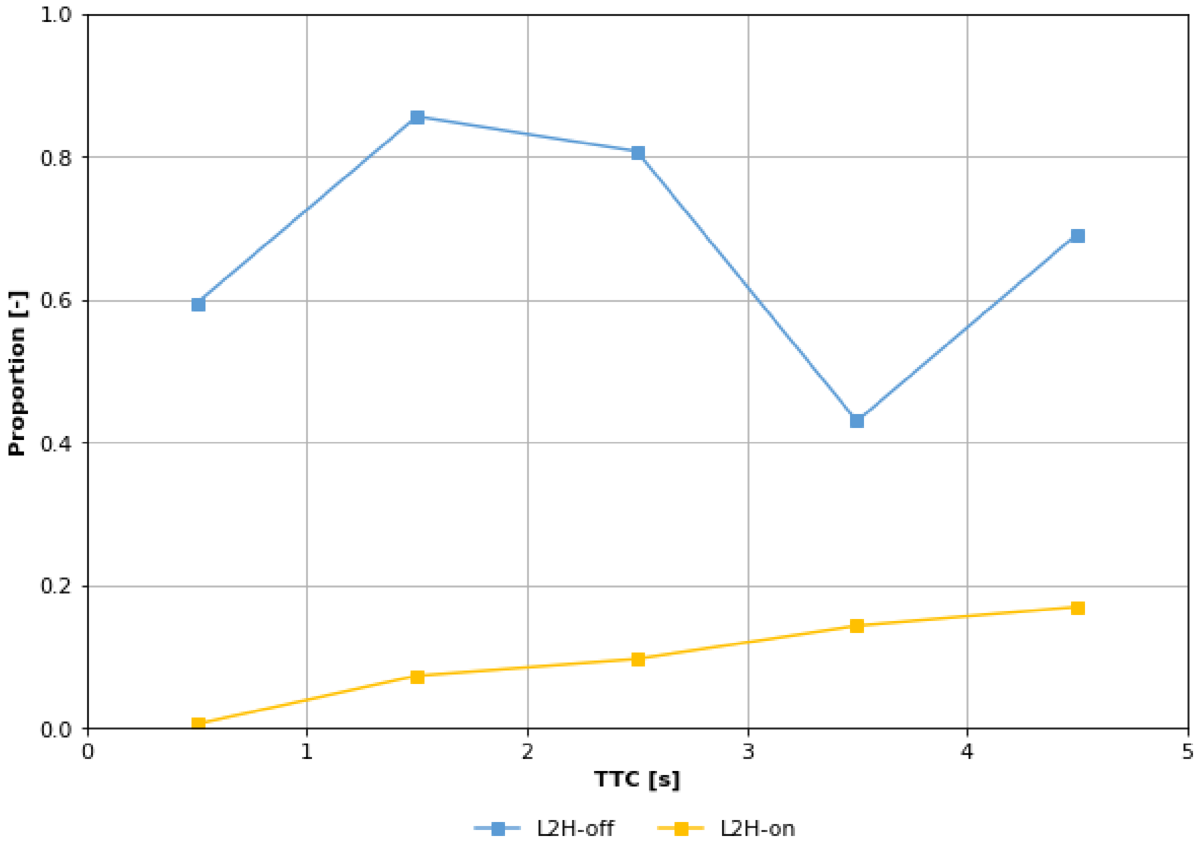


Figure 3-6: Hands-off ratio for the functions L2H-off and L2H-on over TTC.

The number of occurrences in the single bins for the TTC is shown in Table 3-3. For the lowest TTC bin, the number of situations is the same for both system levels, but for higher TTC values the occurrence of critical situations with L2H-off is lower.

Table 3-3: The occurrences of TTC values in the given bins for hands-off behavior and with the mentioned function levels.

Condition / TTC	$[0\ s\ ...\ 1\ s]$	$[1\ s\ ...\ 2\ s]$	$[2\ s\ ...\ 3\ s)$	$[3\ s\ ...\ 4\ s)$	$[4\ s\ ...\ 5\ s)$
L2H-off	6	8	11	24	53
L2H-on	6	31	89	153	205

Looking at the THW, the graph (cf. Figure 3-7) is backed by more data, but drivers do not necessarily change their hands-off behavior with a rising THW. Especially for the lowest THW values with L2H-off enabled, the hands-off proportion compared to higher THW values is slightly lower but for higher THW values it is stable. For L2H-on, a slight tendency to more hands-off with higher THW can be seen. Both results are confirmed by Spearman's r of 0.1020 and 0.0944 for L2H-off and L2H-on respectively, without showing any correlation between the values.

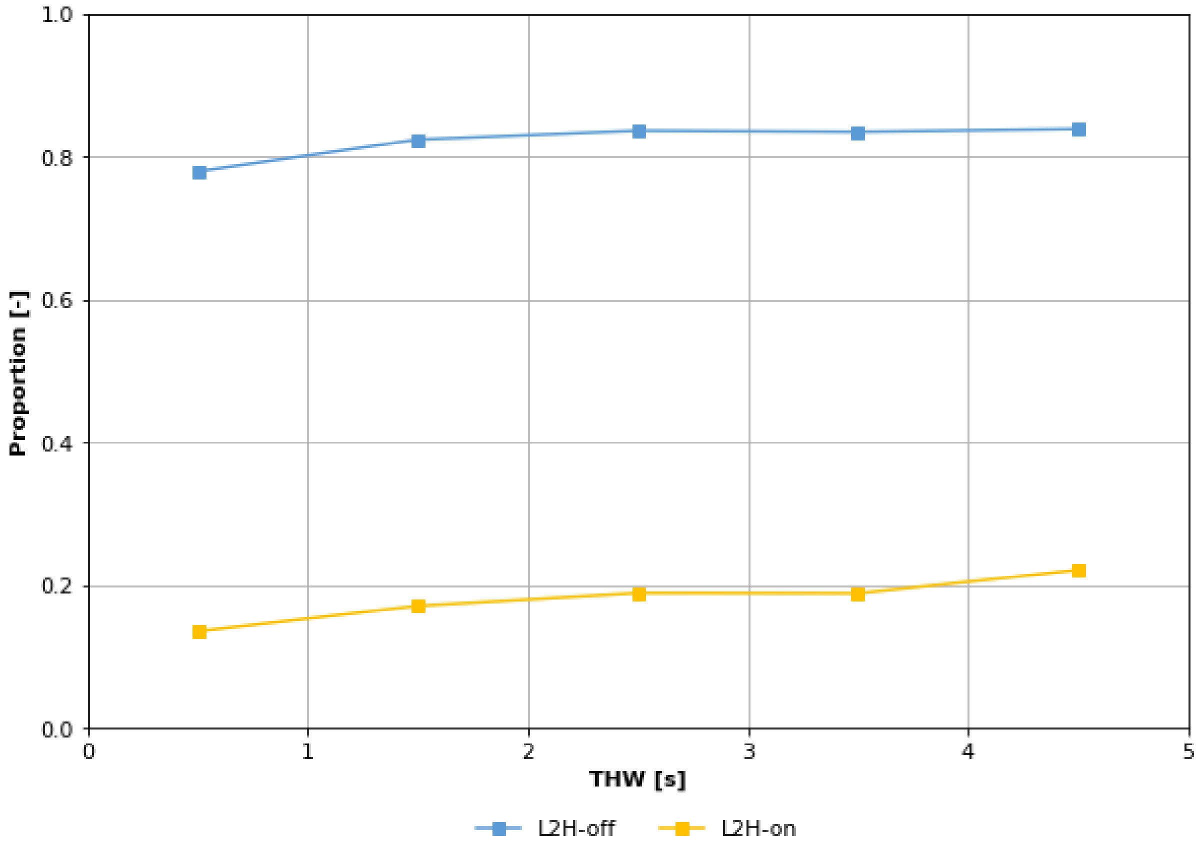


Figure 3-7: Hands-off ratio for L2H-off and L2H-on over THW.

The number of occurrences in the single bins for the THW is shown in Table 3-4. Compared to the TTC for the hands-off position, more situations are available to back the data with L2H-off showing fewer situations, however, the difference is not as notable as for the criticality measured by the TTC.

Table 3-4: The occurrences of THW values in the given bins for hands-off behavior and with the mentioned function levels.

Condition / THW	$[0\,s\ ...\ 1\,s]$	$[1\,s\ ...\ 2\,s]$	$[2\,s\ ...\ 3\,s)$	$[3\,s\ ...\ 4\,s)$	$[4\,s\ ...\ 5\,s)$
L2H-off	408	576	568	547	498
L2H-on	463	613	608	595	548

Overall, L2H-off leads to drivers taking their hands off the wheel more often than with L2H-on activated. At the same time the attention of the drivers does not change with hands-off compared to hands-on.

3.2.3 Aspects Evaluated and Not Reported

As mentioned in Section 3.1, not all aspects of the field data analyzed can be reported.

In addition to the analysis presented in the previous section, an analysis of the L2H-off systems in the data was performed for metrics not reported, resulting in data that was used to design the driving simulator studies.

Among others, the following analyses were also performed on single datasets:

- Take-over times,
- Warning reaction times and driver reactions to warnings,
- Warning modalities,
- Driver input around transitions,
- Coherent gazes towards and away from the windshield,
- Frequency of attention changes and
- Fulfillment of the driving task (e.g. lane centering).

3.2.4 References

Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., et al. (2019). Deliverable 3.3: Evaluation methods. *L3Pilot Driving automation consortium*. Retrieved from https://l3pilot.eu/fileadmin/user_upload/Downloads/Deliverables/Update_07102021/L3Pilot-SP3-D3.3_Evaluation_Methods-v1.0_for_website.pdf

4 Field Data Collection

The third subproject links the state of the art overview conducted in the first two subprojects to data collections on the five potential challenges as conducted within this projects (Figure 4-1). As a starting point, the measurement equipment used for field data collections (Chapter 4.1) has been aligned for all field data collections conducted within this project. In the US expert assessment, current series-production vehicles have been assessed with a focus on potential challenges for interaction and system design (Chapter 4.2). In a field operational test (FOT) conducted in Germany, prototypical L2H-off and series-production L2H-on functions have been assessed by naïve users under the project's research focus (Chapter 4.4). To include input from long-term users of L2 functions, an online survey has been conducted in the USA, targeting the influence of different driver monitoring systems on foreseeable misuse (CQ3; Chapter 4.3).

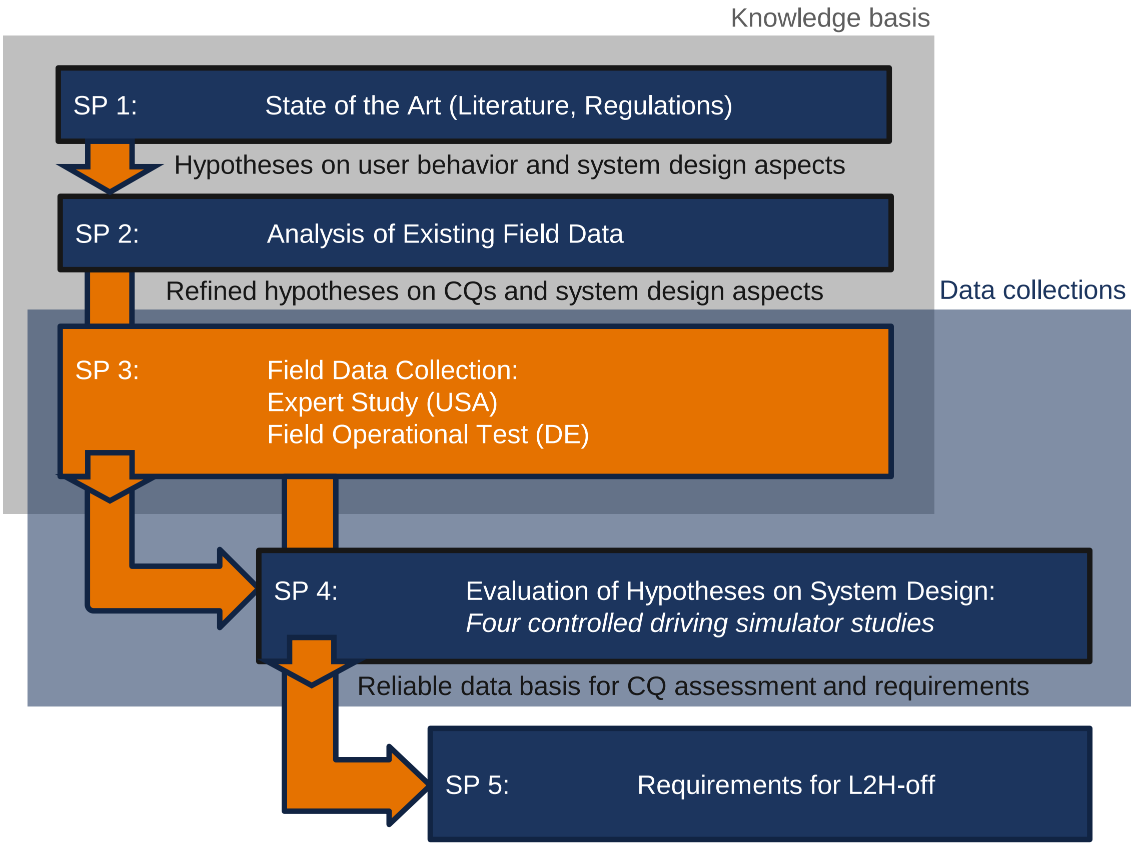


Figure 4-1: Overview on the five subprojects and the role of SP 3 within the project.

4.1 Data Collection Requirements

Documentation by C. Klas (fka GmbH)

The goal of the choice and integration of the data collection equipment was to provide a suitable data source to analyze the driver interaction with the assistance function and possible effects on safety. Therefore, all relevant information about the driver's behavior, the vehicle and function state as well as on the surrounding driving environment shall be provided.

The main considerations for the setup are listed in the following:

- Data on driver behavior shall incorporate information to analyze
    - vehicle guidance operation, especially hands-on/off and interaction with the assistance function,
    - the gaze direction and head turn angle,
    - interaction behavior with relevant vehicle functions,
    - general body posture.
- Data on vehicle and function state shall provide information on
    - Assistance function status,
    - general motion state of vehicle,
    - global location.
- Data on surrounding traffic environment shall provide information to analyze
    - relations to other traffic participants/traffic objects in the vehicle surrounding,
    - position and pose within the current lane and course of the lane.
- Data collection shall be possible to be used in unsupervised test drives and especially
    - activate and run automatically after start of a drive,
    - be able to store data of 1 week with typical driving times (~3h per day).
- Collected data shall be synchronized and be stored with a common timestamp.
- Collected data shall not contain any personal data about individuals who have not consented to a data processing agreement, especially not third parties in the vehicle environment.
- Data collection of all relevant information shall be possible independent of any vehicle-internal signals.
- The collected information shall be comparable between the German FOT and the US Expert Study.
- Equipment shall have a low footprint to not
    - block visibility or influence driving safety in any other way,
    - influence or irritate drivers,
    - impair usability or comfort of the vehicle.
- Equipment shall be robust regarding
    - recording stability to avoid data loss,
    - impairment of vehicle function,
    - mechanical stability.
- Possibility to add markers to the data to

  - indicate audio comments,
  - tag special situations within the data.
- Audio information shall be available
  - to review acoustic signals and warnings,
  - to collect any driver comments.

From these high-level requirements technical specifications were derived. These have then been used for the conception and selection of concrete hardware and software options.

4.1.1 Measurement Equipment

Based on the technical requirements, the measurement hardware and software has been selected and acquired. All vehicles in the German FOT and in the US expert study were equipped with additional hardware to meet the data collection requirements.

The general setup of the measurement equipment for data collection with its main units is shown in Figure 4-1.

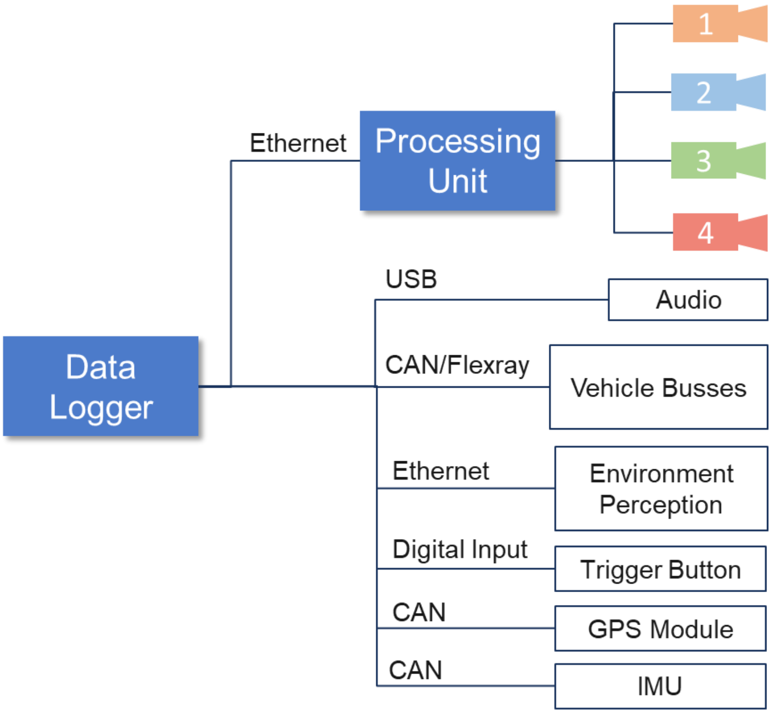


Figure 4-1: Measurement setup overview

The data logger as the central unit collects the data from the different sensors and stores it in a synchronized data stream for later processing and analysis. The main information source for the driver interaction behavior is the interior camera system with four single sensors with different coverage areas. The information about the motion state of the vehicle as well as of the Assistance function could be derived from video data and GPS/IMU information. For efficiency it was also collected from the interfaces to vehicle communication busses. The vehicle surroundings can be analyzed based on the data of a 360-degree LiDAR sensor acting as environment sensor.

The central functional elements of the measurement setup are presented in the following.

4.1.1.1 Data Logger

The datalogging was realized in a Vector CANape log environment based on a Vector VP6450 dedicated logger device (see. Figure 4-2). The logging data was stored in ASAM Measurement Data Format v4 (MDF4) to ensure time consistency of information from all data sources and compact storage of large amounts of information. The data was saved on an 8 TB storage cartridge, which allowed a logging time of one week with about 3 hours of driving per day. The cartridges are removed after logging sessions and the data can be downloaded to a computer.

Figure 4-2: Data logging system Vector VP6450

The logging measurement environment can be configured within the Vector CANape environment on a PC and then transferred to the loggers. The CANape environment can also be used to monitor the function of the logger, download, and inspect the data. Furthermore, the logging configuration can be used in directly on the computer, when the appropriate hardware (e.g. cameras, CAN interfaces, …) is connected. This feature was leveraged in the US Expert Study (see 4.1.3).

4.1.1.2 Video Data

The video data was captured in four synchronous video streams of an Axis FA54 video system. The video system consists of a central processing unit and four camera sensor modules. The sensor modules are capable of collecting video data in all relevant lighting conditions including high-dynamic and low-light conditions. The central unit collects and compresses the data from the sensors and outputs it in four separate video streams over Ethernet. All streams were provided in a resolution of 1080p to ensure good visibility of all relevant details, such as gaze direction, button interaction or cluster display elements.

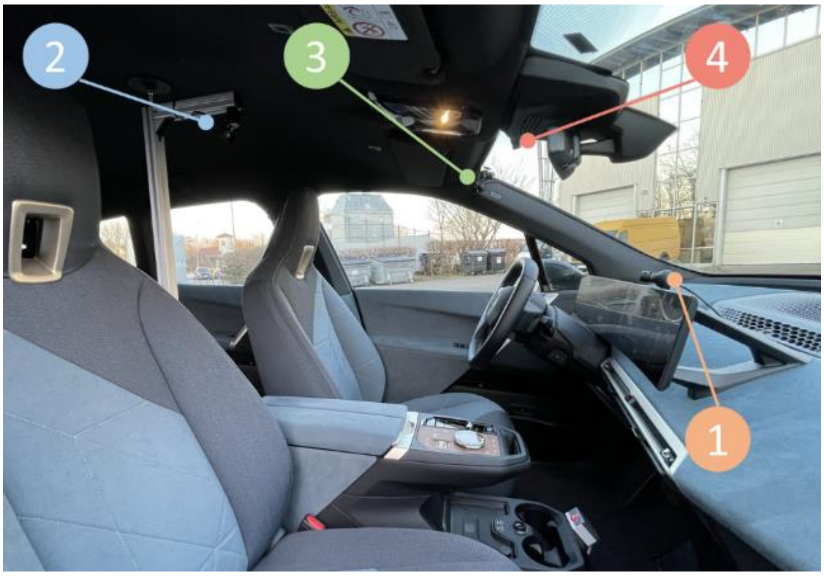


Figure 4-3: Exemplary camera sensor placement (BMW iX), sensor installations and cable routings were done with minimal visibility to avoid irritation or visual blockages for the drivers as seen in Figure 4-3. Robust mounting prevented unwanted alterations of the camera's field of view by drivers, especially during unsupervised test drives.

The single camera views and their purpose are presented in the following sections.

Camera 1: Driver View Direction

Camera 1 is based on an Axis FA4115 sensor unit with tele zoom to provide a view of the driver's head space. The sensor was installed on the right side of the steering wheel as central as possible in front of the driver to ensure good visibility of the driver's gaze direction even at high turn angles of the head. The compact installation was achieved based on a custom mounting bracket for each particular vehicle.

The information can be used to analyze the driver's head position and head direction as well as his/her gaze direction (see Figure 4-4). Furthermore, indications for the driver's state (e.g. drowsiness) and attention can be derived.

Figure 4-4: Exemplary camera sensor placement of Driver View camera (BMW iX)

Camera 2: Driver Body Pose and Interaction

Camera 2 provided a full-body overview based on an Axis FA1105 wide angle camera module. The module was placed on a customized strut installed at the center of the back seat area of the vehicle. The sensor was mounted at an elevated position close to the vehicle roof.

The view provides an overview look of the driver's body posture (see Figure 4-5). Based on this all interaction with all relevant control elements including manipulation of infotainment or comfort system controls can be analyzed. Furthermore, other relevant movements of the driver can be monitored.

Figure 4-5: Exemplary camera sensor placement of Body Pose camera (BMW iX)

Camera 3: Driver Steering Wheel Interaction

Camera 3 was based on an Axis FA1105 wide angle camera module and covered the steering wheel space of the vehicle. The sensor was mounted at a suitable position the top of the inner side of the windshield.

In this way, the position of the driver's hands on the steering wheel, hands-on/-off behavior and steering actions as well as operation of wheel buttons can be closely monitored and evaluated (Figure 4-6).

Figure 4-6: Exemplary camera sensor placement of Steering Wheel camera (BMW iX)

Camera 4: Instrument Cluster

Camera 4 was based on a compact Axis FA1125 pinhole sensor module with wide viewing angle. The sensor was placed at a suitable position to cover the instrument cluster of the vehicle. Depending on the specifics of the vehicle, this position was at the upper windshield or directly at either side of the instrument cluster (see Figure 4-7).

The information of this camera can be used to analyze and verify the state of the ADAS system and any driving notifications and warnings as a redundant information to the signals from the vehicle's communication busses.

Figure 4-7: Exemplary camera sensor placement of Instrument Cluster camera (BMW iX)

4.1.1.3 Environment Sensor

A LiDAR sensor was used to collect information about the vehicle surrounding traffic environment. Based on the main operational domain (highways and similar roads) a 360-degree scanning LiDAR sensor with long detection range was installed on the vehicle roof.

For this purpose, a customized roof rack for each vehicle was installed along with a mounting bracket that allowed angular and height adjustment. Connection cables were integrated in a way to not impair usability, waterproofness or visibility of the vehicle. An example is shown in Figure 4-8.

Figure 4-8: LiDAR sensor installation on roof rack (BMW iX)

The sensor, namely Ouster OS2, provides 3D point cloud data with 128 channels for vertical resolution and 1024 rotational steps for horizontal resolution. This data is output via an Ethernet connection along with an exact timestamp for each frame. The logging software collected the raw data stream to achieve compact storage footprint.

Based on the point cloud data two main information sources can be provided. The first is a camera-like 360-degree view, which allows visual inspection and understanding of the surrounding traffic situation, as shown in Figure 4-9. At the same time, the resolution of this camera-like image does not enable recognizability of any personal data.

Figure 4-9: 360-degree image representation of LiDAR point cloud data

Secondly, the point cloud data is processed to deliver information about surrounding traffic objects (relative position, relative velocities, type, lane assignment, …) and lane marking information (relative position, relative angle, …).

An example visualization of the processed data can be seen in Figure 4-10. On this basis, objective metrics, like Time Headway (THW), Time to Collision (TTC), or Time to Linecrossing (TLC), can be calculated for the relevant situations.

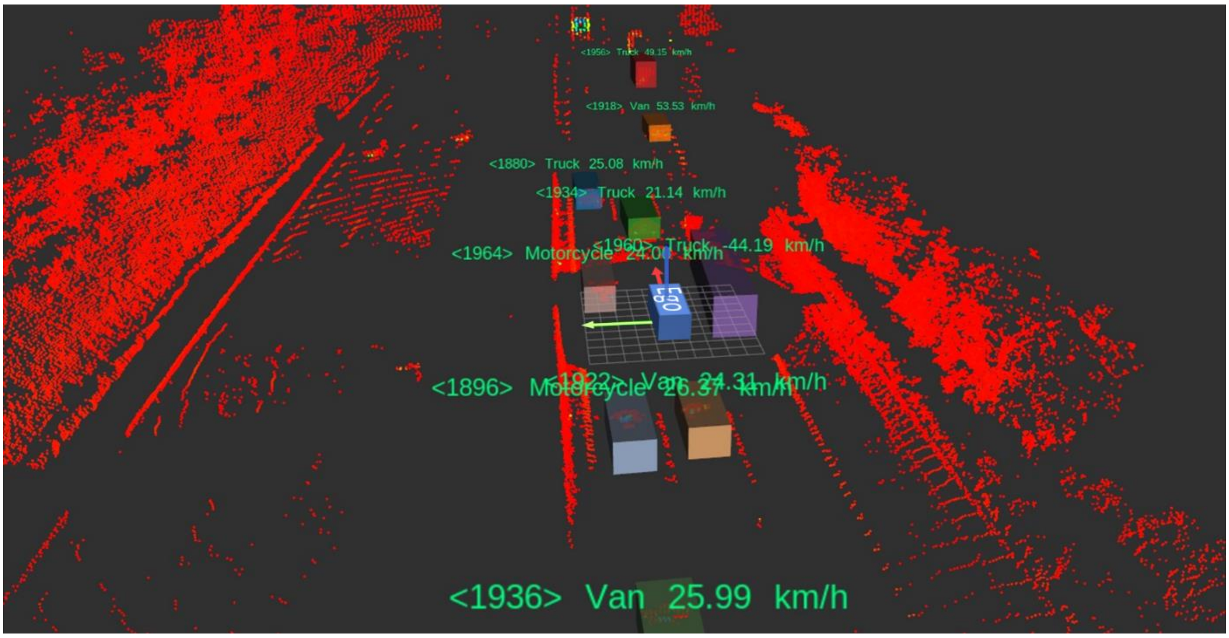


Figure 4-10: Exemplary visualization of extracted object information

4.1.1.4 Vehicle Data

One of the goals was to be able to collect the relevant information completely independent from the vehicle sensors or systems. However, for some well-defined signals, it seemed efficient to collect those based on signals from the vehicle communication busses, especially from FlexRay and CAN. This helped especially to avoid complete visual inspection of the extensive amount video data in order to derive many of the signals.

These signals did especially not cover any data from environment sensors to rely on for calculation of situational criticality measures. A list of the acquired signals is presented in Table 4-1.

Table 4-1: Overview of acquired signals from vehicle communication busses

Signal	Description
Driver Interaction – Primary Driving Task	
Accelerator pedal value	Percentage, voltage, etc. to evaluate driver request
Brake pedal value	Percentage, brake pressure, etc. to evaluate driver request / function cancellation
Brake light switch	Boolean value to evaluate driver request / function cancellation
Steering wheel angle	Assessment of current driving situation, e.g. at time of handover
Steering wheel torque	Assessment of driver's interaction, oversteering behavior
Driver Interaction – Secondary Driving Task	
Wiper control interaction	Driver interaction with wiper controls
Turn signal control interaction	Driver interaction with turn signal controls, e.g. brief/long push, can be used to monitor any interruption/alteration of the automation function
Driver Interaction – Tertiary Driving Task	
Control interaction Infotainment	Any interaction (non-specific) with control elements of infotainment or comfort functions
Vehicle State – Vehicle Motion State	
Longitudinal velocity	Vehicle velocity, true and/or displayed to the driver
Yaw rate	Vehicle motion information to identify specific driving maneuvers
Longitudinal acceleration	Vehicle motion information to identify specific driving maneuvers
Lateral acceleration	Vehicle motion information to identify specific driving maneuvers
Vehicle State – Automation/Assistance Function	
Driving function state	Current status of the driving automation function, may be a combination of multiple signals, e.g. for longitudinal and lateral guidance function
Hands-on detection	Torque-based and/or capacitive hands-on detection
Operation function-related buttons	Driver requests regarding the automation function, such as activation, cancellation, set, resume, etc.
Output Optical Warning	Function-related warning or information output by optical elements/symbols or in textual form, for quick identification of possible relevant events
Output Acoustic Warning	Function-related warning or information output by acoustic signals, for quick identification of possible relevant events
Vehicle State – Other	
Environment illumination	Direct measurement for identification of environmental lighting conditions
Lighting system state	Driver interaction with light system controls, indirect information about environmental lighting conditions
Low beam state	Indirect information about environment illumination in case of automatic light control
Wiper state	Indirect information about precipitation in case of automatic light control

In general, all particular signals could be and were verified by comparison with the redundant information from the additionally installed sources, especially by visual inspection of the video material. Direct collection of this information allows quick automated analysis and filtering of the data to support the identification of relevant situations or events. An additional verification of the vehicle signals could then be done based on the redundant information for the identified snippets of the data.

The generic signals were discussed with the contact persons at the vehicle providers of the project. Suitable, comparable information sources and concrete signals were identified. The necessary information about hardware and software interfacing (especially suitable tapping points and encoding databases) for the in-vehicle communication busses was provided and supported, where necessary, by the particular vehicle manufacturers.

4.1.1.5 Other signals

An additional sensor system was employed to collect the motion data of the vehicle. The microcomputer system PCAN-GPS by Peak-Systems can provide data from precise 3-axis accelerometers and gyroscope sensors. Furthermore, it delivers precise information from GNSS (GPS, Galileo, GLONASS, QZSS, SBAS). This information was provided to the data logger via CAN interface. The data was used as redundant information of the vehicle motion within the German FOT. For the US Expert Study it was the main information source.

The audio stream was collected synchronously directly by the data logger. Depending on the conditions in each vehicle, an omnidirectional microphone was installed in an area close to the driver's head. With this information any acoustic signals of the Assistance function could be identified and reviewed. Furthermore, comments of the driver in special driving situations or verbal assessments of the experts within the US study could be collected.

A trigger button was placed in an area of the cockpit, which is well reachable for the driver and a possible supervisor on the passenger seat as shown in Figure 4-11.

Figure 4-11: Trigger button for marking of special situations (Porsche 911)

The button press information was provided to the data logger via CAN interface to be stored synchronously with all other data streams. Based on this trigger, special situations along the test drives could be marked. This could be e.g. traffic-related or system-related events, indication of audio comments as well as measurement information (like start of actual test phase of a drive).

The distinction of different markings could be made by pressing patterns (e.g. long/short/double press). The feedback for the driver or supervisor about the stored signal was provided via different blink codes of the button light.

4.1.2 Vehicle integration

The focus was to ensure mechanical and electronic robustness as well as the lowest possible footprint in the sense of visibility blockage and driver/passenger usability. The main portion of measurement equipment and necessary peripheral devices were integrated into the vehicles' trunks. Therefore, the hardware was integrated into stable, compact racks, which were mechanically fixed at a defined position (see Figure 4-12). Reachability for sanity checks of the most important components and the exchange of data cartridges was ensured.

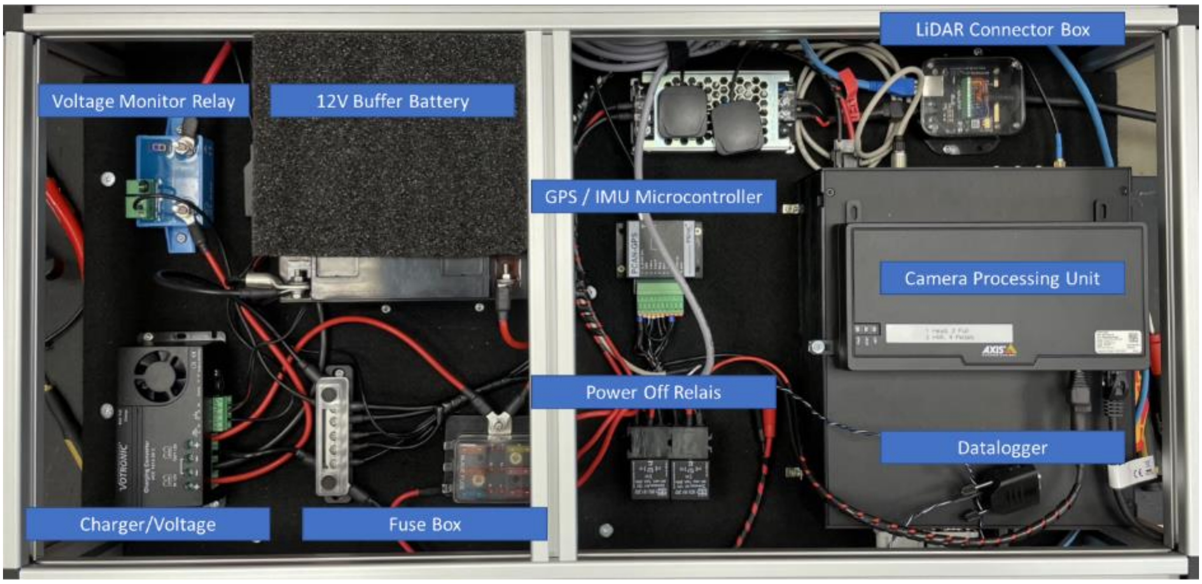


Figure 4-12: Measurement equipment rack for installation in vehicle trunk

The 12 V on-board power supply of the vehicles was utilized as the energy source to power the measurement equipment. To ensure reliable operation of the vehicle function, a voltage monitoring relay was implemented to disconnect the additional equipment before discharging the vehicle starter battery. A battery charger device was implemented to supply a buffer battery and disconnect the additional equipment from the vehicle battery in case the KL15 signal is not present (*Ignition switched off*). The equipment was switched off with a 2-minute delay after turning off KL15 to not turn off measurements for only short stops. After this a shutdown signal was set, to safely power down the devices without losing data.

Aside from collecting and providing the internal sensor data as well as the user button inputs, the GPS/IMU device was used to coordinate the power down behavior of the system.

Furthermore, a central integration step was to establish connections to the specific tapping points of the vehicle bus systems. The hardware interfacing was achieved based on existing breakout connectors or suitable connectors were integrated based on the information provided by the vehicle manufacturers. The cables were routed in a hidden manner towards interface modules of the datalogger system.

4.1.2.1 Overview of equipped vehicles

In total, six vehicles were equipped with the described measurement hardware. The vehicles were prepared and fitted at fka’s workshop in Aachen, tested with regard to their logging functionality and then transferred to LfE in Garching.

Figure 4-13: Overview of equipped Ford Mustang Mach-E

Figure 4-14: Overview of equipped Ford Focus

Figure 4-15: Overview of equipped BMW iX, identical for Hands-on and Hands-off version

Figure 4-16: Overview of equipped Volkswagen Passat GTE

Figure 4-17: Overview of equipped Porsche 911 Cabrio

4.1.3 Adaptations for US Study

4.1.3.1 Preconditions

The data collection within the US Expert Study was conducted at an earlier point in the project. As preparation for the more extensive German FOT, it allowed to verify the measurement concept and data quality. At the same time, the collected data could be used to implement and test the necessary data processing tools. Thus, an identical information basis like for the FOT in Germany could be created.

The design of the study, especially the overall number and duration of test drives, number and brand of vehicles and the location of the drives posed some differences in the preconditions for the data collection.

- Due to the chosen set of vehicles, vehicle-internal communication systems were not accessible.
- Data recording could be conducted and monitored by supervisors, who accompanied the drives.
- The overall duration and number test drives was comparably low, so that manual extraction of certain signals, especially about the system state and driver interaction, from the video stream was feasible.

4.1.3.2 Adaptations of the measurement setup

To collect a similar information basis for both studies the measurement setup for the US study was identical for the most part. The adapted setup can be seen in Figure 4-18.

The same camera system, environment sensors and IMU/GPS module were employed for data collection. As one main difference, the data was recorded in a Vector CANape environment on a laptop computer, which was operated and monitored by a supervisor during each complete test drive. Furthermore, there was no connection to vehicle communication bus systems.

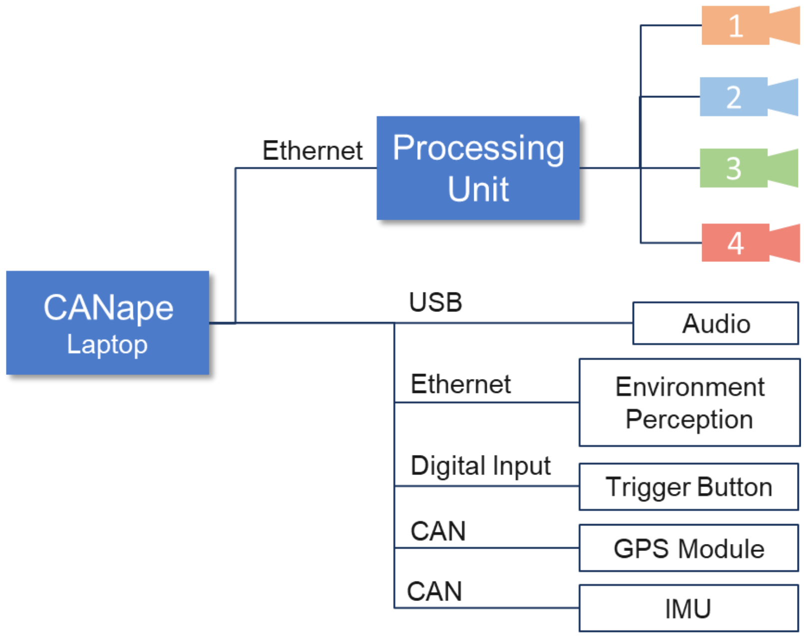


Figure 4-18: Adapted measurement setup overview

4.1.3.3 Adaptations of information sources

The main difference in the set of information sources was the unavailability of signals from the vehicle communication busses. Therefore, the information on the vehicle motion state, the state of the Assistance function and the driver interaction needed to be extracted from other sources.

The vehicle motion state could be reconstructed completely from the data, which was provided by the PCAN-GPS module with its GNNS and IMU information.

For the review of driver behavior and interaction, the camera positions of the cameras 1-3 were generally at the same positions like in the German FOT. For efficient information collection about both the system state of the Assistance function as well as the hands-on/off behavior and steering wheel interaction the coverage of camera 3 was slightly different from the German

FOT. The vehicle sizes and the involvement of only expert drivers allowed flexible placement of the sensor.

Figure 4-19: Camera views within US Expert Study

As the pedal interaction, especially the cancellation behavior by braking could not be observed by any other source, camera 4 was placed within the driver footwell area to provide indications about the pedal utilization.

4.2 US Expert Study

Documentation by Ph. Seewald, J. Josten (fka GmbH)

Although L2H-off systems are already available in some markets, Level 2 automation in Europe and Germany is only available in the form of L2H-on systems. To be able to incorporate an assessment of the current state of L2H-off series-production vehicles into this project, field tests were planned and conducted with non-prototype series production and road-legal L2H-off systems in the US. The US expert study can be considered a combination of a practical state-of-the-art assessment and a first stage of data collection within the project.

Overall, the US expert study shall provide insights on state-of-the-art systems (L2H-off and L2H-on) for transfer into a generic function design to be used in the driving simulator studies. Furthermore, the five challenges and questions related to hands-free driving shall be evaluated and prioritized based on hands-on experience with current series production vehicles. The focus lies on the assessment of subjective data and the derivation of “Do’s and Don’ts” for the system design and operational design domain (ODD). Although objective data is collected as a proof of concept for the field operational test in Munich, it is used only for the detailed analyses of selected scenarios. The results are compared to the state-of-the-art as derived from literature and other (normative or analytical) sources and serve as a basis for the driving simulator studies in SP4.

4.2.1 Test setup and procedure

For the field assessment, three test vehicles were rented and prepared for testing. The vehicles were selected according to differences in the L2 automated driving functions on board, e.g., regarding ODD or driver monitoring system (DMS) features. All vehicles were in series production and publicly available, with the corresponding state-of-the-art automated driving features legally approved in the USA. The selection of L2 systems allowed testing of both L2H-on and L2H-off functionalities, as one vehicle allowed both the hands-free and hands-on use with different DMS criteria. For L2H-on systems, the ODD ranged from widely unrestricted to road type-based. On the other hand, the ODDs of the L2H-off systems allowed comparisons between a map-based architecture and conditions set by current speed, road type and surrounding traffic.

The different systems under investigation featured different technological solutions for similar DMS. Thus, hands-on detection using torque as well as capacitive sensors was included in the selection of vehicles. All L2H-off systems employed a camera-based eyes-on detection with differences between systems in the position of the driver camera as well as the criteria for driver monitoring.

Each vehicle was equipped with additional sensors such as a roof-mounted Lidar sensor, a GNSS receiver, an accelerometer, one central microphone to record the driver’s comments as well as four digital video cameras to capture the driver’s actions and HMI signals such as warnings and system status. Apart from continuous data recording, which was controlled by individual laptops, drivers and front-seat passengers could always create markers within the data stream by pressing a trigger button to highlight scenarios or connect comments to events. The entire test setup was not only implemented for data collection, but also used as a rehearsal of the sensor configuration in preparation for the FOT in Germany (Section 4.1 and Section 4.4).

Testing involved five experts from Germany representing competences in ADAS and automated driving, automotive engineering, ergonomics, and user experience. At least one expert represented each of the entities involved in the project to ensure the transfer of knowledge and experience into later WPs. Overall, 3-4 hours of driving time were scheduled per expert and per vehicle. Depending on the ODD, different test routes were assigned to the vehicles. The main test route (Figure 4-1) featured a round track of about 306 km (190 miles) on highway roads, composed to address systems with map-based ODDs and road type restrictions to highway roads. Furthermore, the main test route included several pre-defined points of interest like e.g., interchanges, sharp bends and estimated ODD limits for post-hoc analysis. For analysis, only data collected on highways and interstates, i.e., collected within the project-relevant ODD, is considered.

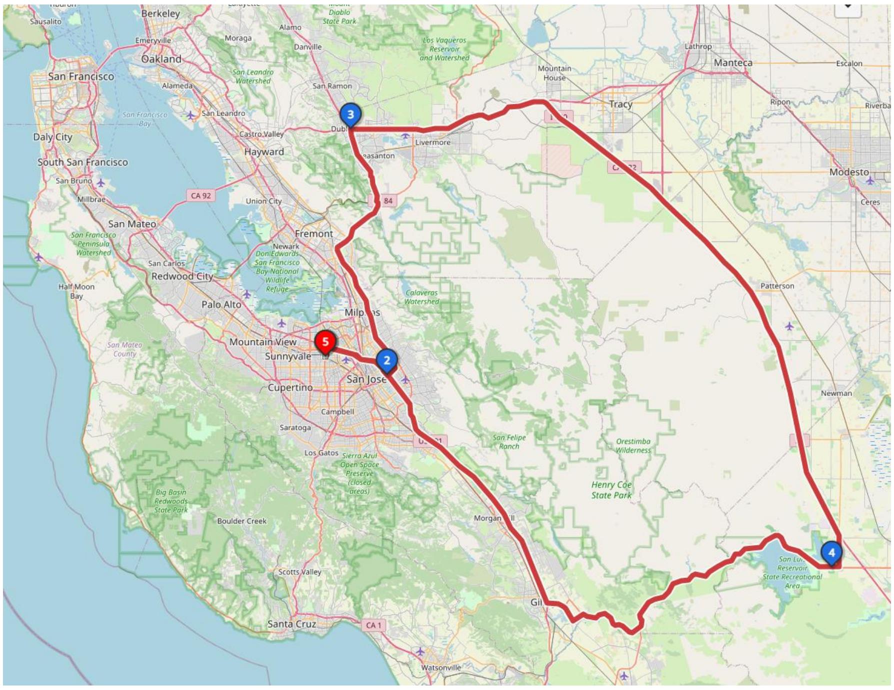


Figure 4-1: Main test route (focus on highway ODD) (Source: © openrouteservice.org by HeiGIT | Map data © OpenStreetMap contributors)

Vehicles using L2H-off systems with ODD restrictions on travel speed, road type and surrounding traffic were tested on alternative test routes (Figure 4-2). Test drives on alternative routes were scheduled preferably during the morning and afternoon rush hours when commuting traffic peaked.

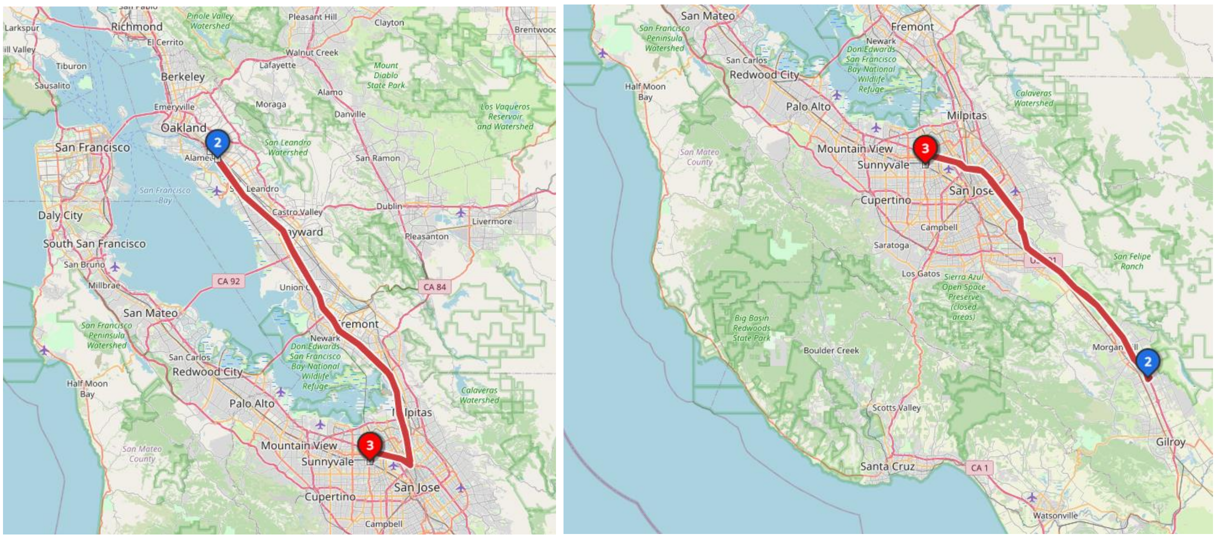


Figure 4-2: Alternative routes (focus on traffic density) (Source: © openrouteservice.org by HeiGIT | Map data © OpenStreetMap contributors)

Testing was carried out in December 2021 over the course of two weeks. Each test run was conducted with the same procedure. Before each drive, the expert received a general briefing on how to operate the L2 system at hand based on a condensed version of the vehicle owner's manual. Furthermore, general instructions were given introducing aspects of interest that should be addressed during the drive (such as route details and test cases)). At the beginning of each test drive, a calibration procedure for the camera system was conducted. The cameras have been installed to link the driver's focus of attention with HMI messages related to the DMS.

A co-driver sitting in the front passenger seat accompanied each expert during the drive. The co-driver instructed defined test cases during the test drive in order to enable a standardized basis for the subjective assessment of DMS and HMI functionality, functional performance and ODD limits. The list of DMS test cases included up to 12 visual areas of interest on which to direct the view, e.g., rearview mirror, center console and passenger. The purpose was to enable the experts to assess the field of view deemed acceptable by the attention-based DMS in direct comparison to the subjective overview of the traffic situation associated with these different gaze areas and the corresponding HMI design. Additional test cases included different actions for overriding and overruling the function in regard to the resulting fallback level (L1 or L0). In the case of system-initiated deactivations, experts were told to assess the comprehensibility of the system's actions and the communication of the driver's tasks. During each test case, the co-driver additionally observed the surrounding traffic and the function's behavior for safety reasons.

Ratings and comments could be highlighted within the data stream by pressing a trigger button. Instructions before the test drive for the use of the trigger button stressed take-over situations, beginning and end of test cases, safety-critical behavior of the function or any safety-critical situation encountered as well as system behavior that was deemed relevant to one of the five challenges and questions.

After the test drive, a post-drive interview was conducted in order to collect an overall system assessment (i.e., qualitative feedback and rating scales) with a focus on the five challenges addressed within this project. The interview was recorded for transcription and subject to a structured analysis. The relevance of potential challenges for each function was derived from specific questions, e.g., relating to the potential for misuse (CQ3) or to neglecting the monitoring task (CQ1) with each of the different DMS. Additionally, the interview covered the different design aspects as derived from SP1. This included questions relating to the design of the DMS, the functional design and the HMI design, i.e., alerts and warnings, the communication of states and transitions as well as take-over requests. After testing two different vehicles, a direct comparison between different functions was encouraged in the interview, formulated as general *Do* (positive aspects) and *Don't* (negative aspects) statements on system design derived from test cases during the drives.

4.2.2 Results

The focus of the analysis lay on qualitative data, i.e., on the subjective assessments of the five experts. Primarily, the goal was to complement the FOT conducted in Germany with a focus on the interaction behavior of naïve drivers (Section 4.4) with a structured assessment of the relation between each challenge and the driver's hand posture. Where possible, assessments were related to specific aspects of the different, state-of-the-art DMS designs. In addition to subjective data, Lidar data and gaze data were analyzed for specific use cases to gather input on functional design and DMS criteria.

In total, over 45 hours of driving data were recorded, resulting in more than 1.7 TB of video and Lidar data. Post-processing of objective data involved a detailed analysis of all situations in which the trigger button was pressed. Each event was reviewed individually and tagged according to the current situation based on video footage from inside the vehicle (e.g., HMI, driver behavior and posture) and the experts' comment. Overall, 22 different types of tags were assigned to classify each trigger button event within the drives by up to three categories, referring to HMI-, DMS-, system- and CQ-related issues.

Based on the tagged and annotated driving data, further analyses were conducted. For all relevant situations, driver gazes, analyzed based on video data, were evaluated to assess the characteristics of the corresponding DMS of the L2H-off system such as detection range, monitoring criteria and timing of alerts as well as warning cascade and system degradation. The results were used as a reference for the DMS design used in the subsequent driving simulator studies (see Section5). Furthermore, driving situations of interest were extracted according to the assigned tags. Lidar data on surrounding traffic in situations classified as potentially safety-relevant was analyzed. The analysis was conducted in the full understanding that some situations were induced by challenging the L2 function in the expert study and do thus not reflect the behavior of naïve drivers in daily traffic. Naïve drivers might, e.g., already react to anticipative cues or deactivate the L2 function in challenging conditions.

Apart from the recorded driving data, interview data provided a substantial part of the insights drawn from the US expert study. In addition to live comments, over six hours of interview data were recorded and transcribed during the post-drive interviews. In a following step, transcriptions were structured according to the project's questions and challenges and additional criteria, similar to the tags applied to the situations marked by the trigger buttons. The tagged driving data served as backup and to re-evaluate specific situations of interest reported during the interviews.

One (hands-on) vehicle needed to be retired early from the expert assessment due to a software fault after being tested by three experts, as it entered a fail-safe mode with no re-activation possible. Additional driving data was generated with the remaining two vehicles in order to compensate for the loss of driving data to be collected in the assessment. Due to this vehicle malfunction, only three out of five experts conducted an interview on both hands-on functions under test.

Interview data as well as event-specific comments on system design were clustered according to the design aspects HMI, function and DMS and interpreted with regard to the potential challenges related to Level 2 hands-free driving:

1. Hands-off = mind-off: Does the function encourage or enable a lesser or a different kind of attribution of attention to the driving task when active?
2. (Prolonged) Transition times: Which influence does the hand posture have on transitions when steering input is once again required or control is transitioned back to the driver?
3. Foreseeable misuse: Which degrees of freedom for hand posture or which DMS criteria seem related to potential misuse?
4. Mode confusion: Which design aspects influence the awareness of the driver of her/his current responsibilities respectively of the current mode of assistance?
5. Safety level: Are there specific aspects of L2H-off systems that seem likely to increase or decrease the safety level in terms of traffic safety?

4.2.2.1 Results from the evaluation of recorded data

Analysis of the tagged video data revealed that attention reminders (AR / hands-on alerts (HOA)) for different L2H-off vehicles at highway cruising speed were issued at about 5 s after detection. This timeout value may extend in traffic jams at low velocities (< 10 km/h) to up to 15 s.

Direct control requests (DCR) for L2H-off vehicles were usually issued at about another 3-5 s after the first AR (see above) throughout the various systems tested. If the driver still does not take over control after additional 2-3 s, systems enact an effective driver lock-out which usually can only be undone after re-starting the entire vehicle. This procedure is also described in the vehicles' manuals.

Hands-on requests (HOR) for L2H-on vehicles varied widely. Even though some systems applied mildly perceptible attention reminders after 5 s similar to the L2H-off features, insisting HORs were issued at the regulatory threshold of around 15 s after hands-off detection. Second stage HORs as well as DCRs followed 5-15 s after the first HOR. In general, some L2H-on systems under test seem to tolerate considerably longer times in which the driver's availability is considered not measurable as compared to the L2H-off systems under test.

The different hands-on detection systems (torque-based vs. capacitive) also showed considerable differences concerning handling and stability of detection. While steering wheels with capacitive sensors tended to sense the drivers' hand position correctly and reliably in almost all cases, torque-based systems showed a higher potential for failure of detection and concept-based handling mistakes. The limited detection range in comparison to capacitive sensors requires the driver to carry his/her hands along the steering wheel at all times. Misjudgments in the force of the grip can more easily lead to unintended deactivation of the system. Furthermore, using torque as a substitute for the knowledge of whether the steering wheel is actually touched or grasped by the drivers' hands was estimated to provide a higher potential for misuse.

Further analysis revealed that repeatedly detected inattention and failure to comply with DMS reminders and requests in some cases led to a change in the timing of DMS prompts and alert characteristics. Although not featured by all systems, some features reduced timeout until the first AR down to about 3 s and applied enforced requirements on monitoring of the driving and surroundings. In case of multiple or repeated neglection of compliance, the functions were deactivated permanently for the duration of the drive. There was no option for re-activation by the driver without stopping the vehicle and setting the driving mode to P or even performing a full restart of the vehicle.

The characteristic AR, HOA, HOR and DCR times were gathered, analyzed, compared to vehicles used in the FOT (Section 4.4) and transferred into Section 5.1 in order to support the design of the driving simulator systems within SP4. Apart from DCRs triggered by the DMS, the systems also presented direct control requests issued by the function for reasons outside of driver behavior (FDCR). Most common reasons for FDCRs were ODD limitations such as leaving a certain road type or mapped area. In these cases, the HMI immediately presented an FDCR without any prior indication and no visible reason within the vehicle's surroundings to the driver. Depending on the system, the driving function remained active until the driver grabbed the steering input device. While the automated lateral control was handed over to the driver in all cases, the ACC stayed active permanently in some systems which lead to remarks on mode confusion on multiple occasions. In one of the analyzed cases, the reason for the FDCR was at least visually comprehensible for the driver as the function needed to stop lateral support due an upcoming end of lane. In this case, the FDCR was given within sufficient time (>10 s) but without any textual explanation.

Further data evaluation from the custom on-boad sensor unit (see Section 4.1) revealed additional insights on accelerations in longitudinal and lateral directions as well as lane positioning and relative distance to other vehicles. Although the test vehicles were capable of performing assisted or automated lane changes, this feature was not considered in the following evaluation.

In most cases, safety-relevant situations leading to longitudinal accelerations worth mentioning were caused by hidden queues at the end of traffic jams as well as vehicle cut-ins. Occasionally, traffic jams required the drivers to increase the level of attention and perform short but strong braking maneuvers manually (about 3-8 m/s² for <1 sec). In general, vehicle cut-in maneuvers lead to similar braking interventions resulting in (theoretical) values for time headway from 0.2 to 1.5 sec. However, due to the testing and challenging character of the expert study, many situations may have been anticipated before becoming safety-relevant by the drivers in this setting. Usually, both kinds of situations seem to have been caused by the forward-facing sensors failing to detect the relevant in-path target within the given time. For further investigations, traffic jam endings were considered relevant situations in the driving simulator studies of SP4 (see Section 5).

In contrast to traffic jams and cut-ins, some test vehicles started to accelerate unintentionally on ramps or interchanges without any preceding vehicles in front. Driving on ramps with no leading vehicle in front lead to inappropriate setspeeds, potentially uncomfortable longitudinal and lateral accelerations and thus, ultimately, to harsh manual braking.

Tight curves and radii (300-500 m) in combination with the current legal speed limit (55-65 mph) lead to lateral accelerations below 3 m/s² and were handeled properly by the tested systems with no need for driver interventions. However, in one occasion a low horizon sun situation blinded the front camera and caused an incorrect steering maneuver of the function leading to lateral deviations from the lane center of up to 1.1 m. The subsequent manual reaction, induced by a corresponding FDCR due to impaired vision, lead to a lateral acceleration of 1.4 m/s² at 65 mph.

Overall, there were little to no occurences that required an immediate intervention by the driver in terms of lateral control within the given ODD (highways and interstates). In contrast to this, traffic jams and cut-ins remain relevant reasons for drivers to take direct longitudinal control of the vehicle, but present challenges well-known from ACC development.

4.2.2.2 Results from the evaluation of interview data

Subjective data was analyzed according to pre-selected design aspects, i.e., HMI, DMS and function, as derived from SP1. Interview data were clustered according to the different aspects and “Do” statements on positive design aspects were derived from the comparative assessment of the functions (see Table 4-1 to Table 4-3). Some of these statements were already realized in all, others were realized only in some of the systems. Furthermore, certain general ideas and recommendations of the experts with regard to the five potential challenges were sometimes not realized in any of the systems, but rather considered helpful or relevant for future system designs in comparison to ’today’s state (‘ideal conception’).

Aspects considered for the HMI design covered the modalities (single or in combination) of warnings and information by the DMS or the function as well as the position and frequency with which the information was conveyed. Furthermore, information on surrounding traffic by depicting the view of the sensors on the current driving situation as well as a clear indication of when the L2 system was available was considered useful.

Table 4-1: Key statements (“*Do’s*”) on HMI design derived from the comparison of different L2 functions. Experts disagree on statements marked with an asterisk (*), albeit the majority (i.e., N=3) tends in the formulated direction.

Aspects	Key statements on HMI design
Modality	• General preference for acoustic over haptic and solely visual signals • Consistent differentiation reduces confusion: AR only visual; DCR including acoustic cues*
Position and frequency of information	• Use of defined color coding on the steering rim or steering wheel spokes indicates the current mode immediately • Use defined color codings for ARs and DCRs • Increase frequency of audible and visual signals in case of persistent misbehavior to convey urgency Ideal conception: • Consistent font size and location of warnings: centered, easy to read in the instrument panel • Large, meaningful icons
Other	• Detailed, dynamic graphics in CID that show surrounding traffic in real time Ideal conception: • Display L2 icon only if the L2 functionality is available

Aspects considered for the DMS design related to the time span after which the system reminded the driver to stay involved with the driving task. Additional remarks targeted the criteria and the driver behavior on which the systems based reminders and warnings.

Table 4-2: Key statements (“*Do’s*”) on DMS design derived from comparisons of different L2 functions

Aspects	Key statements on DMS design
Admissible time for incompatible driver behavior	• Consistent warnings of approximately 4-7 s considered a good trade-off between safety and driving comfort • Warnings after approximately 30 s hands-free driving during L2H-on considered too long
DMS Type	• Attention detection by means of both gaze/head detection and a capacitive sensor considered effective for reducing misuse compared to other, solely hands-on based DMS-solutions Ideal conception: • Capacitive sensor to detect where and how many fingers touch the steering wheel (‘take-over readiness’)
Detection of incompatible driver behavior	• Face occlusions should lead to AR • Glances (at buttons or elsewhere) below approximately 20 degrees should lead to AR Ideal conception: • Glances below the CID or steering wheel should lead to AR (applying 20 degree rule) • Detection range of capacitive sensors should be increased so that touching the steering wheel spoke when being requested to drive hands-on does not lead to further HOR

Aspects considered for the functional design mainly targeted the types of transitions encountered with different systems. This included ACC as a fallback level of L2 compared to a fallback to manual driving (L0). In addition, the benefit of an information exchange between sensors not being part of the L2 system was discussed, e.g., for speed adaptations.

Table 4-3: Key statements (“Do’s”) on functional design derived from the comparison of different L2 functions. Experts disagree on statements marked with an asterisk (*), albeit the majority (i.e., N=3) tends in the formulated direction.

Aspects	Key statements on functional design
Number and type of modes	• Only two modes in total (in difference to three modes) reduce mode-confusion: Level 2 and Level 0, no L2-similar, intermediate states (L1)* Ideal conception: • ODD restrictions for L2H-off similar to current L2H-on functions instead of limited availability that does not reflect the function’s capabilities as experienced within the ODD
Transition from L2 to L1/L0	Ideal conception: • System-initiated transition should be noticeable on functional level, e.g., by moderate recuperation; too strong recuperation is considered critical • No adoption of set speed for L1 after system-initiated transitions from L2
Transition from L0/L1 to L2	• Possibility of activating L2 functionality from L0 with only one button should reduce mode confusion
Information transfer	Ideal conception: • Information transfer between systems for a safer and more comfortable ride, e.g., information from navigation system in order to enable the (earlier/predictive) information of the driver on system or ODD limits or rain sensor to indicate system deactivation due to increasingly bad weather without need for immediate termination.

In addition to clustering statements based on questions to different design aspects, key findings on the five potential challenges were derived from interview data. A comparison to the state-of-the-art (SP1) served as a reference on relevant aspects.

Changes in gaze behavior, with less attention attributed to supervision when using L2H-off, were found in prior studies during hands-free driving without gaze-related monitoring systems (Boos et al., 2020, Josten, 2021, Kraft et al., 2018, Noble et al., 2021, Othersen, 2016, Victor et al., 2018). Current L2H-off series-production vehicles implement “attention-based” DMS, focusing on gaze and head position. In line with publications stating that countermeasures for divergent gaze behavior during hands-free driving are necessary and effective (Blanco et al., 2015, Kurpiers et al., 2019, Llaneras et al., 2017, Victor et al., 2018), the experts stated that the L2H-off systems under test provide earlier reminders for supervision than some L2H-on systems. Attention has thus to be attributed frequently to relevant areas of the roadway. Mind-off (i.e., attention off) was discussed by the experts as a problem evident especially under high monotony, but the L2H-on systems under test offer no solution for this challenge by reminding the user to take the hands back on the steering wheel. A reduction in attention was thus not considered a problem specific to hands-free supervision but related to the implemented DMS design. Earlier warnings for gaze-based DMS in combination with system disengagements or lock-outs as consequences of persisting misuse were considered a more effective strategy to target this first challenge.

Regarding the second challenge, i.e., prolonged transition times after hands-free monitoring, literature agrees on a slight increase in transition times in case of system-initiated transitions after hands-free monitoring (Cahour et al., 2021; Garbacik et al., 2021; Gold et al., 2013; Josten, 2021; Othersen, 2016), resulting in similar crash rates (Victor et al., 2018) and timepoints of driver steering (Pipkorn et al., 2021) with an adapted DMS. Experts stated that to assess the relevance of slightly prolonged transition times, the system‘s ODD would need to be considered. Especially at lower speeds, driver interventions during the field test often required braking instead of steering maneuvers. Studies conducted with naïve users in SP4 provide more insights into whether hands-on wheel precedes longitudinal driver interventions. In addition to the ODD, the DMS has to be factored into considerations on the relevance of hand posture on transition times. Whether drivers can use their hands for other activities that might prolong the transition to the steering wheel seems relevant as well as the analysis of hand postures when monitoring hands-on, as not all hand postures enable a direct intervention (‘take-over readiness’). Furthermore, gaze-based DMS might increase the anticipation of system limits or other types of transition reasons with a supposedly positive effect on transition times. Thus, the relevance of L2H-off on transition times is again closely related to DMS design and the system’s ODD.

Hands-off supervision without an adapted DMS was found to result in a higher likelihood of a secondary task engagement (Victor et al., 2018; Boos et al., 2020; Llaneras et al., 2013; Reagan et al., 2021). After interacting with different L2 systems, the experts considered misuse to be less related to hand posture than to DMS design. As L2H-on systems allowed for longer intervals of undisturbed inattention to the driving task and detected less potentially relevant non-driving related activities in direct comparison to the L2H-off systems under test, foreseeable misuse is not considered a challenge specific to hands-free monitoring in case of an adequate DMS. DMS based on driver attentiveness criteria were considered rather successful in keeping the driver in the loop.

Regarding the awareness of drivers about the current mode, no specific findings were discovered in the literature search that linked mode awareness directly to hand posture. Again, the frequency with which attentiveness-based DMS remind the driver of his/her monitoring responsibilities during L2 operation was seen advantageous. However, especially in the initial phases of use, the better perception of lateral guidance by hands-on supervision might be a factor in getting to know the system and for differentiating between levels of support provided by different assistance modes. Furthermore, hands-on supervision sometimes increased the potential for involuntary oversteering of lateral control. Significant steering input or acceleration was however considered to be a positive feature to deactivate the function. Mode awareness could thus, at least to some extent, be connected to hand posture. More often, however, it was linked to the types of transitions experienced with a system as well as to the availability of feedback on the current level of assistance. Mode awareness was thus deemed important for L2 in general by the experts. Differentiating between different levels of support in lateral guidance was considered more difficult than perceiving the status of longitudinal guidance, being either clearly off (L0) or on (L1/L2). Differences in communicating the system’s status were considered a major influence on interaction success. As stated for the HMI design aspects, the position, size, complexity and modality of feedback on changes in (lateral) guidance is considered essential. Next to information on system status, the number of modes and the specific functional design, e.g., regarding automatic re-engagements after override, was discussed with regard to the potential for mode confusion. Communicating the reason behind a transition to the driver was considered a helpful addition to an HMI displaying the current level of assistance.

According to the subjective assessment of the experts after the field tests, an additional safety benefit might be achieved primarily by adapted DMS, due to establishing a higher involvement with the driving task via the monitoring of the driver’s gaze. Where possible, the misuse of L2 systems has to be prevented. For example, the use of L2 systems in adverse driving conditions, e.g., in heavy rainfall, might challenge the driver’s capabilities for short-termed interventions. Some systems could be used in conditions that likely present handling limits of drivers, which should be avoided by other means than stating conditions of use in the owner’s manual. Furthermore, informing the driver timely about upcoming ODD limits was considered beneficial in comparison to providing mere notice of reaching a system limit.

Another finding from the expert assessment was an increased level of comfort during longer drives when monitoring hands-free albeit a higher frequency of attention reminders. Overall, small details in DMS design were found to make a difference in direct comparison, e.g., the sensitivity or type of hands-off detection technology or the field of view considered as attentive. In addition, the DMS’ tolerance to allow for adjustments to in-vehicle systems such as the entertainment system was considered relevant for acceptance of attention reminders when monitoring the driver’s state.

Overall, the complexity of L2 systems should not be underestimated, especially when various modes are offered. Understanding how exemplary systems work was considered not an easy task, including the process that would enable a user to learn under which conditions the system can be used and how to interact with the system, e.g., finding on/off buttons, understanding the activation logic or DMS criteria.

4.2.3 Summary of the expert assessment

As the first data collection within this project, the field test in the US, conducted with three series-production vehicles equipped with L2H-on and L2H-off functions, served rather exploratory goals. Next to deriving input for data collections in the German field test and the functional designs to be implemented in the simulator studies, the relevance of different design aspects regarding potentially challenging driver behavior in interaction with L2 functions was assessed. To this end, five experts compared the different systems for highway driving scenarios and assessed the relevance of challenges potentially related to hands-free monitoring of L2 functions in combination with adaptations to DMS.

No safety-critical driving situations or interactions during testing were connected to the hands-off supervision of a system. For the tested systems, none of the potential challenges were deemed a factor exclusive to hands-free monitoring. In fact, all challenges but mode confusion were considered to be closely related to the implemented DMS solution. The difference in the implemented DMS, rather than the hand posture during monitoring, was considered to present the major difference between L2H-on and L2H-off systems. Some challenges such as mind-off and misuse seem even likelier for L2 systems with DMS based on hands-off detection, depending on the actual DMS design. Mode confusion, considered here in terms of the awareness of the current mode as communicated by the system or as perceived by functional degradations when changing between assistance modes, was deemed most closely related to transitions and rather unrelated to the driver's tasks during the continued use of L2. In general, transitions, with a focus on lateral assistance levels, as well as the resulting system state need to be clearly communicated to the user by the system in terms of HMI design, but also in terms of perceptible differences in assistance provided by the different modes, especially L2 and L1 (ACC only).

4.2.4 References

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., Fitch, G. M., Russell, S. M., Green, C. A., Cullinane, B., & Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (DOT HS 812 182). https://trid.trb.org/view/1367054

Boos, A., Feldhütter, A, Schwiebacher, J., & Bengler, K. (2020). Mode Errors and Intentional Violations in Visual Monitoring of Level 2. In *23rd IEEE International Conference on Intelligent Transportation Systems,* Virtual Conference.

Cahour, B., Jean-Francois, F., & Arnaud, K. (2021). Driving a Partially Automated Car with the Hands On or Off the Steering Wheel: Users' Subjective Experiences. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Springer eBook Collection: Vol. 221. Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021): Volume III: Sector Based Ergonomics* 1st ed., Vol. 221, (pp.583–592). Springer International Publishing; Imprint Springer. doi:10.1007/978-3-030-74608-7_71

Garbacik, N., Mastory, C., Nguyen, H., Yadav, S., Llaneras, R., & McCall, R. (2021). Lateral Controllability for Automated Driving (SAE Level 2 and Level 3 Automated Driving Systems). In *SAE Technical Paper Series, SAE Technical Paper Series.* SAE International400 Commonwealth Drive, Warrendale, PA, United States. doi:10.4271/2021-01-0864

Gold, C., Damböck, D., Lorenz, L., & Bengler, K.(2013). "Take over!" How long does it take to get the driver back into the loop? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. doi:10.1177/1541931213571433

Josten, J. (2021). *Nutzer und System im Nutzungskontext - Beherrschbarkeit von Teilautomation* [Dissertation, RWTH Aachen University]. Retrieved from https://publications.rwth-aachen.de/record/817292

Kraft, A.-K., Naujoks, F., Wörle, J., & Neukum, A. (2018). The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *52*, (pp. 40–50). doi:10.1016/j.trf.2017.11.012

Kurpiers, C., Lechner, D., & Raisch, F. (2019). The influence of a gaze direction based attention request to maintain mode awareness. In *Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles,* Eindhoven.

Llaneras, R. E., Cannon, B. R., & Green, C. A. (2017). Strategies to Assist Drivers in Remaining Attentive While Under Partially Automated Driving: Verification of Human–Machine Interface Concepts. *Transportation Research Record: Journal of the Transportation Research Board*, *2663*(1), (pp. 20–26). doi:10.3141/2663-03

Noble, A. M., Miles, M., Perez, M. A., Guo, F., & Klauer, S. G. (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. *Accident Analysis & Prevention*, *151*, 105959. doi:10.1016/j.aap.2020.105959

Othersen, I. (2016). *Vom Fahrer zum Denker und Teilzeitlenker: Einflussfaktoren und Gestaltungsmerkmale nutzerorientierter Interaktionskonzepte für die Überwachungsaufgabe des Fahrers im teilautomatisierten Modus* [Dissertation, Technische Universität Braunschweig, Braunschweig]. GBV Gemeinsamer Bibliotheksverbund. Retrieved from https://link.springer.com/book/10.1007/978-3-658-15087-7

Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter? *Transportation Research Part F: Traffic Psychology and Behaviour*, *76*, (pp. 14–25). doi:10.1016/j.trf.2020.10.001

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, (pp. 400–411). doi:10.1016/j.trf.2021.09.010

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, *60*(8), (pp. 1095–1116). doi:10.1177/0018720818788164

4.3 US User Survey

Documentation by D. Schwarze (fka GmbH)

The online user survey, targeting L2 drivers in the USA, is the follow-up on the expert assessment (Section 4.2) of currently available L2H-off and L2H-on series-production vehicles in the US. With its unique sample, the survey's aim is to gather insights within the project that are not accessible through any other methodology or sample in the project. It focuses on reported misuse (CQ: foreseeable misuse) in comparison of different types of driver monitoring systems (DMS).

4.3.1 Research questions

The main motivation behind this survey are insights into the question whether the higher degree of freedom when being allowed to take the hands off the steering wheel leads to an increased occurrence of misuse by the drivers. Foreseeable misuse is difficult to evaluate in controlled studies and a difference in misuse depending on different types of DMS based on the FOT data (see Section 4.4) is potentially influenced by the presence of the safety driver in the L2H-off drives. Furthermore, driving time within the studies of this project lasted seldom longer than an hour in total. Thus, the survey aims to assess if the actual use of L2H-off functions (with EOD = eyes-on detection) leads to other or more frequent kinds of misuse compared to the use of L2H-on functions (with HOD = hands-on detection, Figure 4-1).

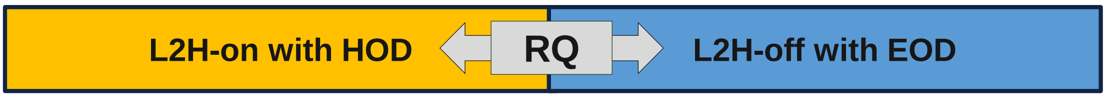


Figure 4-1: Overview of intended function comparisons to answer the research questions (RQ).

Different aspects of foreseeable misuse are addressed based on the types of misuse defined in ISO 21448:2022:

1. *Non-driving related tasks (NDRTs)*: Which NDRTs do drivers engage in during use of L2 functions (indirect misuse)?
2. *Situation of use*: In which situations of use, especially those outside the designated operational design domain (ODD), do drivers activate or rely on the support by L2 functions (direct misuse)?
3. *Driver role*: Are drivers aware of the drivers' responsibilities (e.g., supervision of the function and attention attributed to the road) during L2 function use (indirect misuse)?

Further, the survey addresses whether experienced drivers report changes in their interaction behavior with the L2 function over time. The responses provided shall answer the following questions:

1. Do drivers develop strategies for interacting with L2 based on experience with the function?
2. Do drivers change their usage behavior, especially behavior relating to foreseeable misuse?

Additional aspects of the survey are the participant's experience with their L2 function and the hand position during use to conclude if drivers take their hands off the steering wheel or, in case of the L2H-on functions, keep their hands on the steering wheel.

4.3.2 Method

The survey link was shared from April 6th until August 31st on different platforms (i.e., Facebook, LinkedIn, and forums for automated driving topics). Recruiting strategies focused initially on the Bay Area in California to be able to reach out to participants for the driving simulator study (Section 5.5) and then, additionally, in other states such as, for example, Ohio and NY.

The survey took approximately 20 min to conclude, depending on the participant's response behavior. Beginning April 29th, a voucher of $25 was offered to all participants who completed the survey. The participants who completed the survey had the opportunity to participate in the driving simulator study (Section 5.5) by providing fka SV (Silicon Valley) with their contact information for this purpose.

4.3.2.1 Sample

The target sample consisted of experienced L2H-off and L2H-on users living in the USA. To be eligible to participate, participants had to be at least 18 years of age and needed to indicate to have one of the L2 functions of interest equipped in their vehicle as well as regularly use it. In addition, participants were required to indicate correctly which kind of assistance the L2 function normally provides when activated on highways or interstates. This validation criterion was used to differentiate between L2 and L1 functions, e.g. adaptive cruise control (ACC).

Table 4-1: Overview on the L2 functions participants have experience with.

Functions participants have experience with…	
L2H-off functions	• **SuperCruise (Cadillac): *n* = 27** • Blue Cruise* (Ford): *n* = 18 • Extended Traffic Jam Assist** (ETJA; BMW): *n* = 12
L2H-on functions	• **Autopilot (Tesla): *n* = 28** • Driver+ (Rivian)/ProPilot Assist (Nissan): *n* = 3 each • Lane Tracing Assist (Toyota): *n* = 7 • Pilot Assist (Volvo)/Highway Driving Assist (Kia)/Highway Driving Assist II (Hyundai): *n* = 1 each • Others (e.g., Subaru Crosstrek; Mercedes GLE450): *n* = 11

Note. * Not possible to verify to what extent the function is already available in the vehicles due to dependency of over-the-air updates; ** Can be used as H-off as well as H-on function in different ODDs.

A total of *N* = 353 individuals opened the shared survey link. Of these, *n* = 164 either refused to participate after reading the introduction, did not complete the questionnaire in its entirety or were directed to the end of the questionnaire due to not meeting the criteria described above. Seventy-seven participants did not pass validation control criteria as described in 4.3.2.2

(Experience with L2 functions). The final sample included in the data analysis consisted of *N* = 112 participants, including *n* = 57 L2H-off users and *n* = 55 L2H-on users (see Table 4-1).

4.3.2.2 Survey structure

The survey is divided into different sections (Figure 4-2). The introduction stated the purpose of the survey and provided information about the opportunity to participate in the anchor study (Section 5.5), followed by the general definition of L2 functions to provide a clear distinction between SAE L2 and SAE L1. Afterwards, the primary goals of the survey were addressed whereby types of misuse were differentiated into NDRT, situation of use, and driver role. Changes in the usage behavior over time were targeted subsequently.

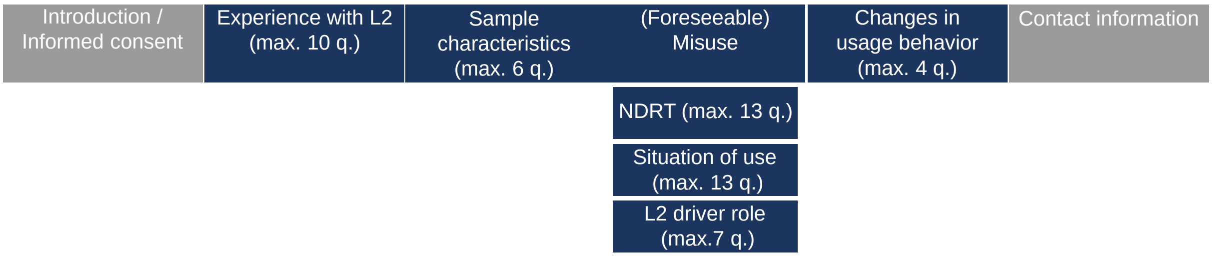


Figure 4-2: Overview of the survey structure. The maximum number of questions per section is listed in parentheses. Depending on the response behavior, less questions are possible.

Where appropriate, scales and foci were aligned with the minimal data set (Section 5) and expert assessment (e.g., sample characterization). The following provides a more detailed explanation of the content of each section. The detailed questions are listed in the Appendix.

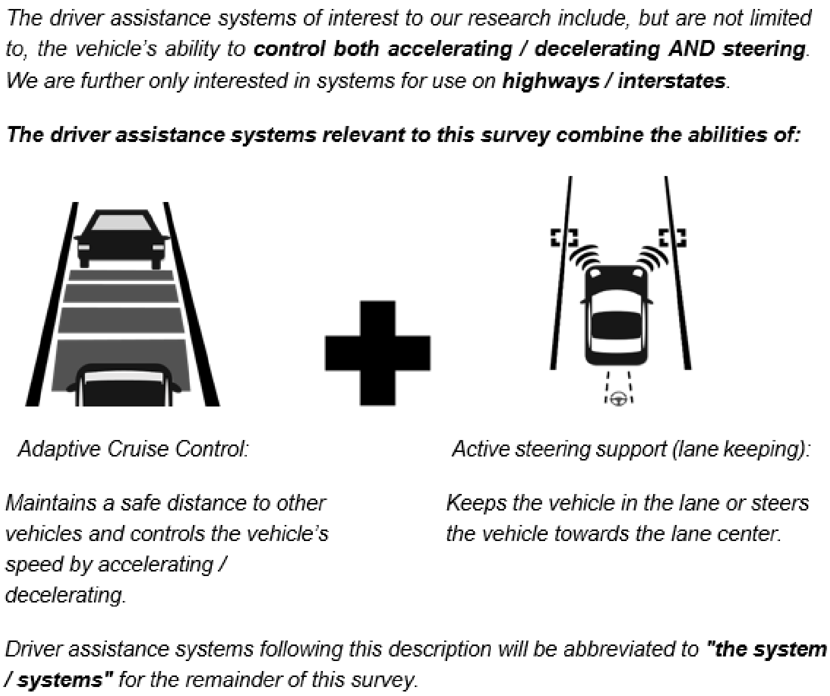


Figure 4-3: Illustration of the definition of L2 functions used in the survey based on Brannon et al. (2020).

Experience with L2 functions. One major challenge of the survey was to ensure that the intended sample was included, i.e., experienced L2 users, in particular experienced L2H-off users. Misunderstandings about the assistance systems that were in the focus of research needed to be avoided. The validation strategy was based on a survey by McDonald, Carney, and McGehee (2018), in which the authors addressed drivers with experience with Advanced Driver Assistance Systems (ADAS). Following this approach, we developed a general definition of L2 functions based on Brannon et al. (2020) to encourage only L2 users to participate (see Figure 4-3).

In order to ensure that (only) L2 function users participated in the survey, filter questions were used. After agreeing to the definition, participants also had to confirm that they have experience with ACC and the combination of ACC and active steering support. For participants who answered "no" to these questions, the survey terminated early after thanking them for their interest and explaining why they cannot partake in the survey. Likewise, the survey terminated early for participants who answered incorrectly to the question of what assistance their L2 function provides on highways / interstates (i.e., maintaining a minimum safe distance to the vehicle in front of you or keeping the vehicle in the lane).

Table 4-2: Functions that have been accepted as L2

L2H-off functions	L2H-on functions
Super Cruise (available in, e.g., Cadillac CT4, CT5, CT6, Escalade & Chevrolet Bolt EUV)	ProPilot Assist (available in Nissan Rogue, Leaf, Altima)
BlueCruise (available in Ford Mustang Mach E & F-150)	I.Q Drive / Travel Assist (available in Volkswagen 2019 models)
Driver+ (available in, e.g., Rivian R1T)	Full Self-Driving / Autopilot (available in Tesla Model 3/S/X/Y)
Extended Traffic Jam Assist / Assisted Driving Plus (part of Assisted Driving Professional; available e.g. in BMW X7, X5 & X3)	Assisted Driving Mode (with or without Traffic Jam Assistant; available e.g. in BMW X3 & X4)
	Dream Drive (available in Lucid Air)
	Pilot Assist (available in, e.g., Volvo S90, XC90 & V90)
	Highway Driving Assist (part of Kia Drive Wise, available in, e.g., Kia Telluride)
	Highway Driving Assist II (part of Smart Sense, available in, e.g., Hyundai Elantra, Sonata, Kona & Tucson)
	Lane Tracing Assist (part of Toyata Safety Sense 2.0, available in, e.g.,Toyota Corolla, Highlander, Prius & Sienna)

To assign participants to one of the two groups, i.e., L2H-on or L2H-off, a pre-selected list of current L2 functions was created, including especially those L2 functions that have been given a specific name by the manufacturer and could thus be differentiated from L1 or other ADAS functions (Table 4-2). This list is based on Consumer Reports' overview of ADAS system names. An additional question at the end of the survey clarified whether the DMS continuously monitors hand posture during L2 use, i.e., inquired about the function’s reaction to the absence of contact to the steering wheel, to cross-check responses for assumed L2H-off users. L2H-on functions, however, only seldom have unique names. Sometimes the assistance system packages names are unique, but not exclusive for the L2 functionality. For all further makes and models, which also provide L2 functions but are not named specifically, the participants were asked to indicate the vehicle model, the vehicle make and the manufacturing year. With this information, we determined whether L2 is included or not.

L2 experience was defined by several factors, since the survey was intended to capture the usage behavior and interaction of regular L2 users. In general, the definition of L2 experience was adopted from the approach of Llaneras (2006) where different levels of ACC experience were defined based on the frequency of use as well as the miles driven since the vehicle was purchased. Therefore, L2 experience in the survey was quantified by

- the frequency of use in daily life,
- the percentage of actual activation (whenever the function can be activated, i.e., within the functions ODD),
- the function experience over time as well as
- the function experience over miles.

Several respective questions are shown in the following Figure 4-4.

Please estimate your experience with the system in months as precisely as possible (e.g., I have used the system for 6 months).

☐ I gathered experience for less than a month.
☐ I gathered experience for ________months.

How often do you use the system in your daily life since you first used it?

Rarely or never	Less than once per month	Several times per month	Several times per week	(Almost) Daily
o	o	o	o	o

Of the time the system would be available for usage, I choose to activate the system approximately ...

_______________% of the time

[…]

Approximately how many miles did you drive in the last twelve months?

- ☐ 0 miles (no drive)
- ☐ 1 mile – 500 miles
- ☐ 501 miles – 3,000 miles
- ☐ 3,001 miles - 8,000 miles
- ☐ 8,001 miles - 15,000 miles
- ☐ 5,001 miles – 30,000 miles
- ☐ 30,001 miles - 60,000 miles
- ☐ more than 60,000 miles

Approximately how many miles did you drive on highways / interstates in the last twelve months?

__________miles

Figure 4-4: Selection of questions to quantify L2 experience.

Foreseeable misuse. After gathering L2 experience and additional demographic data, the primary goal of the survey, foreseeable misuse, was addressed. Misuse was further differentiated for activities unrelated to the driving task, situations of use not recommended in the owner's manual and the perception of the driver's responsibilities.

First, participants were asked about their change in the amount of attention to NDRTs when the L2 function is active compared to manual driving. A follow-up question, when they stated a difference, asked them to specify on which specific activitie(s) they focus more attention. Subsequently, the frequency of engaging in NDRTs while driving with the activated function on highways / interstates was assessed using a 6-point Likert scale (1 = *never* to 6 = *very frequently*). For this purpose, based on the observations and surveys of Pfleging et al. (2016), Petermann-Stock (2015) and primarily those by Metz et al. (2014), a list of potential NDRTs during L2 function use was established. For NDRT-categories considered especially critical during L2 use, participants were asked to specify the exact type of task they engage in more often. Furthermore, they were asked if these specific activities were ever interrupted by alerts from the DMS. To capture the effect of different DMS systems, participants also specified whether they ever received an alert from the DMS because of engaging on activities unrelated to the driving task and, if yes, what impact this alert had on their engagement in NDRTs (e.g., stopped the engagement or reduced the engagement). Furthermore, they were asked if they could think of and specify any strategies that drivers might apply to engage in these activities to remain undisturbed by function alerts.

To capture misuse of L2 functions apart from distraction by NDRTs, participants were also asked whether they are aware of any conditions or situations that are not recommended by the owner's manual. If so, they were asked if they could elaborate on any conditions not recommended by the owner's manuals and whether they have experience using the function under any of these conditions or situations. If participants were not aware of any such conditions, they were asked whether their function had ever behaved in a way they did not expect and, if so, in which situations or conditions this happened. Moreover, the participants were asked whether they had stopped using the function in any of the mentioned conditions or situations after this event. If they continue to activate the function or have experienced the function under conditions or situations for which it is not recommended, they were asked to indicate on a 5-point Likert scale (1 = *decreases* to 5 = *increases*) in what way their usage behavior changes in these conditions (e.g., attention, readiness to take over, and the degree of contact with the steering wheel).

As an additional aspect, the general interaction with L2 functions in daily life was collected. For this purpose, participants were asked to state on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*) to what extent statements reflect their daily use of the function on highways and interstates (e.g., the need to be prepared to take over the driving task or to keep contact to the steering wheel). The participants' reported interaction with the L2 function was compared to their perception on behavior generally required by L2 users. This information was used to differentiate between knowledge on L2 use and participants' own behavior. To gain some insights into the extent to which users of different L2 functions use these functions hands-free, participants were asked to what percentage they use the function on highways and interstates with both hands, with one hand, with less than one hand but at least one finger or with no contact on the steering wheel. Additionally, participants were asked whether they tend to rely on their function or monitor the function and provide similar attention to the driving task as when driving without assistance systems.

Changes in usage behavior. To address changes due to usage experience over time, participants were asked to provide information on whether a particular behavioral pattern or attitude towards the function decreased, remained the same, or increased (5-point Likert scale; see Figure 4-5).

The next questions should provide information on your change in usage behavior over time since you first experienced the system.

In comparison to the first days / weeks / months of using the system…

	(generally) Decreased	Slightly decreased	Did not change	Slightly increased	(generally) Increased
…the amount of attention to the driving task…					
...the amount of attention I allocate to activities unrelated to the driving task...					
...the contact to the steering wheel...					
...the number of different driving conditions in which I use the system...					
...the level of trust in the system...					
...the number of system alerts I receive until I adapt my behavior as told by the system...					
...the time that expires until I react to a system alert...					

Figure 4-5: Changes over time over different categories, rated on a 5-point Likert scale from 1 = *(generally) Decreased* to 5 = *(generally) Increased.*

4.3.3 Results

For further analysis, L2 functions were separated into L2H-off and L2H-on functions to capture effects of different DMS solutions on potential misuse. However, the separation based on L2 experience is not 100% accurate as, e.g., functions like the Extended Traffic Jam Assistant (ETJA) can be used both hands-on and hands-off. Where a split into user groups was not feasible, e.g. due to few responses, the results of both groups are reported combined.

The distribution of age, gender, miles driven in the last 12 months on highways, and the L2 experience per group is shown in Table 4-3. The two L2 groups were generally similar in terms of demographic data. Participants were predominantly male (82%). L2H-on users were slightly younger than L2H-off users (see Table 4-3). Moreover, both groups reported a rather similar extent of experience with their L2 functions over miles and also stated an activation frequency on highways when the function is available of more than 60% of the available time (M = 68.24%, SD = 26.37%, range: 5-100%). Usage frequency ranges between several times per month and several times per day (Md = 3.00, M = 2.88, SD = 1.10, range: 0 = rarely (or never) to 4 = (almost) daily). The limited and shorter availability of L2H-off functions on the market was reflected by a shorter duration of use for this user group.

Table 4-3: Sample demographics, L2 experience, and miles driven in the last 12 months.

	Sample size	**Gender**	**Age [years]**	**L2 experience [month]**	**Miles driven in the last 12 months on highways [mile]**
Overall	n = 112	18 Women 92 Men 2 Other	M = 41 SD = 13.03	M = 17.5 SD = 17.89 *(excl. n = 16 less than a month)*	M = 11,462 SD = 10,624.2 range: 3-60,000
L2H-on function	n = 55	9 Women 45 Men 1 Other	M = 39 SD = 12.23	M = 28 SD = 18.79 *(excl. n = 6 less than a month)*	M = 11,986 SD = 11,998.2 range: 3-60,000
L2H-off function	n = 57	9 Women 47 Men 1 Other	M = 44 SD = 13.41	M = 12 SD = 16.85 *(excl. n = 10 less than a month)*	M = 10,956 SD = 9,186.22 range: 3-60,000

4.3.3.1 Foreseeable misuse

Non-driving related tasks

On average, NDRTs were not rated as receiving significantly more attention during the use of L2 functions compared to when driving unassisted. Activities explicitly stated to receive more attention during either manual or assisted driving could be clustered post-hoc according to the categories by Metz et al. (2014) with one exception, i.e., sightseeing (n = 8 out of n = 49), which receives more attention when driving with L2 functions only.

Overall, there was no obvious difference regarding reported engagement in NDRTs between the functions (Figure 4-6). Both L2H-on and L2H-off users stated rather frequent involvement in activities such as talking on the mobile phone when it is fixated, i.e., can be used hands-free, as well as in vehicle related inputs (e.g., operating the integrated navigation system, adjusting settings in the infotainment system), eating or drinking, and interacting with passengers.

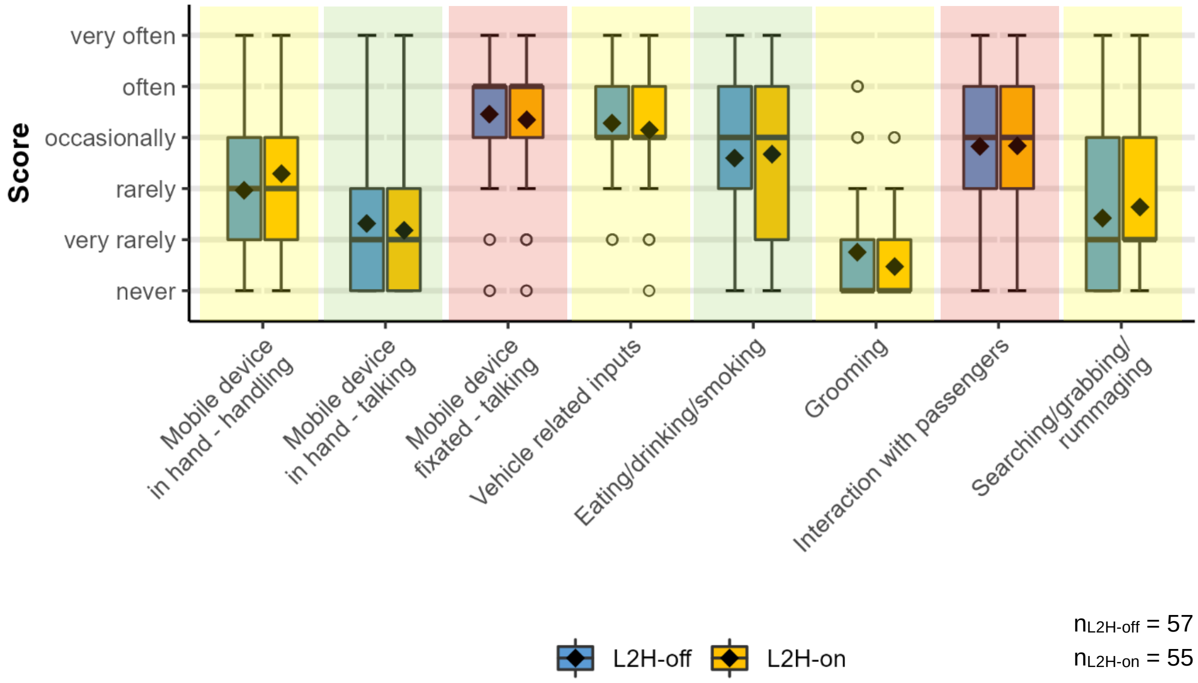


Figure 4-6: Reported frequency to which L2 users engage in certain activities on highways when the function is active, rated on a 6-point Likert scale from 1 = *never* to 6 = *very often*. The activities were based on Metz et al. (2014). The yellow shading represents predominantly visual-motoric activities, the green shading predominantly motoric and the red shading predominantly auditory activities.

For activities concerning the use of mobile devices, those participants who indicated to engage in these NDRTs at least occasionally (i.e., mobile device in hand – handling, $n_{L2H\text{-}off}$ = 40%, $n_{L2H\text{-}on}$ = 49%; mobile device in hand – talking, $n_{L2H\text{-}off}$ = 25%, $n_{L2H\text{-}on}$ = 18%; mobile device fixed – talking, $n_{L2H\text{-}off}$ = 86%, $n_{L2H\text{-}on}$ = 78%) were asked to specify the exact type of task they engage in. For more details, the follow-up questions can be found in the appendix. However, the sub-samples are rather small, rendering a meaningful interpretation of differences difficult. Furthermore, differences were not distinct. For example, slightly more L2H-on users indicated that they text and browse when the L2 function is activated, while slightly more L2H-off users indicated that they read and view video clips during L2 use.

Of the *n* = 84 participants who stated to receive alerts issued by the DMS when engaging in NDRTs, *n* = 23 L2H-off users reported that they stopped or reduced their engagement in predominantly visually distracting tasks such as texting, reading, or video calls. The engagement in predominantly visual and motoric distracting tasks such as browsing, rummaging, or watching videos was stopped or reduced by L2H-on users when receiving an alert (n = 39). In fact, the alert issued by the DMS also affected the hand position of L2H-on users. Additionally, three L2H-on users stated that they no longer took their hands off the steering wheel when engaging in NDRTs.

When asked if the participants could think of any strategies to remain undisturbed by the DMS, n = 18 out of n = 39 participants said they could imagine placing weights or objects on the steering wheel. Seven would keep at least one finger, hand, or knee on the steering wheel to avoid receiving hands-off wheel alerts. To avoid DMS alerts when using hands-free functions, the placement of the smartphone or other objects behind the camera was mentioned (n = 3) to simulate looking into areas accepted by the DMS, or the development of a gaze strategy to shift attention between the road and the NDRTs (n = 5).

Situation of use

Almost 80% of the participants either did not read the owner's manual or were not aware of any conditions not recommended in it (n = 90 out of N = 112). More than half of these participants (n = 54 out of n = 90) reported that they experienced situations or conditions the function did not handle as well as expected. Regardless of the L2 function used, participants reported both longitudinal and lateral situations, such as adaption to stationary or slow-moving vehicles and objects (n = 12) or a steady adjustment of active steering support to maintain the center of the lane (called "pinball" behavior by participants in the survey where they would have expected different function performance; n = 8). However, some participants also reported conditions with limited sensor or camera detection due to inappropriate road conditions (e.g., curvy roads; n = 10), bad weather conditions (n = 5) or not clearly visible lane markings (n = 7). One L2H-off user also reported the experience that DMS detection was unexpectedly limited due to being too short in body height, resulting in a poor detection of the viewing angle.

Participants who did not read the manual, but experienced unexpected behavior were asked whether they use the function under conditions or in situations where it does not behave as expected. Of the n = 54 participants, n = 40 stated that they have stopped using the L2 function in these conditions and situations, while a minority of n = 14 still considered it acceptable to use the function under conditions or situations in which they experienced imperfect or unexpected function behavior. However, asking these n = 14 participants to state their usage behavior changes in these conditions or situations on a 5-point Likert scale (see Table 4-4), the reported results showed a slight increase in the participants' effort to be consciously aware of the driving task, the supervision of the function performance, and the readiness to take over the driving task at all times.

Only a minority of L2H-off (n = 11 out of N = 112) and L2H-on (n = 11 out of N = 112) users stated to be aware of function limitations addressed by the owner's manual. Asked if they use the L2 function in at least one condition not recommended, 50% of the L2 users indicated that they use the function in bad weather conditions or construction zones (n = 11 out of n = 22). When comparing the changes in behavior of participants who have read the owner's manual with those who did not, a similar trend can be seen in their response. Both reported an increase in their readiness to take over the driving task, as well as an increased supervision of the function performance and higher effort to be consciously aware of the driving task. Furthermore, ratings of the L2H-off group ($n_{manual\ known}$ = 5 out of $n_{manual\ known}$ = 11 and $n_{manual\ unknown}$ = 10 out of $n_{manual\ unknown}$ = 14) showed a tendency to increase the degree of contact with the steering wheel outside the function's recommended ODD. Participants who were aware of the function limitations addressed in the owner's manual showed a greater decrease in their awareness of NDRTs in such usage situations compared to participants who indicated that they were unaware of the limitations.

Table 4-4: Changes in usage behavior of participants who use the L2 function under conditions not recommended in the owner's manual (top) or in which the L2 function did not behave as expected (bottom) for the respective conditions.

Changes in usage behavior in the subsample that is aware of limitations in the owner's manual	**Overall (n = 11)**		**L2H-off group (n = 5)**		**L2H-on group (n = 6)**	
	Mean	SD	Mean	SD	Mean	SD
…the amount of attention I allocate to activities unrelated to the driving task.	2.00	1.55	2.00	1.73	2.00	1.55
…the readiness to take over the driving task at all times.	3.09	0.54	3.00	0.71	3.17	0.54
…my effort to be consciously aware of the driving task.	3.45	0.93	3.00	0.71	3.83	0.93
…my supervision of the system performing the driving task according to the current condition.	3.55	1.04	3.00	0.71	4.00	1.04
…the degree of contact I keep with the steering wheel.	3.36	1.03	4.00	1.00	2.83	1.03
Changes in usage behavior in the subsample that continues to use the function in non-ideal conditions	**Overall (n = 14)**		**L2H-off group (n = 10)**		**L2H-on group (n = 4)**	
	Mean	SD	Mean	SD	Mean	SD
…the amount of attention I allocate to activities unrelated to the driving task.	2.93	1.33	2.70	1.16	3.50	1.73
…the readiness to take over the driving task at all times.	3.57	1.16	3.60	1.17	3.50	1.29
…my effort to be consciously aware of the driving task.	3.36	1.15	3.30	0.16	3.50	1.29
…my supervision of the system performing the driving task according to the current condition.	3.50	1.22	3.50	0.27	3.50	1.29
…the degree of contact I keep with the steering wheel.	3.43	1.45	3.50	1.51	3.25	1.50

Note: Rated on a 5-point Likert scale from 1 = decreases to 5 = increases.

The results show that both L2 function groups use the functions under conditions or in situations they are aware of in which the functionality does not perform sufficiently (e.g. in rain or in road works). However, the results also show that users of both functions seem to increase their attention to the driving task, as well as the supervision of the function in such situations and conditions, and thus show anticipatory behavior or an awareness that the function might not be able to handle the situation itself. Nevertheless, with only 20% of participants stating that they were aware of the function limitations addressed in the owner's manual, the findings imply that the awareness of function limitations and the ODD should be increased.

Driver role during L2 use

When asked how they interact in daily life with their function on highways, both L2 function groups showed rather rule-compliant behavior. Further, they also showed a good to a very good understanding of the responsibility of a driver in general when the L2 function is activated. Overall, however, it seems that the theoretical understanding of the responsibilities of the driver during L2 use (Figure 4-7) is slightly higher than the own perceived responsibility stated during daily use of the function (Figure 4-8). Both L2 function groups ($M_{L2H\text{-}off}$ = 3.35, $SD_{L2H\text{-}off}$ = 1.32, $M_{L2H\text{-}on}$ = 3.69, $SD_{L2H\text{-}on}$ = 1.18; Figure 4-8) indicated the lowest agreement with the statement that they allocate the same amount of attention to the road as when driving without any driver assistance systems in their daily lives. However, when they were asked what drivers in general are required to do the agreement to this statement was higher ($M_{L2H\text{-}off}$ = 4.13, $SD_{L2H\text{-}off}$ = 1.00, $M_{L2H\text{-}on}$ = 4.29, $SD_{L2H\text{-}on}$ = 0.98; Figure 4-7).

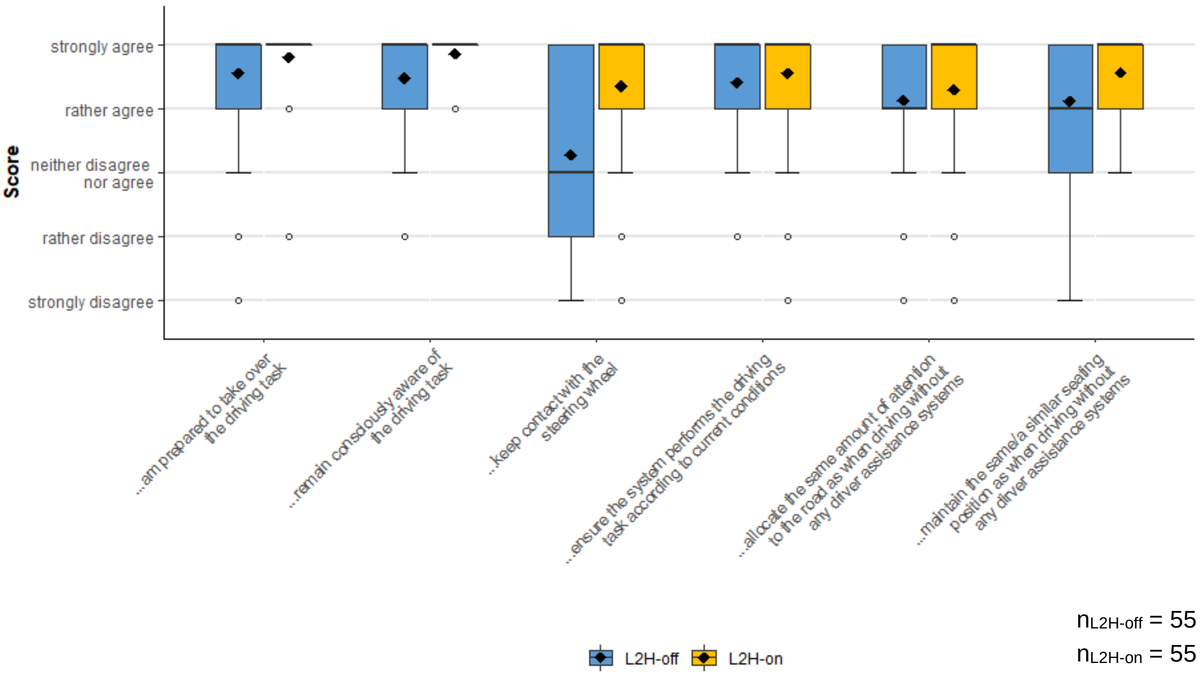


Figure 4-7: Perceived requirements for drivers when using the L2 function on highways. Two L2H-off users terminated the survey before the end and are not included here. Rated on a 5-point Likert scale from 1 = *strongly disagree* to 5 = *strongly agree*.

With regard to their own perceived driver role, L2H-off users generally showed slightly less agreement with all asked statements than L2H-on users (see Figure 4-8). However, descriptively the answers are in the agreeing ranges (means between 3 = neither agree nor disagree and 5 = strongly agree) and, therefore, do not indicate more misuse or disuse for L2H-off functions. A slightly lower understanding for the L2H-off group is descriptively shown by the question about the general responsibilities of the driver when the L2 function is active.

However, the descriptive difference between both groups is small and the responses remain within a high agreement, excluding the item about the hand position (see Figure 4-8).

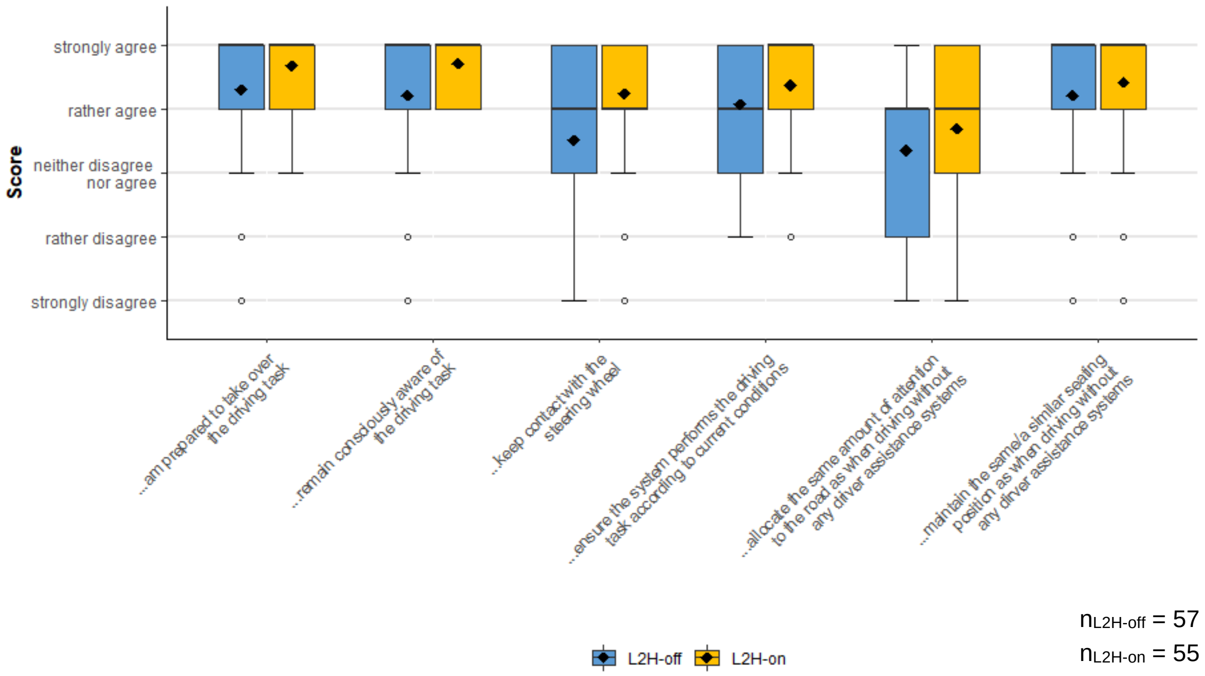


Figure 4-8: General interaction with the L2 function in the participant's daily life. Rated on a 5-point Likert scale from 1 = *strongly disagree* to 5 = *strongly agree*.

As expected, a significantly lower agreement of L2H-off users was found for the statement that contact with the steering wheel is required while driving on highways (U = 1045, p = .001, r = 0.3; $M_{L2H\text{-}off}$ = 3.51, $SD_{L2H\text{-}off}$ = 1.27; $M_{L2H\text{-}on}$ = 4.24, $SD_{L2H\text{-}on}$ = 0.90). The data show a tendency towards stronger agreement than disagreement in the L2H-off group. When answering the questions regarding the perceived requirements for drivers the variation is somewhat higher than when rating one's own usage behavior. On the one hand, this result could be explained by the fact that the separation between hands-off and hands-on functions is not 100% accurate because, for instance, the ETJA can be used as a hands-off as well as a hands-on function. On the other hand, the results could be interpreted in terms of L2H-off users being aware of the opportunity to take their hands off the steering wheel, but actually do not use it in everyday life as often as they could. The fact that the L2H-off users do not use the opportunity to take their hands off the steering wheel at all times is also shown by the percentage of the time the participants have either both hands, one hand, at least one finger or no contact at all with the steering wheel when L2 is active. The L2H-off group reported keeping at least some amount of contact to the steering wheel around 70% of the time. (Table 4-5). Therefore, a certain awareness of the need to keep the hands on the steering wheel under certain conditions and situations could be assumed.

Comparing the data of the L2H-on group with the data of the L2H-off group, the difference regarding their hand position is less than one might expect due to the possibility to take the hands off the steering wheel when driving with the L2H-off function activated. While the L2H-off group seems to make rather less use of their opportunity to not keep contact with the steering wheel than expected, the L2H-on group seems to make more use of their opportunity to take their hands off the steering wheel than expected. Additionally, *n* = 14 out of *n* = 55 L2H-on users answered "no" to the question if the function insist them to keep contact to the steering wheel while it is active. This outcome is contractionary to earlier responses about the understanding of the driver's role, where a high level of agreement was given that drivers both theoretically (Figure 4-7) and in daily use (Figure 4-8) maintain contact with the steering wheel when the function is active. However, the figure also shows that a few L2H-on users (*n* = 3 in Figure 4-8; *n* = 4 in Figure 4-7) did not agree with this statement. A possible lack of awareness of the fact that the function requires contact with the steering wheel is, however, only apparent for two subjects who indicated no agreement in both questions. Six out of *n* = 16 L2H-on users seemed to be aware of the responsibility of the driver to keep contact to the steering wheel but showed lower agreement with regard to their own perceived driver role. These outcomes may indicate indirect misuse. The nature of the question and its interpretation, however, could also have led to misunderstandings. Overall, the results indicate that at least some participants adapt their hand posture to the current situation of use and that a DMS based on hands-off wheel detection does not result in continuous contact to the steering wheel.

Table 4-5: Proportion of hand positions with activated function on highways in percent.

Hand position [%]	**Both hands**		**One hand**		**At least one finger**		**No contact**	
	L2 H-off	**L2 H-on**	**L2 H-off**	**L2 H-on**	**L2 H-off**	**L2 H-on**	**L2 H-off**	**L2 H-on**
Mean	24.55	32.45	34.24	49.56	10.44	12.92	30.78	7.98
SD	24.46	30.12	28.61	29.09	16.11	12.92	12.32	17.05
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Max	80.00	100.00	90.00	95.00	50.00	60.00	50.00	100.00

Participants were asked what best described their interaction with the function based on four statements, ranging from paying more attention to the function and the driving task than when driving manually to relying on the function and paying little or no attention to the driving task. No differences between L2 groups were found regarding the level of attention to the driving task when the L2 function is active compared to driving without any driver assistance systems or the reliance on the function. Overall, the minority of the participants reported relying on the L2 function (*n* = 18 L2H-off users and *n* = 15 L2H-on users) and paying less attention to the driving task than when driving without driver assistance systems.

4.3.3.2 Changes in usage behavior

In order to capture behavioral changes due to increased usage experience over time, participants were asked to provide information on whether a particular behavioral pattern decreased, stayed the same or increased.

For both L2 groups, the amount of attention attributed to the driving task was reported to slightly decrease with experience ($M_{\text{L2H-off}} = 2.63$, $SD_{\text{L2H-off}} = 0.9$, $M_{\text{L2H-on}} = 2.65$, $SD_{\text{L2H-on}} = 1.08$), while the amount of attention to NDRTs slightly increased ($M_{\text{L2H-off}} = 3.46$, $SD_{\text{L2H-off}} = 0.79$, $M_{\text{L2H-on}} = 3.41$, $SD_{\text{L2H-on}} = 0.86$). However, there is no difference between L2H-on and L2H-off functions with their respective DMS regarding the redirection of attention (Figure 4-9, see Item *…number of system alerts received* […]). Furthermore, the contact to the steering wheel decreased with experience not only for L2H-off users, but also slightly for L2H-on users ($M_{\text{L2H-off}} = 2.31$, $SD_{\text{L2H-off}} = 1.02$, $M_{\text{L2H-on}} = 2.81$, $SD_{\text{L2H-on}} = 0.7$). Additionally, the frequency of function usage under different conditions increased for both groups, as well as the level of trust in their function. Whereby the L2H-on group shows a slight tendency for greater trust in the function than the L2H-off group ($M_{\text{L2H-off}} = 3.91$, $SD_{\text{L2H-off}} = 1.00$, $M_{\text{L2H-on}} = 4.02$, $SD_{\text{L2H-on}} = 0.9$). Overall, it should however be taken into account that L2H-on users have been able to gain experience with their function for a longer period of time. On average, the number of alerts issued by the DMS does not show any changes. However, there is a tendency that alerts issued by the DMS seem to decrease for the L2H-on group ($M_{\text{L2H-off}} = 2.94$, $SD_{\text{L2H-off}} = 1.02$, $M_{\text{L2H-on}} = 2.79$, $SD_{\text{L2H-on}} = 0.82$).

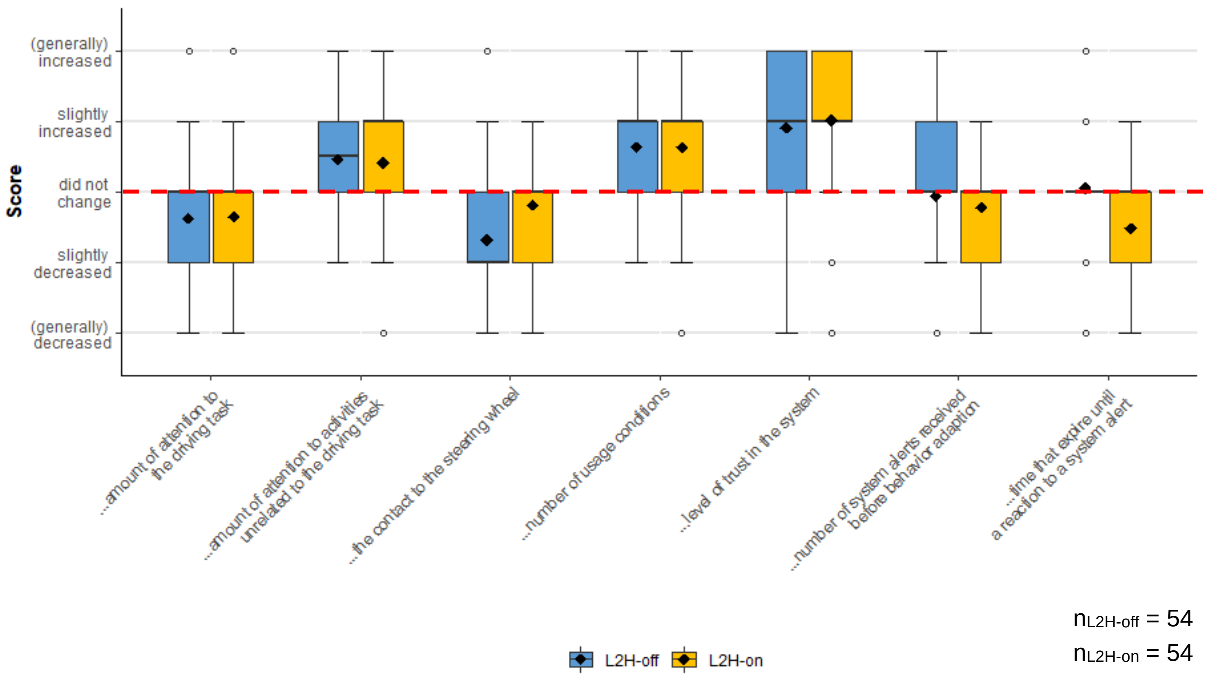


Figure 4-9: Behavioral changes due to increased usage experience over time for L2H-off users and L2H-on users. The red dotted line indicates "no change" over time. Rated on a 5-point Likert scale from 1 = *(generally) decreased* to 5 = *(generally) increased*.

Additionally, four L2H-on users reported that they learned the strengths and weaknesses of the function and become more familiar with it. One L2H-on user stated that, because of increased trust, he gives the function a chance to handle the situation before actually taking over. Two L2H-off users reported that their usage frequency has increased, and that they sometimes choose to drive other routes in order to activate the function more.

As a follow-up question, participants who reported a change in conditions of use ($n = 63$ out of $n = 108$) were asked to indicate how their willingness to use the function had changed and under what conditions (e.g., fog, heavy rain, or situations where road conditions are unclear). In general, both groups showed similar tendencies in the use of the L2 function in the different conditions. However, the use of both functions increased for conditions such as fog, heavy density traffic (> 40 mph), traffic jams (> 40 mph), night time drives, and situations where the participant feels uncomfortable when driving without any driver assistance systems (Figure 4-10). L2H-on users showed a slightly higher descriptive tendency to use their L2 function in fog ($M_{\text{L2H-off}}$ = 3.08, $SD_{\text{L2H-off}}$ = 1.18, $M_{\text{L2H-on}}$ = 3.31, $SD_{\text{L2H-on}}$ = 1.46) or when they feel uncomfortable within a situation ($M_{\text{L2H-off}}$ = 3.19, $SD_{\text{L2H-off}}$ = 1.42, $M_{\text{L2H-on}}$ = 3.23, $SD_{\text{L2H-on}}$ = 0.94). L2H-off users seem to use their L2 function more often than L2H-on users in traffic jams with a speed higher than 40 mph ($M_{\text{L2H-off}}$ = 4.15, $SD_{\text{L2H-off}}$ = 1.12, $M_{\text{L2H-on}}$ = 3.67, $SD_{\text{L2H-on}}$ = 1.41).

One the other side, participants decreased their usage in heavy rain, on slippery or icy roads, or in situations where the road conditions are unclear. For each condition, L2H-on users showed slightly higher decrease than L2H-off users. That tendency could however be explained by higher experience with their L2 function in such situations.

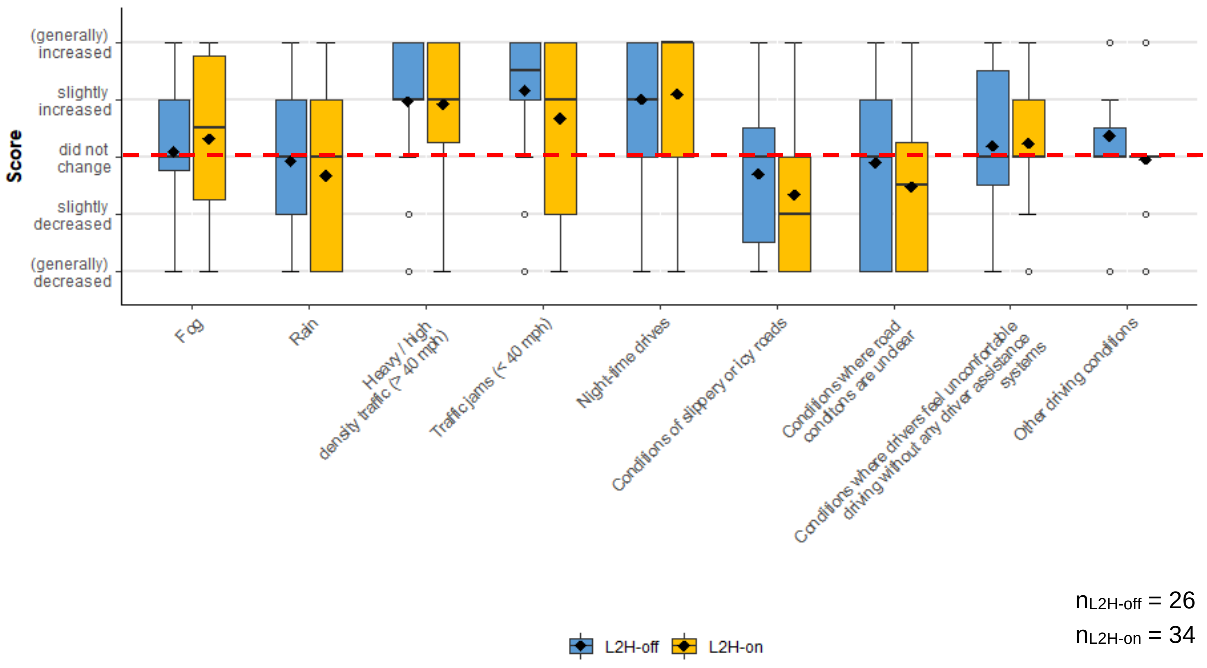


Figure 4-10: Changes in usage behavior for recommended and not recommended conditions in the owner's manual for L2H-off functions and L2H-on functions. The red dotted line indicates "no change" over time. Rated on a 5-point Likert scale from 1 = *(generally) decreased* to 5 = *(generally) increased*.

If participants indicated a change in their attribution of attention, they were asked to explain why their amount of attention to the driving task changed after gathering more experience with their L2 function. Since the responses of the different groups overlapped, they are reported combined and not comparative for each L2 function. The majority of the participants stated that the amount of attention changed because the level of trust in the function increased and they have more confidence in the performance of the function (*n* = 23 out of *n* = 49). Twelve participants reported that their attention changed because they learned when more or less attention is generally needed and that they adapt their attention to the particular situation or condition. However, two participants also stated that they focus more on the environment and less to the driving task after gathering experience with the function. One L2H-on user and one L2H-off user in particular stated that they believe that the function is in some way superior to the human driver and that this is why they changed their usage of the function over time.

Overall, these results show that the participants rely more on L2 functions and show a higher level of trust with more experience. However, these behavioral changes in use need to be contextualized by the participants' statement that they learned when to pay more or less attention depending on the particular situation or condition. Furthermore, increasing experience seems to not lead to more direct misuse by activating L2 functions more frequently in situations or conditions where they are not performing sufficiently.

4.3.4 Conclusion

In contrast to the studies by Llaneras et al. (2013) or Noble et al. (2021), the results of the survey do not show an increased reported engagement in NDRTs during L2 use compared to manual driving. However, this result could also be an effect of the underlying method. In our survey, participants were asked about the difference between manual driving without any driver assistance system and driving with the L2 function. An individual consideration of NDRTs drivers engage in during L0 and L2 driving has, however, not been considered. However, simulator study 1 (Section 5.2.3.1.6; subjective NDRT engagement) also showed no increased involvement to NDRTs comparing L2 and manual driving (L0).

The present study also indicates that there is no difference between L2H-on functions and L2H-off functions in terms of the drivers' involvement in NDRTs. For the categories investigated in this project i.e., mobile device in hand - handling, mobile device in hand - talking, mobile device fixated - talking, and interaction with passengers, other surveys as that by Mueller et al. (2022) did neither show any differences between the L2 functions. The tasks that were indicated as frequent during L2 use for both L2 functions are NDRTs which are not very critical activities in terms of directing attention to the driving task during task engagement (i.e., vehicle-related inputs or talking on the phone). Similar results were found in the objective NDRT engagement of the FOT (Section 4.4.3.3: foreseeable misuse). Alerts issued by the DMS that request to shift the attention to the road or to put the hands back on the steering wheel seem to interrupt secondary task interactions such as texting or watching videos on the phone and are therefore an effective countermeasure. This is also confirmed by Llaneras et al. (2017).

Users in both groups activate the L2 function under conditions in which they are aware of the function's insufficient performance (e.g., in rain or in road works). However, the results show that L2H-off functions do not seem to lead to increased direct misuse compared to L2H-on functions. Moreover, the results indicate that both groups increase their attention to the driving task, as well as the supervision of the function in such situations, and thus show anticipatory behavior or an awareness that the function might not be able to handle the situation itself. Nevertheless, with only 20% of the participants stating that they were aware of the L2 function limitations addressed in the owner's manual, the findings imply that the awareness of L2 function limitations and the ODD should be increased. Overall, the results also show that the function is used in situations outside the ODD, but no differences between the L2 functions could be observed. Furthermore, the L2 users seem to increase their trust with experience, but it does not per se lead to higher reliance on the L2 function or to more direct misuse by activating L2 functions more frequently in situations or under conditions where their performance is not sufficient.

L2H-off users also reported an increased amount of contact to the steering wheel in situations in which the function does not perform sufficiently. In general, L2H-off users do not always make use of the opportunity to take their hands off the steering wheel in all situations, even with experience. In contrast, L2H-on users seem to use the opportunity to remove their hands from the steering wheel temporarily due to DMS design and are sometimes not even aware that they are required to keep their hands on the steering wheel. These finding are consistent with on-road studies (e.g., Banks et al., 2018; Morando et al., 2021) where some people appeared to use hands-on functions as hands-off functions. These results indicate that at least some participants adapt their hand posture to the current situation of use and that a DMS based on hands-off wheel detection does not result in continuous contact to the steering wheel (Section 6.1.3: foreseeable misuse).

This outcome is in contrast to the good understanding of the driver’s responsibility regarding the contact to the steering wheel when the L2 function is active. However, when considering the results of the reported hand position on highways and the possible misconceived permission by the DMS design to temporarily take the hands off the steering wheel during L2H-on use, this could indicate intentional abuse/misuse of the L2H-on function. In general, the participants seem to have a good understanding of their responsibilities. This finding is further supported by the simulator studies and the FOT (see Section 4.4 (FOT) Section 5 (Simulator studies 1-4) and Section 6.1.4 (CQ4: Mode confusion)).

Further results show a tendency to rely more on L2 functions with more experience, which seems to lead to a shift of the participant’s attention away from the driving task. However, this outcome should be contextualized by the participants’ statement that they learned when to pay more or less attention depending on the particular situation or condition with more experience as well. This outcome, considering the other findings (NDRT engagement, situation of use, and increased level of trust), could be an indicator of a strategy to interact with L2 functions. It seems possible that the shift of attention away from the driving task is not continuous, but is used as a result of learning effects in situations or conditions considered appropriate for it. However, confirming this assumption goes beyond the scope of the survey. In general, this survey cannot provide insights into whether or how these changes in behavior translate into the occurrence of incidents when using L2 functions, especially since anticipative, strategic behavior seems common based on the results and might counteract reported NDRT activity. Taking into account the limitations of an online survey, i.e., low control over the care with which users respond or the missing opportunity to address comprehension questions, the survey indicates no increased occurrence of misuse by L2H-off users compared to L2H-on users.

4.3.5 References

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied ergonomics*, *68*, 138-145. doi:10.1016/j.apergo.2017.11.010

Brannon, G., Funkhouser, K., Epstein, A., & Kolodge, K. (2020). *CLEARING THE CONFUSION: Recommended Common Naming for Advanced Driver Assistance Technologies. Retrieved from* https://www.sae.org/binaries/content/assets/cm/content/miscellaneous/adas-nomenclature.pdf

Consumer Reports (2020). *Active Driving Assistance Systems: Test Results and Design Recommendations.* Retrieved from https://data.consumerreports.org/wp-content/uploads/2020/11/consumer-reports-active-driving-assistance-systems-november-16-2020.pdf

International Organization for Standardization. (2022) Road vehicles— Safety of the intended functionality. *(*ISO Standard 21448). Retrieved from https://www.iso.org/standard/77490.html.

Llaneras, R. E. (2006). Exploratory study of early adopters, safety-related driving with advanced technologies*.* Draft final task 2 report*: In-vehicle systems inventory, recruitment methods & approaches, and owner interview results* (No. HS-809 972)

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway. In *Driving Assesment Conference* (Vol. 7, No. 2013). University of Iowa. doi:10.17077/drivingassessment.1472

Llaneras, R. E., Cannon, B. R., & Green, C. A. (2017). Strategies to assist drivers in remaining attentive while under partially automated driving: Verification of human–machine interface concepts. Transportation research record, 2663(1), 20-26. doi:10.3141/2663-03

McDonald, A., Carney, C., & McGehee, D. V. (2018). *Vehicle owners' experiences with and reactions to advanced driver assistance systems* (Technical Report). Washington, D.C.: AAA Foundation for Traffic Safety.

Metz, B., Landau, A., & Just, M. (2014*).* Frequency of secondary tasks in driving–Results from naturalistic driving data*. Safety science, 68,* 195-203. doi:10.1016/j.ssci.2014.04.002

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). Visual attention and steering wheel control: From engagement to disengagement of Tesla Autopilot. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* Vol. 65(1), 1390-1394. Sage CA: Los Angeles, CA: SAGE Publications.

Mueller, A. S., Cicchino, J. B., & Calvanelli Jr, J. V. (2022). Habits, attitudes, and expectations of regular users of partial driving automation systems.

Petermann-Stock, I. (2015). *Automation und Transition im Kraftfahrzeug - Nutzerzentrierte Gestaltung von Übergabe- und Übernahmesituationen innerhalb eines mehrstufigen Automationsansatzes* [Dissertation, TU Braunschweig]. doi:10.24355/dbbs.084-201511041101-0

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th international conference on mobile and ubiquitous multimedia* (pp. 91-99).

4.3.6 Appendix

L2 experience – filter question

Please evaluate the following questions **only** regarding your experience with using the following system:

Super Cruise
(available in, e.g., Cadillac CT4, CT5, CT6, Escalade & Chevrolet Bolt EUV)

Depends on the stated experience

Please do **not** refer to experience you may have gathered with any other system you might be familiar with.

Please indicate which assistance the system normally provides when you use it on highways or interstates.

When the system is activate…

	Yes	No
…it maintains a minimum safe distance to the vehicle in front of you.	O	O
… it controls the vehicle´s speed (according to set speed or speed limits)	O	O
… it keeps the vehicle in the lane.	O	O
… it conducts lane changes (driver- or system-initiated) without the need for the driver to steer.	O	O

(foreseeable) Misuse – NDRT *(depending on stated experience with system)*

In the following, we are interested in your usage of

IQ.Drive / Travel Assist
(available in Volkswagen 2019 models and younger)

Please remember that **responses** to this survey will be **aggregated** and **analyzed anonymously**.

For the following questions, please compare your **behavior** when driving **on highways / interstates <u>without</u>** any **driver assistance system (i.e., manual driving)** to your behavior **when driving <u>with</u> the system**.

In comparison to driving without any driver assistance systems, please indicate how using the system changes the amount of attention you attribute to activities unrelated to the driving task (e.g., talking on the phone, eating).

- ☐ I spend significantly **less** attention to activities unrelated to the driving task.
- ☐ I spend **slightly less** attention to activities unrelated to the driving task.
- ☐ There is **no significant change** in the amount of attention attributed to activities unrelated to the driving task.
- ☐ I spend **slightly more** attention to activities unrelated to the driving task.
- ☐ I spend significantly **more** attention to activities unrelated to the driving task.

Follow up question if question is filled in as indicated above

When driving **without** any driver assistance systems, I **attribute more attention to the following activitie(s)** that are unrelated to the driving task:

- ☐ None
- ☐ I attribute more attention to the following activities:_________________

Follow up question if question is filled in as indicated above

When driving **with** the system, I **attribute more attention to the following activitie(s)** that are unrelated to the driving task:

☐ None
☐ I attribute more attention to the following activities: ___________________

Please answer the following questions regarding your system usage in terms of your everyday driving experience on highways / interstates.

Please indicate the frequency to which you engage in the following activities on highways / interstates when the system is active.

	Never	Very infrequently	Infrequently	Occasionally	Frequently	Very frequently
Mobile phone or similar device (e.g. laptop, mobile navigation system, tablet) **in hand – handling** (e.g. read or type SMS / WhatsApp messages)	o	o	o	●	●	●
Mobile phone or similar device (e.g. laptop, tablet) **in hand – talking** (e.g. talking on the phone with the device in my hand; recording voice message)	o	o	o	●	●	●
Mobile phone or similar device permanently installed or connected to **hands-free system – talking** (e.g. talking on the phone with hands-free; recording voice messages via voice commands)	o	o	o	●	●	●
Vehicle related inputs (e.g. operating the integrated navigation; adjusting settings in the infotainment system, the seat, the air conditioning)	o	o	o	o	o	o
Eating/Drinking/Smoking (e.g. opening a can, eating an apple, lighting a cigarette)	o	o	o	o	o	o
Body related (e.g. styling hair, make up, nail care)	o	o	o	o	o	o
Passenger related (e.g. talking, gesturing, looking at passenger)	o	o	o	o	o	o
Searching/Grabbing/Rummaging (e.g. in a bag)	o	o	o	o	o	o
Other activitie(s) not listed above	o	o	o	o	o	o

Follow up questions if question is filled in as indicated above

Please select **all** specific activities that you **engage in at least occasionally** when the **system is active** while **handling a mobile phone or a similar device holding it in your hand**.

Follow up question: *Mobile phone or similar device (e.g. laptop, mobile navigation system, tablet)* ***in hand - handling*** *(e.g. read or type SMS / WhatsApp messages).*

(The list only shows some examples, please feel free to add more activities you engage in.)

- ☐ Texting (e.g. WhatsApp / SMS / E-mail)
- ☐ Browsing
- ☐ Reading (e.g. messages, eBooks)
- ☐ Viewing (e.g. videos, clips)
- ☐ Playing games (e.g. video games)
- ☐ Taking pictures (e.g. selfies)
- ☐ Other activitie(s) not mentioned in the list:______________________

When engaging in one of the above activities, the system generally requests me ...

- ☐ ...to put my hands on the steering wheel (with or without a request to stay attentive)
- ☐ ...to stay attentive (e.g., to shift my attention to the road).
- ☐ None of the above.

Follow up questions if question is filled in as indicated above

Please select **all** specific activities that you **engage in at least occasionally** when the **system is active** while **talking with / on a mobile phone or a similar device holding it in your hand.**

Follow up question: *Mobile phone or similar device (e.g. laptop, tablet)* ***in hand - talking*** *(e.g. talking on the phone with the device in my hand; recording voice messages).*

(The list only shows some examples, please feel free to add more activities you engage in.)

- ☐ Talking or listening
- ☐ Recording voice messages
- ☐ Phone calls without video
- ☐ Video Calls
- ☐ Other activitie(s) not mentioned in the list: ____________________________

When engaging in one of the above activities, the system generally requests me ...

- ☐ ...to put my hands on the steering wheel (with or without a request to stay attentive)
- ☐ ...to stay attentive (e.g., to shift my attention to the road).
- ☐ None of the above.

Follow up questions if question is filled in as indicated above

Please select **all** specific activities that you **engage at least occasionally** when the **system is active** while **talking with / on a mobile phone or a similar device with your <u>hands free</u>.**

Follow up question: *Mobile phone or similar device permanently installed or connected to **hands-free system - talking** (e.g. talking on the phone with hands-free; recording voice messages via voice commands).*

(The list only shows some examples, please feel free to add more activities you engage in.)

- ☐ Talking or listening
- ☐ Recording voice messages
- ☐ Phone calls without video
- ☐ Video calls
- ☐ Other activitie(s) not mentioned in the list: ______________________________

When engaging in one of the above activities, the system generally requests me ...

- ☐ …to put my hands on the steering wheel (with or without a request to stay attentive)
- ☐ ...to stay attentive (e.g., to shift my attention to the road).
- ☐ None of the above.

Have you ever received an alert from the system because of engaging in activities unrelated to the driving task?
(Please refer **only to alerts** from the **system** that request you to put your **hands back on the steering wheel** or to **shift** your **attention** to the road.)

- ☐ No, I have never received an alert from the system because I engage in activities unrelated to the driving task.
- ☒ Yes, I have received an alert from the system to put my **hands back on the steering wheel** (with or without a request to stay attentive).
- ☒ Yes, I have received an alert from the system to stay **attentive / shift my attention** to the road.

Can you think of any strategies drivers could apply to engage in activities not related to the driving task to remain undisturbed by system alerts?

- ☐ No, I am not aware of any potential strategies
- ☐ Drivers could apply these strategies to engage in the following activities: (Please enter **strategies** and the **tasks** for which they could potentially apply.)_________________________

Follow up question if question is filled in as indicated above / if at least one "Yes" is applied

Please select your responses to alerts issued by the system.

In general, the alerts (e.g., to put hands back on steering wheel or to stay attentive) issued by the system...

- ☐ **..stopped me** from engaging in the following activities unrelated to the driving task: ____________________
- ☐ **...led to a reduction** of my engagement in the following activities unrelated to the driving task: ________________
- ☐ ...**had no effect** on my engagement in the following activities unrelated to the driving task: ________________
- ☐ ...**had no effect** on my engagement in activities unrelated to the driving task in general, but I **engage** in the following activities **in a different way**: ________________
- ☐ **...led to an increase** of my engagement in the following activities unrelated to the driving task:

(forseeable) Misuse – Situations of use

Are you aware of any conditions or situations of use that are not recommended by the owner's manual?

- ☐ Yes
- ☐ No
- ☐ I did not read the owner's manual.

Follow up questions if question is filled in as indicated above

Can you please elaborate on any conditions not recommended by the owner's manual?

____

Do you have experience using the system under conditions that are not recommended by the owner's manual?

- ☐ I either do not use or I deactivate the system in conditions not recommended by the owner's manual
- ☐ I use(d) the system in at least one condition that is not recommended by the owner's manual, but I would not recommend this to other users.
- ☐ I use(d) the system in at least one condition that is not recommended by the owner's manual and see no reason why I or others should not be doing so.

Follow up questions if question is filled in as indicated above

Please state in what way, if any, your usage behavior changes in conditions not recommended compared to when the system is activated in conditions the system is designed to operate / assist as described by the owner's manual.

My behavior changes regarding…

	Decreases	Slightly decreases	Does not change	Slightly increases	Increases
…the amount of attention I allocate in activities unrelated to the driving task	o	o	o	o	o
…the readiness to take over the driving task at all times	o	o	o	o	o
…my efforts to be consciously aware of the driving task	o	o	o	o	o
…my supervision of the system performing the driving task according to the current conditions	o	o	o	o	o
…the degree of contact I keep to the steering wheel.	o	o	o	o	o

Follow up questions if question is filled in as indicated above

Are there situations that the system does not handle as well as you would have expected?

- ☐ Yes
- ☐ No

Follow up questions if question is filled in as indicated above

Can you please elaborate which situations / conditions the system does not handle as well as you have expected?

____

Have you stopped using the system in situations the system does not handle as well as you have expected?

☐ Yes
▣ No

Follow up questions if question is filled in as indicated above

Please state in what way, if any, your usage behavior changes in conditions you know the system does handle well <u>compared</u> to when the system is activated in conditions the system is designed to operate / assist.

My behavior changes regarding…

	Decreases	Slightly decreases	Does not change	Slightly increases	Increases
…the amount of attention I allocate in activities unrelated to the driving task	o	o	o	o	o
…the readiness to take over the driving task at all times	o	o	o	o	o
…my efforts to be consciously aware of the driving task	o	o	o	o	o
...my supervision of the system performing the driving task according to the current conditions	o	o	o	o	o
…the degree of contact I keep to the steering wheel.	o	o	o	o	o

(forseeable) Misuse – L2 driver role

In the next questions, we are interested in your general interaction with the system and the vehicle in your daily life.

The following statements refer to your daily use of the system. Please indicate how much this reflects your daily use of the system on highways / interstates.

While using the system on highways / interstates, I normally...

	Strongly disagree	Rather disagree	Neither disagree nor agree	Rather agree	Strongly agree
…am prepared to take over the driving task.	o	o	o	o	o
…remain consciously aware of the driving task.	o	o	o	o	o
…keep contact with the steering wheel.	o	o	o	o	o
…ensure the system performs the driving task according to current conditions (e.g. speed limits, distance to other road users, road conditions).	o	o	o	o	o
...allocate the same amount of attention to the road as when driving without any driver assistance systems (e.g. road condition, interaction with other users).	o	o	o	o	o
...maintain the same / a similar seating position as when driving without any driver assistance systems.	o	o	o	o	o

At all times when using the system on highways / interstates, users are required to...

	Strongly disagree	Rather disagree	Neither disagree nor agree	Rather agree	Strongly agree
…be prepared to take over the driving task at all times.	o	o	o	o	o
…remain consciously aware of the driving task.	o	o	o	o	o
…keep contact with the steering wheel at all times.	o	o	o	o	o
…ensure the system performs the driving task according to current conditions (e.g. speed limits, distance to other road users, road conditions).	o	o	o	o	o
...allocate the same amount of attention to the road as when driving without any driver assistance systems (e.g. road condition, interaction with other users).	o	o	o	o	o
...maintain the same / a similar seating position as when driving without any driver assistance systems.	o	o	o	o	o

Please indicate to which percentage you use the system on highways / interstates...

(Please note: Your ratings need to result in a total of 100%. **Example**: If you apply each of the four hand positions equally often, select 25 % for each option.)

...with **both hands** on the steering wheel (e.g., both hands placed in the top half / lower half of the steering wheel or one hand in the top half and one in the lower half).	%
...with only **one hand** on the steering wheel (e.g., in the top half / lower half, on the left / right side of the steering wheel).	%
...with **less than one hand** but at **least one finger** in contact to the steering wheel (e.g., 1-4 finger(s) are embracing the top / lower rim of the steering wheel).	%
...with **no contact** to the steering wheel with my hands or fingers.	%

Does the system insist you keep contact to the steering wheel while it is active (e.g. that you more or less constantly keep at least one finger on the steering wheel rim)?

- ☐ Yes
- ☐ No

Which statement would best describe your interaction with the system on highways / interstates?
Please select the most applicable.

- ☐ I permanently monitor the system and attribute **more attention** to the driving task than when driving without any driver assistance systems.
- ☐ I monitor the system and attribute **similar attention** to the driving task than when driving without any driver assistance systems.
- ☐ I rely on the system and attribute **less attention** to the driving task than when driving without any driver assistance systems.
- ☐ I rely on the system and attribute **little to no attention** to the driving task.
- ☐ None of the above.

Changes in usage behavior

The next questions should provide information on your change in usage behavior over time since you first experienced the system.

In comparison to the first days / weeks / months of using the system…

	(generally) Decreased	Slightly decreased	Did not change	Slightly increased	(generally) Increased
…the amount of attention to the driving task…	o	o	o	o	o
…the amount of attention I allocate to activities unrelated to the driving task…	o	o	o	o	o
…the contact to the steering wheel…	o	o	o	o	o
…the number of different driving conditions in which I use the system…	o	o	o	o	o
…the level of trust in the system…	o	o	o	o	o
…the number of system alerts I receive until I adapt my behavior as told by the system…	o	o	o	o	o
…the time that expires until I react to a system alert…	o	o	o	o	o

We would like to understand why you **use the system differently after you gathered more experience** with the system. Please try to explain why your usage of the system changed over time by describing for example an influencing situation/condition:

____

Follow up question if question is filled in as indicated above

We would like to understand the change in your attention to the driving task. Could you please explain why your **amount of attention to the driving task** changed after you gathered more experience with the system?

____

Follow up question if question is filled in as indicated above

Please indicate how your willingness to use the system for the following conditions changed over time since you first experienced using the system.

My willingness to use the system in...

	(generally) Decreased	Slightly decreased	Did not change	Slightly increased	(generally) Increased	Did not experience
…fog	o	o	o	o	o	o
..heavy rain	o	o	o	o	o	o
...heavy / high density traffic (speed above 40 mph)	o	o	o	o	o	o
...traffic jams (speed below 40 mph)	o	o	o	o	o	o
...night-time driving	o	o	o	o	o	o
...conditions of slippery or icy roads (e.g. slush) or snowfall	o	o	o	o	o	o
... conditions where road conditions are unclear (e.g. no lane markings)	o	o	o	o	o	o
...conditions in which I feel uncomfortable driving without any driver assistance systems	o	o	o	o	o	o
...other driving conditions not listed above	o	o	o	o	o	o

4.4 Field Operational Test (FOT)

Documentation by N. Grabbe, D. Albers, B. Biebl, A. Feierle, M. Hübner, T. Hecht, K. Bengler (Lehrstuhl für Ergonomie, TU München)

The goal of the FOT DE is to gain insights into short-term naturalistic driver behavior for vehicles equipped with L2 automated driving systems (ADAS) either as L2H-on or L2H-off. Manual driving was not considered due to the following reasons:

- in the FOT L2H-on is set as the reference for L2H-off,
- the safety effects of L2 in comparison to the manual driver (L0) can be found in the general literature,
- the driving behavior and performance of manual driver vs. driver interacting with L2 is conceptually difficult to compare in merely one short-term study,
  and one additional group of L0 users would unnecessarily increase the resources and the duration of the FOT.

The approach is exploratory and aims to provide evidence regarding the five CQs as well as to derive hypotheses for the subsequent simulator studies planned in this project. Documents such as questionnaires, instructions, and experimenter guides are attached in Appendix B-D (4.4.6.2-4.4.6.4).

4.4.1 Research questions

Four research questions are formulated to create the scope for the study design:

I. …
   a. How are L2H-on and L2H-off (conceptually) implemented in vehicles in Germany?
   b. How do current driver monitoring systems perform in real traffic conditions?

These questions aim to systematically capture the state of the art of L2 systems currently or soon to be available on market. The results are used for the design of the L2 ADAS in the simulator studies (Section 5).

The following three research questions will be examined for the five challenges & questions (CQs) addressed in the project (Section 2.1).

II. Are there differences between first contact L2H-on and first contact L2H-off with regards to the challenges?

This question lays the focus on users naïve to L2 systems: Both groups should have no or only a little knowledge about ADAS. The outcome is expected to help identify training and instruction needs specific to L2h-off. The intervals to be compared are indicated in orange in Figure 4-1.

III. Are there differences when switching from L2H-on familiar to first contact L2H-off with regards to the challenges?

It is expected that L2H-off vehicles will be bought and used predominantly by users who are already familiar with L2H-on (or ADAS at the least). Potential effects on driving behavior when switching from L2H-on to L2H-off are addressed in this research question. The intervals to be compared are indicated in green in Figure 4-1.

IV. Are there differences in first contact L2H-off with and without prior familiarization with L2H-on with regards to the challenges?

In addition to the previous question, differences in the driving behavior of drivers encountering L2H-off for the first time will be examined for the two groups 'users naïve with ADAS' and 'users familiar with L2H-on. The intervals to be compared are indicated in yellow in Figure 4-1.

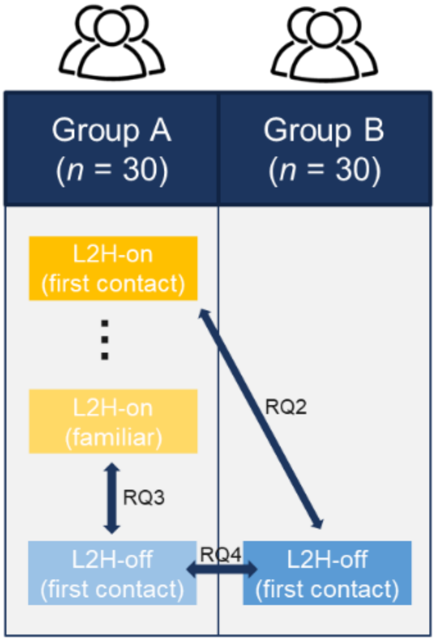


Figure 4-1: Visualization of study design, participant groups, and relevant comparisons between intervals.

4.4.2 Method

4.4.2.1 Experimental Design and Procedure

A between-subject design is planned. The between-subject factor is whether participants are already experienced with L2H-on when experiencing L2H-off or whether they have no prior experience with L2H-on when experiencing L2h-off (RQ2 and 4). A within-subject factor exists within group A through the increasing level of experience with L2H-on functions or the use of L2H-off after prior experience with L2H-on (RQ3).

Group A (*n* = 30) experiences L2H-on before testing L2H-off. In group A, two intervals are used for the analysis of the L2H-on experience; the first dataset covers an interval at the beginning of the experiment where the driver is inexperienced with L2; the second dataset is recorded at the end of the L2H-on rental when the driver is already familiar with L2H-on. Group B (n = 30) only experiences L2H-off.

The four intervals have a duration of at least 45 min. Risser and Brandstätter (1985) state that participants show normal driving behavior after about 15 min driving.

Procedure

The data for the FOT is collected between February 28, 2022 and May 27, 2022. The study procedure for participants of group A and group B is outlined in Figure 4-2.

Group A:

- The first contact consists of informing the participant about the study's contents and conditions. In addition, participant data on demographics and driving experience are collected. This part takes about 1 hour and consists of reading this document and filling out the questionnaire.
- At the second contact, the participant and the experimenter discuss the handover of the L2H-on test vehicle. At this appointment, the date for the vehicle handover will also be set. The part takes about ½ hour and takes place via phone or video call.
- The third contact consists of handing over the L2H-on test vehicle to the participant for approximately five days. This involves a short introduction and instructed familiarization drive for the L2H-on test vehicle. Furthermore, the requirements for the driving task are discussed. This part takes about 1 hour and takes place at the Chair of Ergonomics in Garching.
- Between the third and the last contact, the participant’s task within the scope of the experiment is to perform a certain driving performance/to produce a certain mileage with the use of the L2H-on function and to fill in an online survey twice.
- The fourth contact consists of the return of the L2H-on test vehicle. When the vehicle is returned, the participant takes part in an interview to report impressions during the use of the L2H-on test vehicle. This is followed by an accompanied test drive with a L2H-off test vehicle. The test ends with an online survey and a final interview. This part takes about 3.5 hours and takes place at the Chair of Ergonomics in Garching.

Group B:

- The first contact consists of informing the participant about the study's contents and conditions. In addition, participant data on demographics and driving experience are collected. This part takes about 1 hour and consists of reading this document and filling out the questionnaire.
- At the second contact, the participant books an appointment for the experiment via a calendar tool. This part takes about 10 min.
- At the third and last contact, the participant receives a short introduction and instructed familiarization drive for the L2H-off test vehicle. This is followed by an accompanied test drive with a L2H-off test vehicle. The test ends with an online survey and a final

interview. This part takes about 3.5 hours and takes place at the Chair of Ergonomics in Garching.

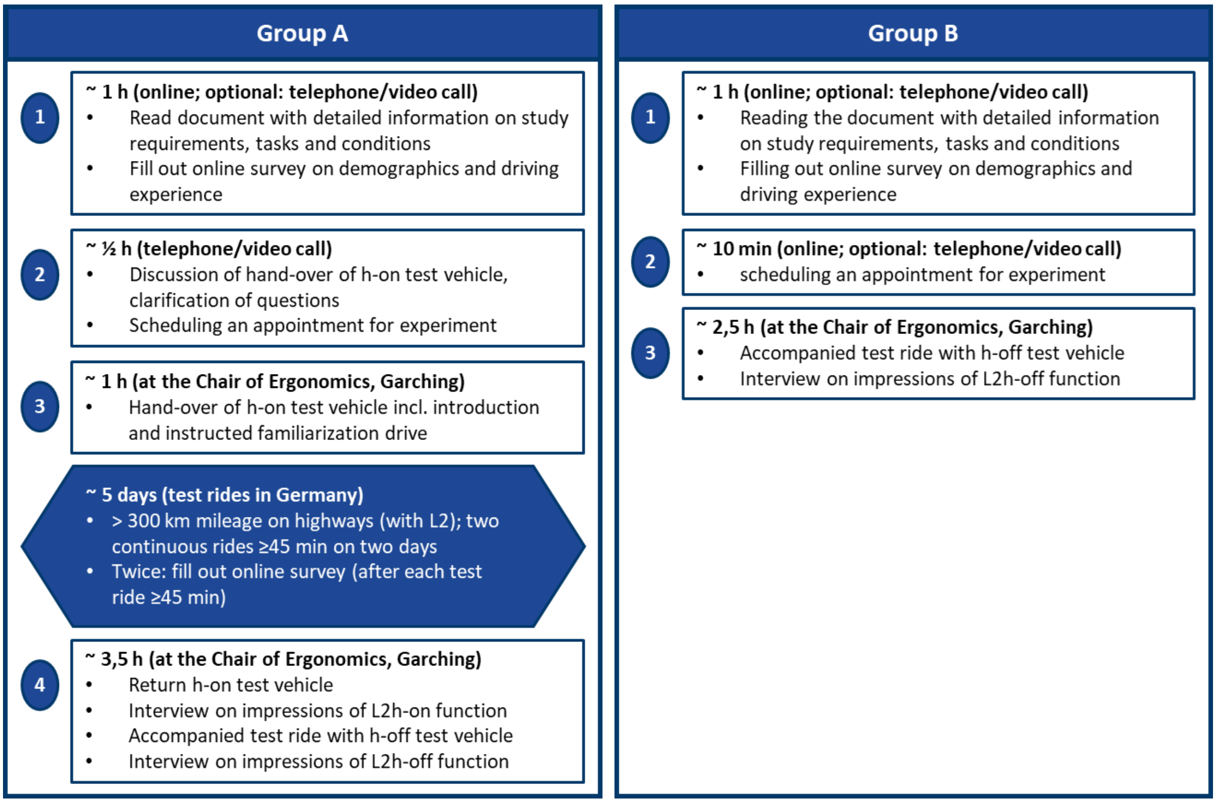


Figure 4-2: Outline of study procedure for participants of group A and group B.

Instructions

H-on test rides:

- Participants are instructed to follow the traffic rules and to avoid safety risks at all times. Furthermore, participants are instructed to activate L2 whenever this is possible.
- Test drives are only allowed in Germany. Other than that, there are no route requirements. Participants are instructed to drive on highways and to activate L2 whenever possible. Participants are allowed to also drive on other road types, these distances are not considered in the driving task or the data analysis.

Group A:

- A L2H-on vehicle is provided for about five days in order to generate the data for the intervals Figure 4-3. These rides are unaccompanied. The participant is instructed to continuously drive ≥ 45 min just at the beginning of his/her L2H-on experience, i.e., starting the interval within the first 50 km driven on a highway. After the ride, the participant is instructed to complete an online survey (~20 min).
The participant conducts more rides on highways while using L2 in order to gain experience with the ADS.
The second dataset has a duration of ≥ 45 min. The participant starts this interval after 200 km – 350 km mileage on highways. After the ride, the participant is requested to fill out another online survey (~20 min). See Figure 4-3 for visualization purposes.

- The test vehicles are equipped with trigger buttons. Participants are requested to press the button and verbally comment on situations that they considered special, e.g., critical situations, surprising system behavior, or evaluations of system behavior in particular situations.
- The participant is instructed to mark the beginning of intervals 1 and 2 by pressing the trigger button three times in a row.

L2H-off test rides:

- Participants are instructed to follow the traffic rules and to avoid safety risks at all times. Furthermore, participants are instructed to activate L2 whenever this is possible.
- The test vehicles are equipped with trigger buttons. Participants are requested to press the button and verbally comment on situations that they considered special, e.g., critical situations, surprising system behavior, or evaluations of system behavior in particular situations.
- The routes are determined by the experimenter.

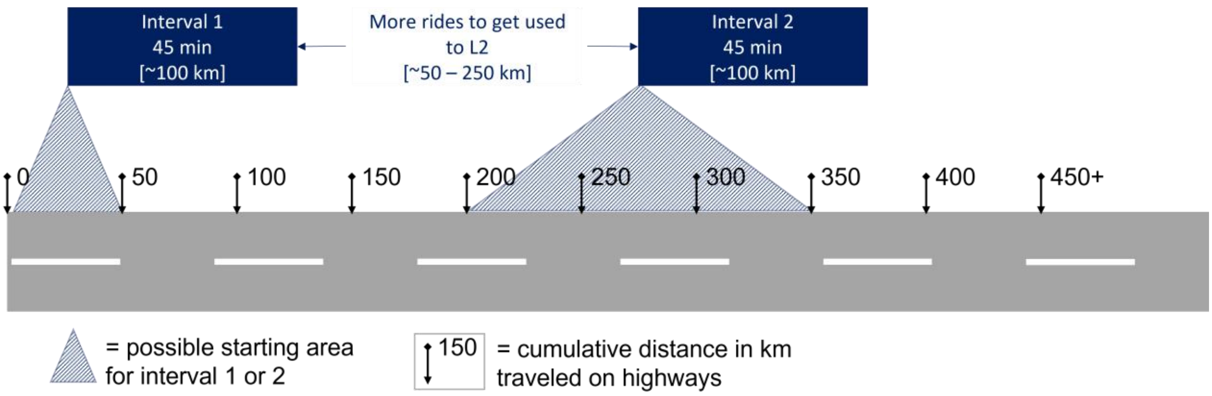


Figure 4-3: Visualization of L2H-on driving task for participants of group A.

4.4.2.2 Test Vehicles and routes

<u>Test vehicles</u>

The six test vehicles comprise four L2H-on test vehicles and two L2H-off test vehicles (see Figure 4-4). The test vehicles are manufactured by four different manufacturers in total. One of the LH-off test vehicles is the same model as one of the L2H-on test vehicles. The other L2H-off test vehicle shares a brand with one L2H-on test vehicle but is another model. Both L2H-off test vehicles are electric vehicles, one L2H-on test vehicle is electric, and one L2H-on test vehicle is hybrid. The other two L2H-on test vehicles have combustion engines. One of the L2H-on test vehicles and one of the L2H-off test vehicles can conduct automatic lane changes (only in L2H-on mode). The settings of the driver assistance systems are predetermined and adjusted to increase the comparability between the test vehicles. The layout setting of the instrument cluster is switched to display the driver assistance system (instead of displaying navigation, infotainment, or other information).

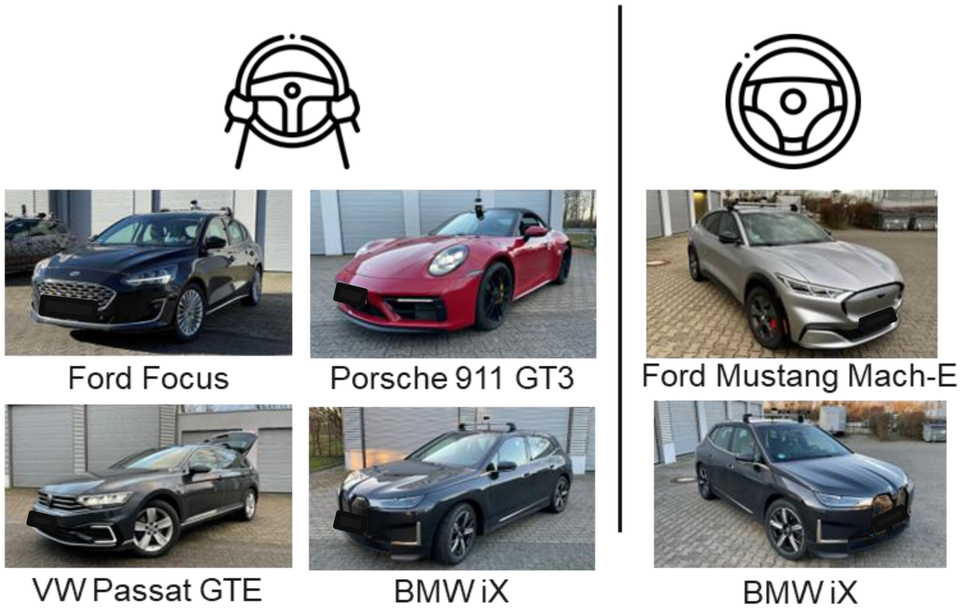


Figure 4-4: Overview of the four L2H-on test vehicles (left) and two L2H-off test vehicles (right) (icon source: Flaticon.com)

Short-term cancellations, weather conditions, vehicle availabilities, technical problems, and participants' parking situations (safety precautions for measurement equipment require private parking places) led to the decision to allocate the vehicles to participants randomly and not based on the privately owned car (or other familiar car models). Maximum organizational flexibility is ensured and the variability among the sample and their prior experiences is increased. For an exploratory approach, different backgrounds of participants may help to assess intuitiveness and to identify effects caused by brand knowledge and prior experiences. In order to reduce learning effects within the experiment, participants of group A are never assigned to a L2H-off test vehicle that was produced by the same manufacturer as the L2H-on test vehicle.

L2 system design: HMI and DMS

For the control of the L2 system and the HMI of the six test vehicles, we refer to the manuals of each test vehicle. In general, for the L2H-on test vehicles one can switch between L0, L1 (ACC), or L2 (ACC+LCA) whereas for the L2H-off vehicles the same general logic applies except for one additional switch within the L2 mode between L2H-on and L2H-off. The four L2H-on test vehicles are only equipped with a HOD as a 3-stage cumulative warning cascade (Stage 1: HOR_1: 15 s, Stage 2: HOR_2: 20 s, Stage3: DDCR: 25 s) detecting the hands position on or off the steering wheel. The HOR_1 is a visual warning, the HOR_2 is a visual warning combined with an acoustic warning, and the DDCR is a visual warning with a continuous acoustic alert and potential haptic alert (brake pulses) in case of no driver reaction. If the driver takes the hands on the steering wheel again after the first and second HOR, then L2H-on is still activated. If the driver reacts to the third warning, then L2H-on is deactivated and the driver has to activate L2H-on again. If the driver does not react at all, then the vehicle comes to a controlled stop or slow down. Instead, one of the L2H-off test vehicles is equipped with a DMS as a 3-stage cumulative warning cascade that is speed dependent (EOR: 3-5 s, HOR: 7-9 s, DDCR: 10-12 s) detecting in addition to the hand position also whether the driver's gaze is directed on-road to the driving scene. The other L2H-off test vehicle has a similar warning cascade but with fixed times: 5s (EOR), 8 s (HOR), and 13 s (DDCR). The EOR is either a visual warning (driver looks at the instrument cluster) or a visual warning with a single acoustic alert (driver does not look at the instrument cluster), the HOR is a visual warning combined with an acoustic warning, and the DDCR is a visual warning with a continuous acoustic alert and potential haptic alert (brake pulses) in case of no driver reaction. If the driver directs the gaze to the road scene again after EOR, then L2H-off is still activated. If the driver directs the gaze to the road scene and also takes the hands on the steering wheel again after HOR, then L2H-off is still activated. If the driver reacts to the third warning, then L2H-on is deactivated and the driver has to activate L2H-off again. If the driver does not react at all, then the vehicle comes to a controlled stop or slows down.

Manuals

L2H-on & L2H-off test vehicles:

The experimenter gives instructions on the settings of the seat, mirrors, and steering wheel, the handling of the engine and gearbox, light, and wipers. Important information in the instrument cluster is explained, e.g., the mileage, the time, fuel, or the charging gauge. An overview of the menu in the central information display is given incl. air condition settings, navigation, and media. Participants are instructed not to change the settings of driver assistance systems or the layout.

Participants receive general information on L2 automated driving functions and the resulting allocation of responsibilities for the driving task. After clarification of questions, participants receive information on the L2 system specifically for the test vehicle. This information includes the manufacturer's name of the L2 function, and the maximum speed (140 km/h for L2H-on test vehicles, 129/130 km/h for h-off test vehicles). The information is available in a short manual that is stored inside the car. Participants may look at it at any time. The manual comprises a photo of the instrument cluster with the activated L2 function. The icons reflecting the status of L2 are highlighted. In cases of different variants (e.g., a different icon if the lead vehicle affects the current speed vs. no lead vehicle), icons for all variants are displayed. The manual reminds participants of their responsibility for the driving task and the limits of L2.

The manual also includes instructions for selected operations when using L2: (1) activation of L2, (2) deactivation of L2, (3) speed adjustment, (4) lane change, and (5) adjustment of minimal distance to lead vehicle. The list includes other remarks, e.g., concerning other ADAS that are controlled via buttons needed for L2 handling or buttons in close-proximity. A photo of the relevant buttons on the steering wheel is added with references in the instructions.

L2H-on test vehicles only:

Participants receive information on the position of the fuel tank / charging cover and safety equipment (warning triangle, warning vest, and first aid kit).

<u>Routes:</u>

L2H-on test drives:

A familiarization ride is conducted on the highway A9 between exit 70 Garching Nord and exit 67 Allershausen and back (~38 km), see Figure 4-5(a). Otherwise, the participants are free to accomplish their test drives on all highways in Germany.

L2H-off test drives:

A familiarization ride is directly followed by the test ride. The ride is conducted on the highway A9 exit 70 Garching Nord to exit 63 Manching and back (~107 km), see Figure 4-5(b). In cases of traffic jams or construction works, the test route the experimenter may adjust the route in order to ensure a familiarization drive of about 15 min and a test drive of about 45 min.

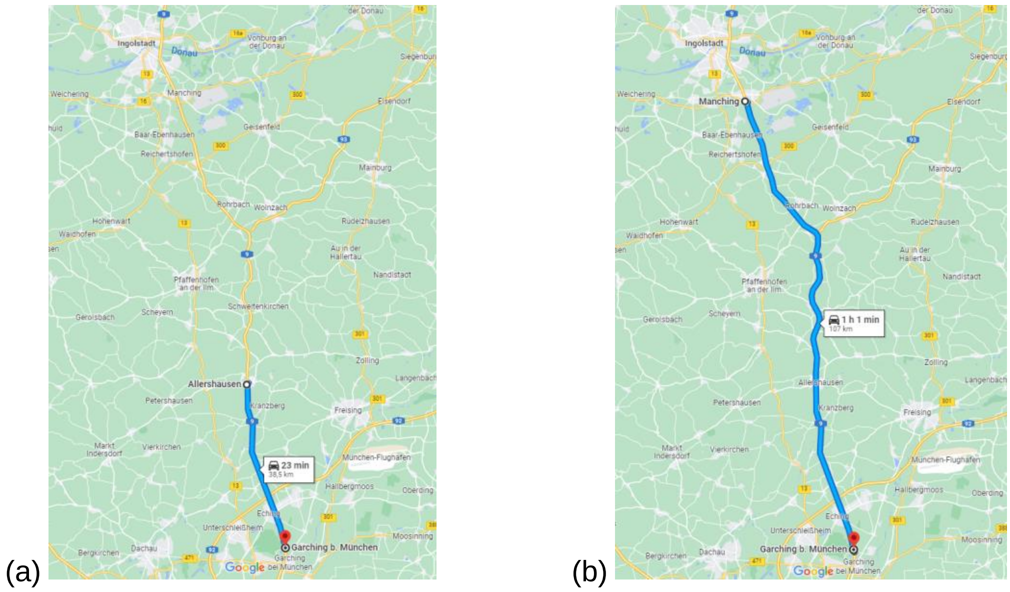


Figure 4-5: Screenshot of standard routes for (a) familiarization rides with L2H-on test vehicles[1] and (b) familiarization and test rides with L2H-off test vehicles[2] (Source: Google Maps, Map Data © 2023 GeoBasis-DE/BKG (©2009),Google)

4.4.2.3 Measurement equipment

In the following, the measurement equipment is briefly described. For more details, we refer to Section 4.1 (chapter about measurement equipment in detail). The test vehicles are equipped with measurement technology for data recording. A total of four cameras are installed here

[1] https://www.google.de/maps/dir/48.2644915,11.646496/48.4280455,11.5879255/48.2666185, 11.6455 157/@48.5318979,11.4511445,10.13z/data=!4m2!4m1!3e0

[2] https://www.google.de/maps/dir/48.2645297,11.6464832/48.7099806,11.4835457/48.2663168, 11.6456199/@48.2640781,11.6526056,14.96z/data=!4m2!4m1!3e0

(see Figure 4-6). Camera 1 is installed frontally to the test person and records the test person's head and face in order to record their attention (gaze directions and movements). The second camera is located diagonally behind the driver at the ceiling near the passenger seat to get a more holistic picture of the participant's usage behavior such as driver interactions with the steering wheel and infotainment system or body postures. Finally, a third and fourth camera is installed near the dashboard and steering wheel to view the displays in the instrument cluster and the steering wheel interaction in more detail.

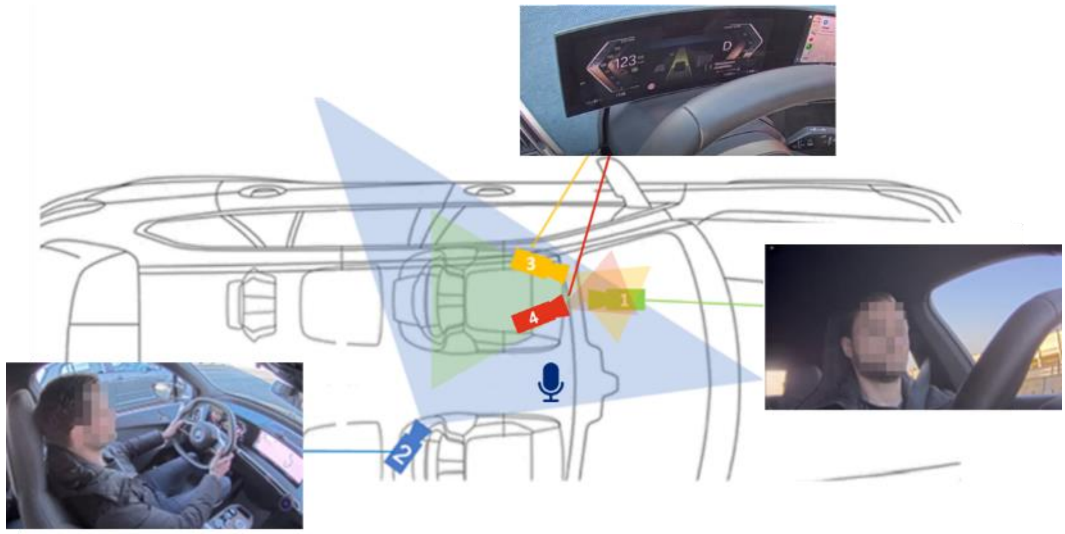


Figure 4-6: Camera positions in test vehicles.

In addition, a 360° LiDAR sensor is installed on the vehicle to provide a comprehensive view of the traffic environment, including road users and infrastructure. Furthermore, a microphone is installed to record the audio signals in the vehicle (verbal reactions of the driver). In addition, the vehicle bus signal of the vehicle is recorded in order to be able to analyze driving data such as speed and acceleration but also system status. In addition, a GPS module is installed to collect the position data of the vehicles. The vehicles are also equipped with a "trigger button" so that participants can mark specific situations and mark the beginning of test drives. The measurement equipment is compactly stowed and fixated in the trunk or the back of the vehicle, respectively (see Figure 4-7).

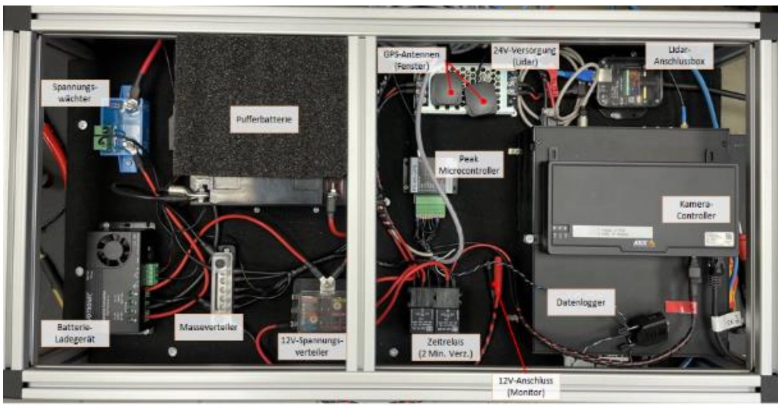


Figure 4-7: Measurement equipment as installed in the vehicle.

Equipment check

L2H-on & L2H-off test vehicles:

The experimenters check the status of the measurement equipment before handing the test vehicle over to the participant. When participants return to the test vehicle, the experimenter checks the status again. The check includes the position of the cameras, the status of the data logger, the lidar sensor, and the trigger button.

L2H-on test vehicles only:

Participants receive written instructions on how to check the status of measurement equipment before starting a test drive. Participants are instructed to inform the study team if there are deviations from the nominal state.

4.4.2.4 Recruitment & Sample Description

In total, the study requires 60 participants with 30 participants in each group. The sample characteristics of both subsets should be comparable in order to minimize the probability of affecting the results.

Recruitment

A balanced age distribution is aimed for as well as a minimum of 30% female participants. No novice drivers are recruited for the experiment. In addition to the driving experience, this is due to safety reasons. The span of five years (holding a driver's license) is chosen concerning the German program "accompanied driving at age of 17" ("Begleitetes Fahren ab 17 Jahre", Fahrerlaubnis-Verordnung, 2010) where the supervising passenger needs to hold a driver's license for at least five years. Potential participants are required to drive regularly (at least once a week) and to have no prior experience with L2 driving. Participants of group A are required to have a private parking place as a safety precaution for the measurement equipment.

Participants are recruited through social media advertising, notices on the campus, and the subject database of the Chair of Ergonomics. Participants fill out an online survey and provide organizational data such as preferred date for the experiment or parking situation as well as data on demographics and driving experience.

Participants of group A receive 220€ if they completed the study and fulfilled the requirements (driving task and online surveys). If they fail to complete the study, participants receive 50€. The compensation includes expenses for fuel or charging the vehicles. Participants of group B receive a compensation of 50€.

Sample

The mean age of participants is 42.97 years for group A (*SD* = 18.11, 21-75) and 36.33 years for group B (*SD* = 13.89, 21-70). The age distribution is visualized in Figure 4-8. In group A 19 participants are male and 11 are female, in group B 21 participants are male and 9 participants are female. All participants of group A are right-handed. In group B 27 participants are right-handed, 2 participants are left-handed, and 1 participant is two-handed. About half of the participants do not need a visual aid when driving a car (group A: *n* = 14, group B: *n* = 15) and the other half of the participants report using their visual aid during the test rides. None of the participants in group A and two participants in group B state to have a color vision deficiency. Only one participant of group B indicates to have a hearing deficiency. They report correcting it during the test rides.

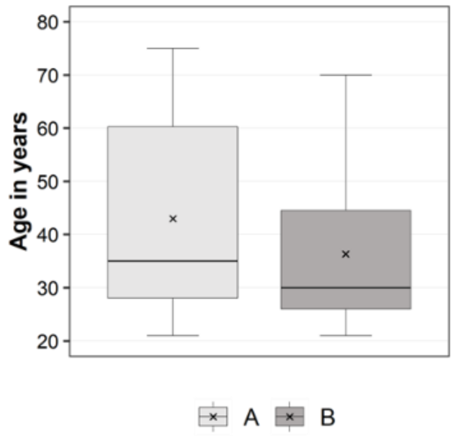


Figure 4-8: Visualization of the age distribution and the mean age in years for groups A (*n* =30) and B (*n* = 30). The sample size indicates the number of participants

Participants hold their driver's license on average since 1996.2 (group A, *SD* = 17.68, 1964-2017) or 2003 (group B, *SD* = 13.82, 1970-2017), respectively. The distribution is visualized in Figure 4-9.

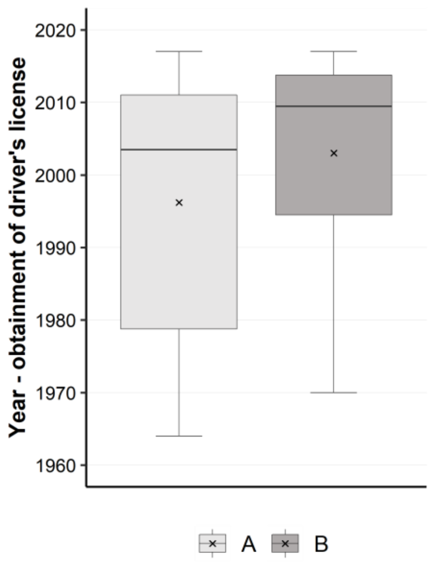


Figure 4-9: Visualization of the distribution and the mean year of obtainment of the driver's license for groups A (*n* = 30) and B (*n* = 30). The sample size indicates the number of participants

All participants report driving a car more often than once a month (*Median* = >1x/week). For highways, two participants in each group indicate to use them less than once a month and most of the participants indicate to drive more than once a month (group A, n = 14) or more than once a week (group B, n = 15). Figure 4-10 visualizes the distribution of participants for the driving frequency in general and on highways.

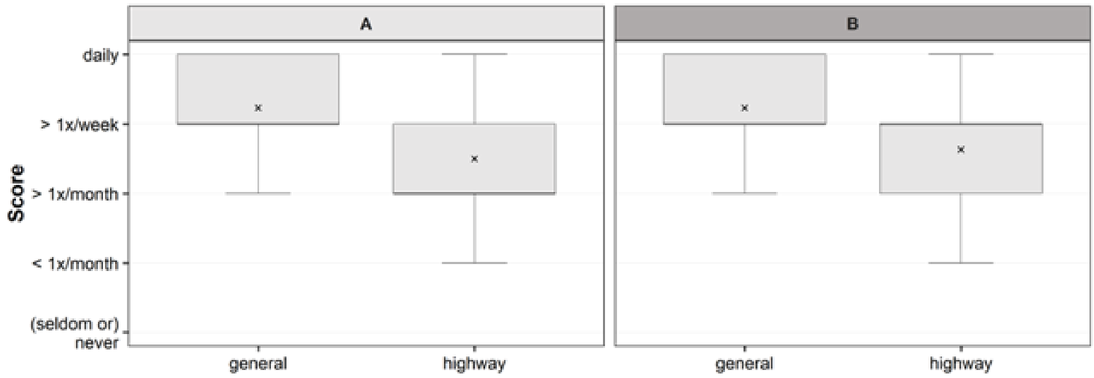


Figure 4-10: Visualization of the response distribution for the driving frequency in general (top) and on highways (bottom) for groups A (n = 30) and B (n = 30). The sample size indicates the number of participants

The reported mileage driven within a year ranges between 1 km and 50,000 km for all types of roads as well as for highways. Most participants indicate to drive between 5,001 km and 20,000 km. Figure 4-11 visualizes the distribution of participants for the mileage per year in general and on highways.

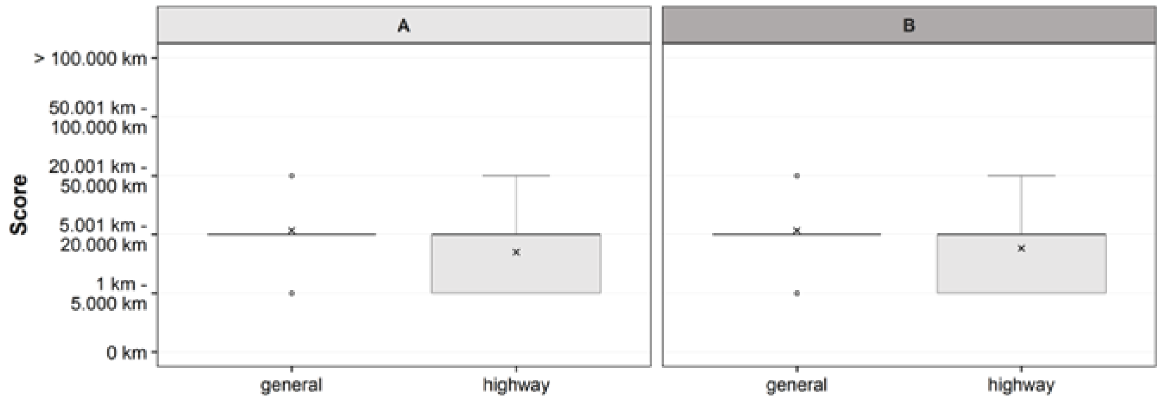


Figure 4-11: Visualization of the response distribution for the mileage within one year in general (top) and on highways (bottom) for groups A (n = 30) and B (n = 30). The sample size indicates the number of participants

Figure 4-12 visualizes the distribution of responses by participants with regard to their familiarity with six different driver assistance systems. The driver assistance system most familiar among the participants is cruise control (CC). All participants report knowing the system and the majority report using it regularly (group A: n = 17, group B: n = 19). The other five assistance systems are less familiar among the participants. The majority of participants indicate for each of these assistance systems knows it but has never used it before. Assisted cruise control (ACC) is used regularly by n = 4 (group A) or n = 4 (group B) participants, respectively.

None of the participants indicates to use the lane-keeping assistant (LKA) or the parking assistant (PA) regularly. The traffic jam assist (TJA) is used regularly only by one participant (group A) and seldomly used by one participant in group A and group B. In group A and group B one participant each indicates to have used L2 seldomly (group B) or even regularly (group A). When discussing this matter with the participants, it is clarified that prior experience is gained years ago only by isolated test drives with rental cars (group A) or trucks while acquiring a driver's license (group B), respectively. Consequently, the participants remain in the sample.

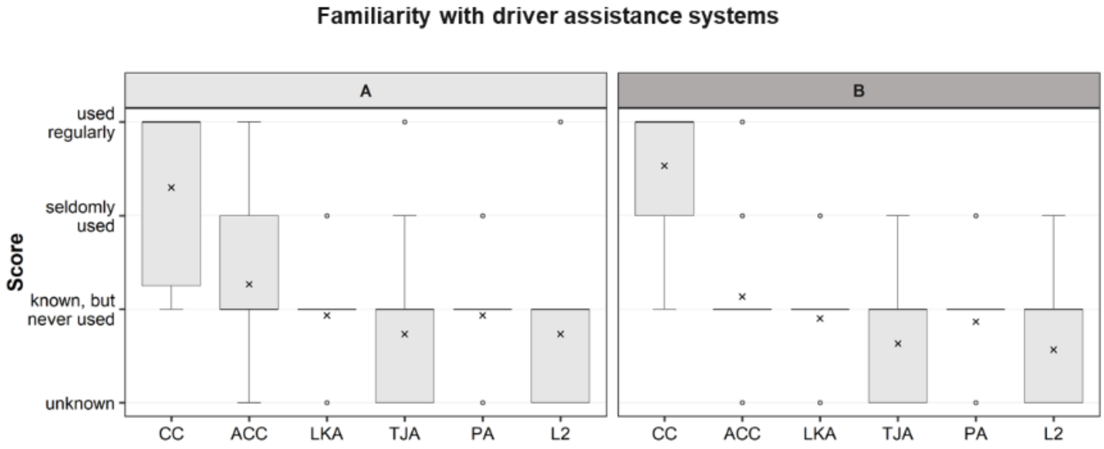


Figure 4-12: Visualization of the response distribution for the familiarity with driver assistance systems for groups A (n = 30) and B (n = 30). The sample size indicates the number of participants. Responses are shown for six driver assistance systems from top to bottom: cruise control (CC), adaptive cruise control (ACC), lane keeping assistant (LKA), traffic jam assist (TJA), parking assistant (PA), and L2.

When participants indicate to use a system seldomly or regularly they are requested to report the manufacturer(s) of the respective familiar system. Most often participants list BMW (among all systems: group A: n = 22, group B: n = 22) followed by VW (among all systems: group A: n = 18, group B: n = 17), Mercedes (among all systems: group A: n = 9, group B: n = 17), Audi (among all systems: group A: n = 10, group B: n = 10), Volvo (among all systems: group A: n = 8, group B: n = 5), Opel (among all systems: group A: n = 7, group B: n = 5), Ford (among all systems: group A: n = 5, group B: n = 7), and Skoda (among all systems: group A: n = 4, group B: n = 7). Other car manufacturers are listed less than 10 times in total.

The driving style questionnaire (French et al., 1993; translation by fka, ika, & LfE) helps to rate the comparability of the subsamples with regard to their driving style. The subscales, their boundaries, and the distributions including the mean values are visualized in Figure 4-13. The subscales have different boundaries of either 3-18 (Speed ↑ = tendency for speeding, Calmness ↑ = calmer, and Focus ↑ = more focus) or 2-12 (Social Resistance ↑ = higher dislike for advice, Planning ↑ = more planning, Deviance ↑ = more deviance), respectively. The subscales for social resistance and planning show slightly different distributions for group A and group B with the difference between the means of Δ = 1.20 (social resistance) and Δ = 1.26 (planning), respectively. All other subscales produce similar distributions as well as means.

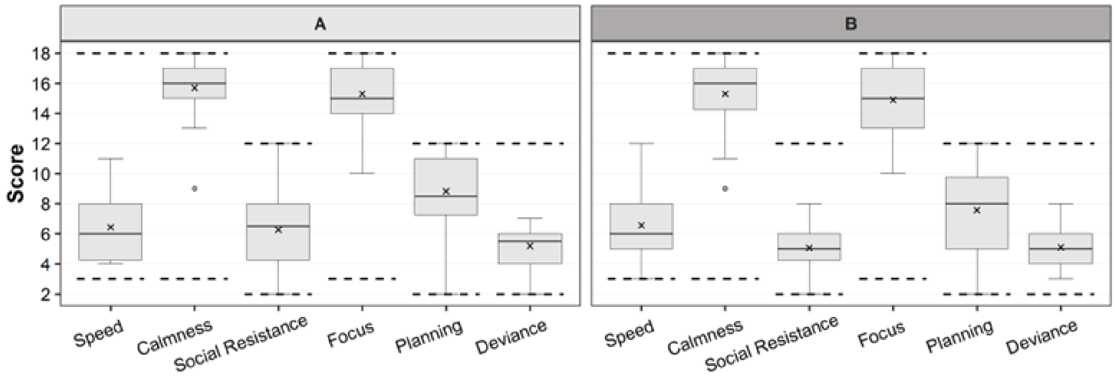


Figure 4-13: Visualization of the distribution and the mean response of the subscales of the Driving Style Questionnaire for groups A (n = 30) and B (n = 30). The sample size indicates the number of participants. The subscales are Speed, Calmness, Social Resistance, Focus, Planning, and Deviance. The boundaries for the subscales are visualized by horizontal lines.

The Affinity for Technology Interaction Scale (Franke et al., 2019) is applied in a general context and not specifically limited to L2 or ADAS. Participants of both groups show a high affinity for technology interaction with means of 4.57 (group A, SD = 1.01) and 4.75 (group B, SD = 0.82). The variance is slightly higher within group A. The distributions are visualized in Figure 4-14.

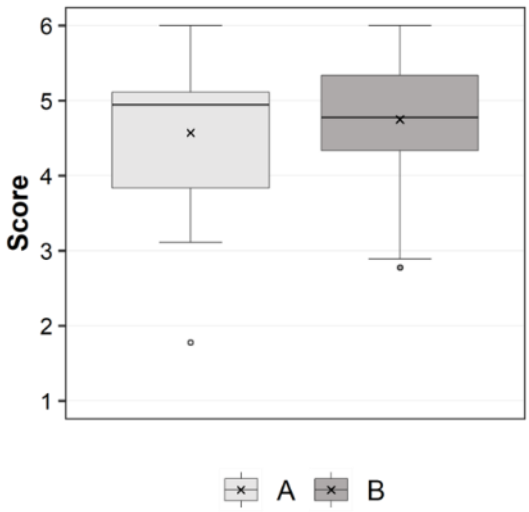


Figure 4-14: Visualization of the distribution and the mean response of the scale Affinity for Technology Interaction (Franke et al., 2019) for groups A (n = 30) and B (n = 30). The higher the score, the higher the affinity. The sample size indicates the number of participants

Driving profile

The median of the minimum speed is similar at 0 km/h (see Figure 4-15). The median of the mean speed is similar between the four groups at 100 km/h but a higher dispersion exists for A H-on drives. The median of the maximum speed is similar at 130 km/h but also shows a high variance for the A H-on (fam) group.

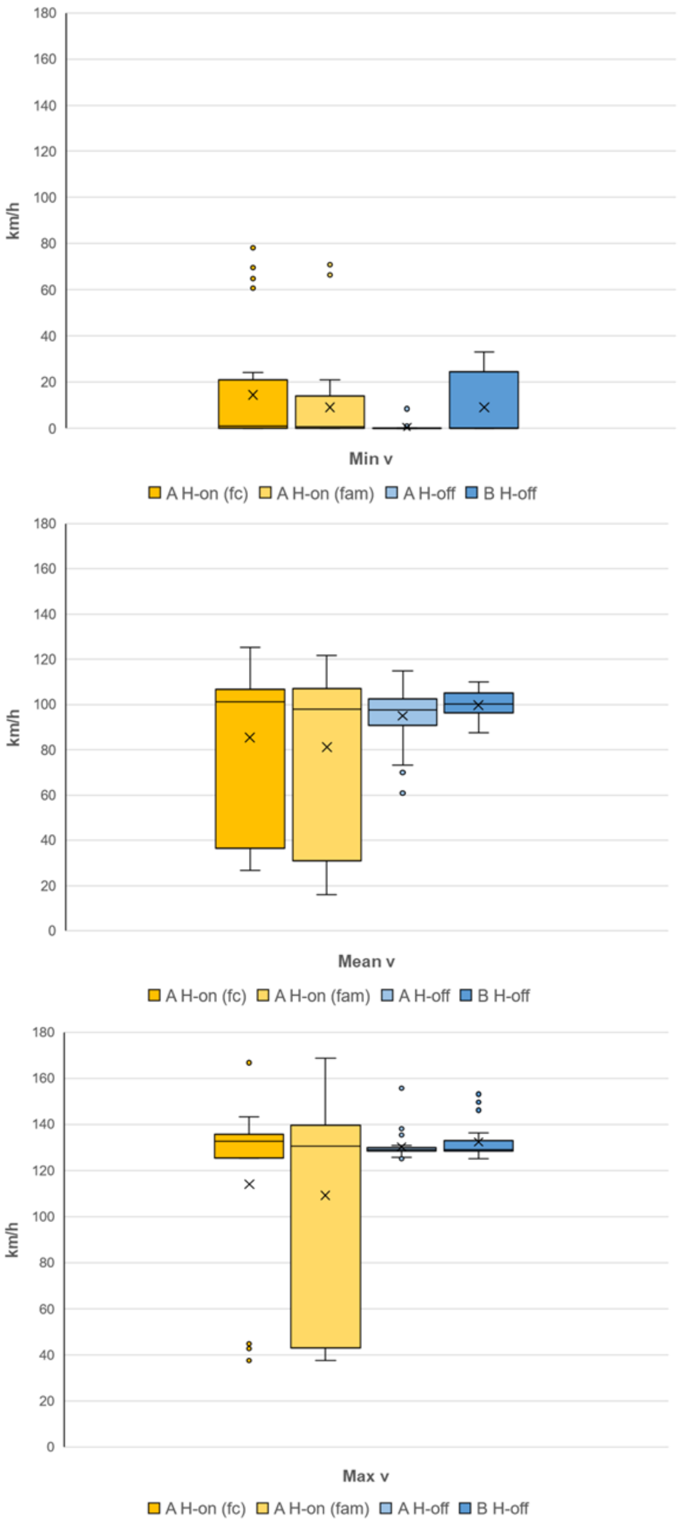


Figure 4-15: Visualization of the distribution of the min, mean and max speeds for A H-on (fc) (n = 27), A H-on (fam) (n = 28), A H-off (fc) (n = 29) and B H-off (fc) (n = 29). The sample size indicates the number of participants

Figure 4-16 depicts the availability of the system levels during each 45min-interval. We see that the L2H-off mode was available on average for 82-83 %, the same applies to the L2H-on mode for the L2H-on test vehicles. We also see a few outliers with relatively short duration of L2-function availability which is due to bad weather conditions. In addition, there is a small proportion for L2H-on mode (3-5 %) for the L2H-off vehicles indicating switches between L2H-off and L2H-on modes. We can also observe 15-18 % proportions of L0/L1 indicating phases of deactivated functions. However, the L0/L1 proportions are similar between the groups.

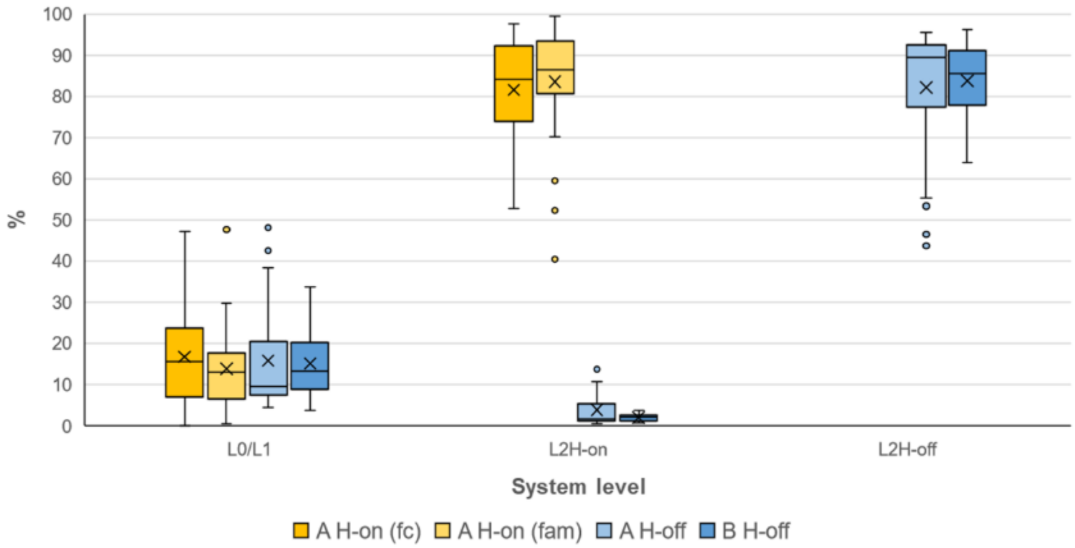


Figure 4-16: Visualization of the distribution and the mean availability of system levels during the 45min-intervals for A H-on (fc) (*n* = 28), A H-on (fam) (*n* = 27), A H-off (fc) (*n* = 29) and B H-off (fc) (*n* = 29). The sample size indicates the number of participants.

Overall, we can conclude that the drives between the groups are comparable in terms of speed profile and availability of the L2 function.

4.4.2.5 Dependent Variables

The collected data can be categorized into observational and self-reported data. A comprehensive overview can be found in Table 4-1 and Table 4-2, respectively. Figure 4-17 gives an overview of the temporal process of the rides in order to understand for which periods or events the metrics apply. Overall, four different 45min intervals are of interest. There are two intervals for L2H-on rides differing in the duration of function use (experience) as first contact vs. familiar which is defined in terms of driven kilometers and day of use. The other two intervals belong to L2H-off rides, one for each group. Before every interval, except the L2H-on familiar interval, a demographic questionnaire had to be completed as well as a briefing and a training drive took place which is not relevant for the driving performance evaluation. Also, the periods between the two L2H-on intervals are not considered in the analysis. The observational metrics cover the entire 45min intervals independently from the mode (L0, L1, L2) and/or apply for specific transitions which are driver-initiated activations (DA)/deactivations (DD)/lane changes (DL) and system-initiated deactivations (SD). For the transitions, a period of 30 s before and 10 s after the event is considered differentiating between a pre and post-stage. For the metrics regarding the construct of endangerment, the period is decreased to 5 s before and 5 s after the event. After each interval, a follow-up questionnaire had to be completed which covers the self-reported data. Finally, after the L2H-on rides as a whole and after the L2H-off rides interviews were conducted in order to add qualitative insights.

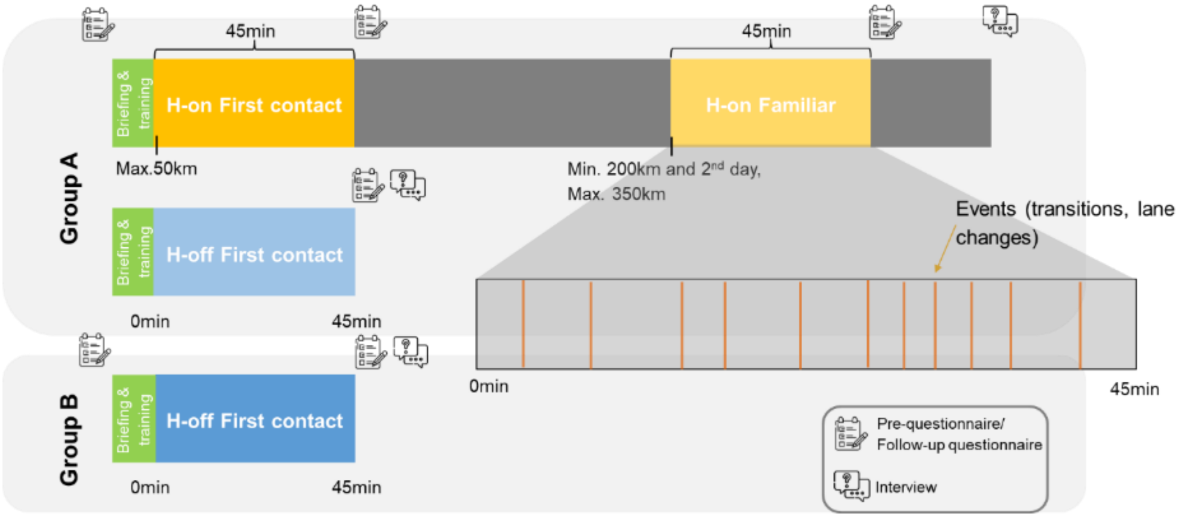


Figure 4-17: Overview of the temporal process of the rides and the analyzed periods and events (icon source: Flaticon.com)

Table 4-1: List of observational metrics assigned to CQs

Construct	Metric	Unit	Time/Event	Database	CQ
Visual attention; perceptual readiness	Eyes-off road glances above 2 s	Number	45 min interval, Transitions (pre, post)	Video annotation	1
	Attention ratios (Eyes-on road, instrument cluster/steering wheel, other)	Percentage in %			
The motoric ability for safe vehicle guidance	Hands-on/-off proportion		45 min interval, L2H-off mode		1, 5
	Hand position		45 min interval		1
Motoric readiness	Level of motoric control rating	0-8 rating	Transitions (pre, post)		1, 2
Monitoring	Hands-off/Eyes-off warnings	Number	45 min interval	Vehicle bus; Video annotation	1
Reaction time	Reaction time to H-off-/Eyes-off warnings	Time in s Mean, maximum	DMS warnings, System-initiated deactivations	Vehicle bus	2
	H-on reaction time	Time in s		Video annotation	
	Direct control time				
	Difference between H-on reaction time and direct control time				
Type of driver interventions	Driver-/system-initiated deactivations	Number	Driver- & system initiated deactivations	Vehicle bus	
	Type of intervention in case of driver-initiated deactivation	Percentage in % Steering, brake, button		Vehicle bus; Video annotation	

	Type of intervention in case of system-initiated deactivation	Percentage in % Steering, brake, throttle			
Outcome of transitions	Controllability (TOC)-rating (Naujoks et al., 2018)	1-10 rating		Video annotation	2, 5
	Level of Service → Highway capacity manual (2000)	A-F rating			
	TTC (Lead/rear object (LO, RO))	Time in s Minimum	Driver- & system initiated deactivations (pre, post)	Vehicle bus; Lidar	
	Lateral distance	m Minimum			
	Longitudinal and lateral acceleration	m/s2 Minimum, maximum			
Distraction	NDRT engagement	Percentage in %	45 min interval; Transitions (pre, post)	Video annotation	3
Misuse	Hands-off although not allowed (only L2H-on)		45 min interval		
Behavior-based confusion	Hands-off although in L0/L1 mode (only L2H-off)	Time in s Mean, maximum		Vehicle bus; Video annotation	4
	Hands-off although in L2H-on mode (only L2H-off)				
	Attempted activations of L2 although not available	Number			
Objective safety (safety-I)	Accidents, incidents	Number	45 min intervals; incident candidates based on TTC, THW, longitudinal and lateral distance, and longitudinal and lateral acceleration	Vehicle bus; Lidar	5
	Incident classes	Number (Dynamic-based long/lat; vehicle dynamic-based long/lat)			
	Safety criticality level	0-4 rating normal driving, increased risk, crash-relevant, near-crash, crash)		Video observation	
	Controllability TOC-rating (Naujoks et al., 2018)	1-10 rating			

Observational metrics

The observational metrics are used to evaluate the driving performance using quantitative indicators. This metric is based on vehicle bus data and/or video annotation. The video annotation process will be explained in Section 4.4.2.6. In the following, some specific metrics are explained in more detail. The visual attention metrics are based on three areas of interest (AOIs): road, instrument cluster/steering wheel, and others (see Figure 4-18). Here, the road scene is defined by the driver-sided windshield. Mirrors and the passenger-sided windshield is not considered due to difficulties in precisely annotating the eyes-on-road gazes if these AOIs would have been considered as road scene.

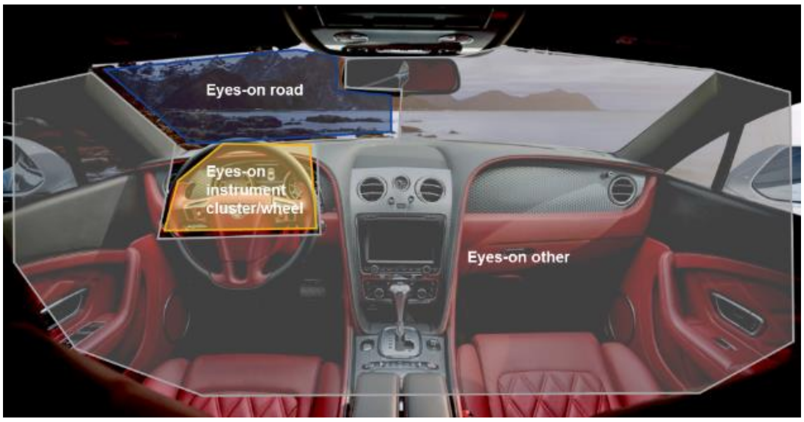


Figure 4-18: Visualization of the three areas of interest for gaze behavior (image source: Pixabay.com)

The hand positions are distinguished into three categories: both hands-on, both hands-off, and different (one hand-on and one hand-off). In addition, the basic hand positions are precisely annotated according to the following scheme (see Appendix C – Taxonomy hand positions):

<u>Hand-on:</u>

- Position steering wheel (Jonsson, 2011): 0 (Bottom), 1, 2, 3 (Top-right), 10, 11, 12 (Top-left)
- Types of grip (Götz, 2007): Contact grip, Grasp, Grasp grip, and other
    - 1-5 finger (F), hand, thumb, Ball of the thumb, wrist, knee, other

<u>Hand-off:</u>

- Grasp space: A, B, C
- Activity (Fleischer & Chen, 2020): working, resting

The hand position control rating is based on a scale from 1-8. The higher the value, the higher the motoric control in terms of steering by the driver. Each hand can achieve a maximum of four points according to the following scheme based on the grasp space or activity in case of hand-off and the position on the steering wheel in case of hand-on which is inspired by Walton & Thomas (2005):

- Hand-off: C or working = 0
- Hand-off: B = 1
- Hand-off: A = 2
- Hand-on: Bottom = 3
- Hand-on: Top = 4

The Level of Service is a metric for traffic density which consists of six levels from A-F according to the Highway capacity manual (2000):

- Level A: Free flow
- Level B: Reasonably free flow
- Level C: Stable flow, at or near free flow
- Level D: Approaching unstable flow
- Level E: Unstable flow, operating at capacity
- Level F: Forced or breakdown flow

In terms of objective safety, the 45 min intervals, inter alia, were filtered for so-called incident candidates indicating potential critical situations. The basic idea behind the analysis of potentially critical situations is the assumption that such situations are predecessors of accidents. Furthermore, accidents happen rarely which is why critical situations are more appropriate as an indicator.

It is also expected that a measured reduction of less severe events (e.g., incidents) allows concluding a reduction of more severe events such as accidents (Faber et al., 2012). Therefore, through an analysis of changes in incident frequency, impacts on accident frequency can be inferred. The process of detecting and analyzing the incident candidates is threefold (see Figure 4-19). First, incident candidates are detected and classified in a rule-based manner based on vehicle data by threshold criteria (see Figure 4-20 and Figure 4-21).

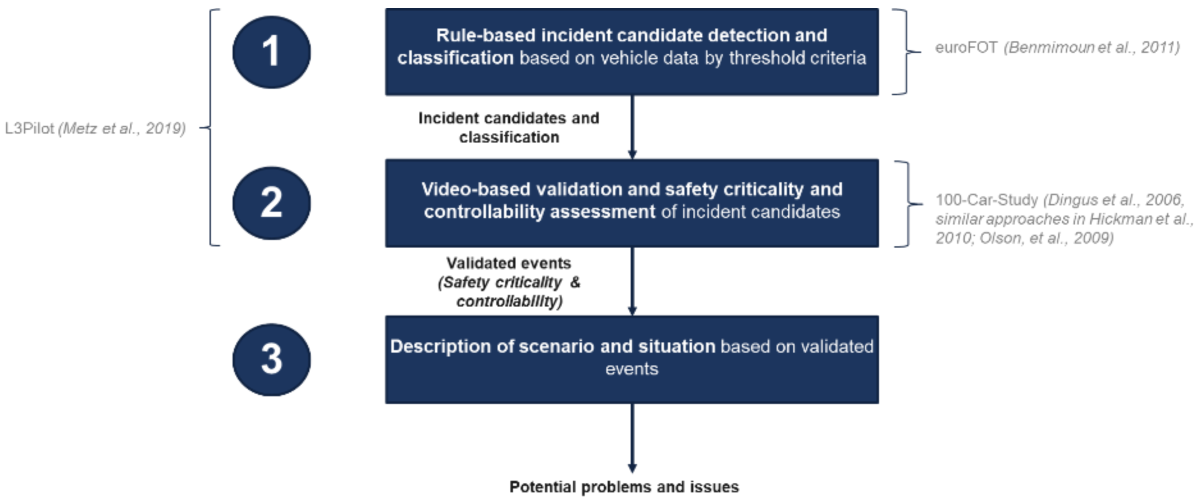


Figure 4-19: Process of detecting and analyzing incident candidates

Incident type		Metrics	Criteria
Distance-based	Front	THW [sec] TTC [sec] Δv [km/h]	Forward THW < 0.35 s & Δv < 20 km/h Forward THW < 0.5 s & Δv > 20 km/h Forward TTC < 1.75 s
	Side	Distance [m] TTC to rear [sec]	Distance to side vehicle < 0.5 m & projected TTC to vehicle in target lane < 1.75 s to vehicles approaching from rear (in case of lane change)
	Rear	THW to rear [sec] TTC to rear [sec] Δv to rear [km/h]	Rear THW < 0.35 s & Δv < 20 km/h Rear THW < 0.5 s & Δv > 20 km/h Rear TTC < 1.75 s
Vehicle dynamics-based		ax [m/s²] ay [m/s²]	Longitudinal acceleration: ax < - 6 m/s² (at 50 km/h) ax < - 4 m/s² (at 150 km/h) Lateral acceleration: ay >= 2.5 m/s² (at 0 km/h) ay >= 7 m/s² (at 50 km/h)

Figure 4-20: Threshold criteria for incident types and metrics, adapted by L3Pilot (Metz et al, 2019)

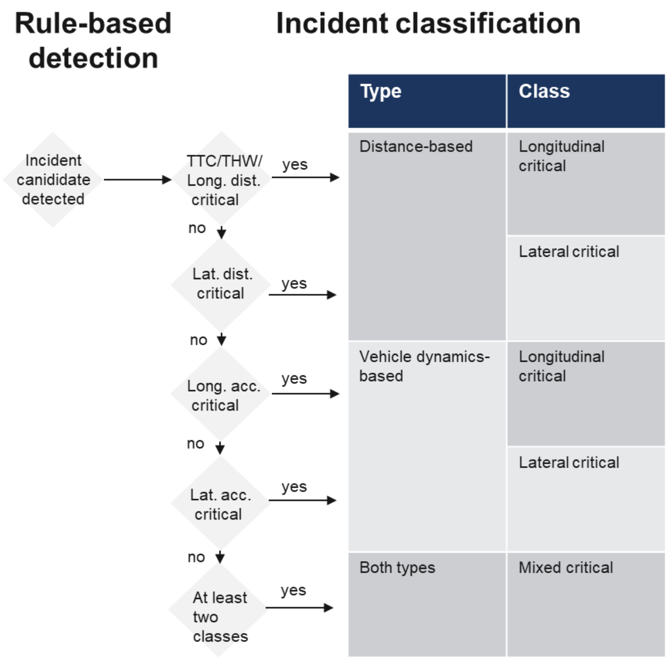


Figure 4-21: Process level 1 of incident classification, adapted by L3Pilot (Metz et al, 2019)

It should be noted that the lane deviation is not considered as a filtering metric due to technical processing issues. Hence, it is possible that a very few potential safety-critical situations are overseen that are not identifiable through the considered metrics. Second, the identified incident candidates are validated through video data in order to evaluate the safety criticality and controllability (see Figure 4-22 and Figure 4-23). This process is adapted by the euroFOT (Benmimoun et al, 2011), the 100-Car-Study (Dingus et al., 2006, similar approaches in Hickman et al. (2010) and Olson et al. (2009)), and the L3Pilot (Metz et al., 2019). Finally, safety-critical events are analyzed in more detail in order to identify potential systemic problems and issues.

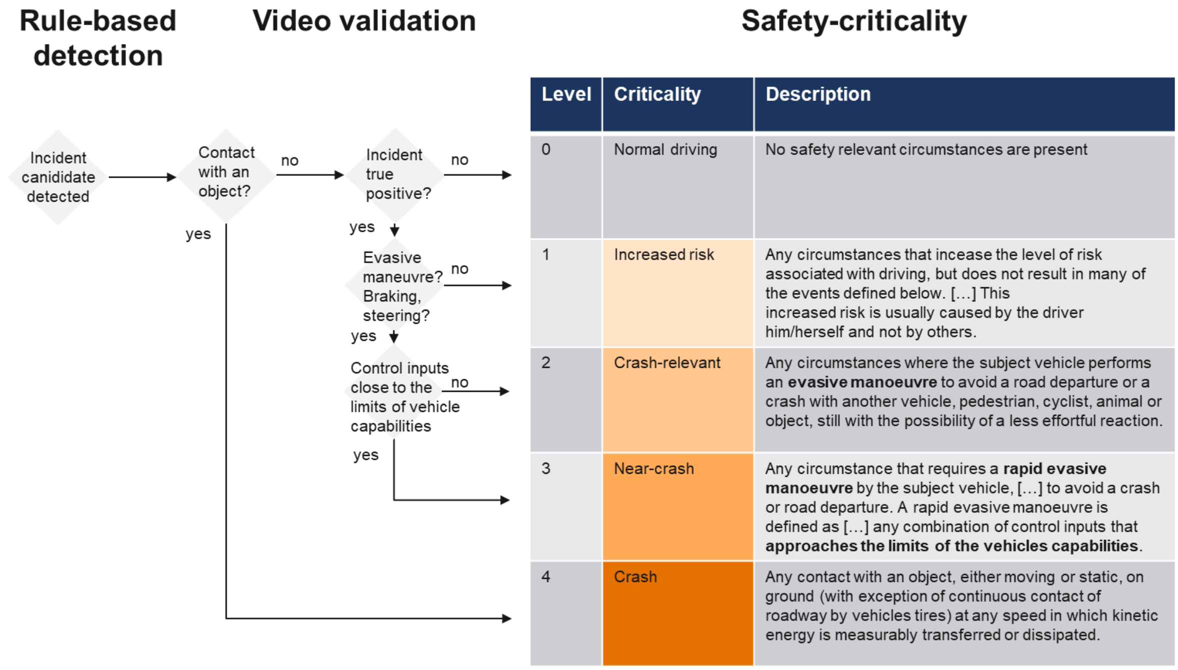
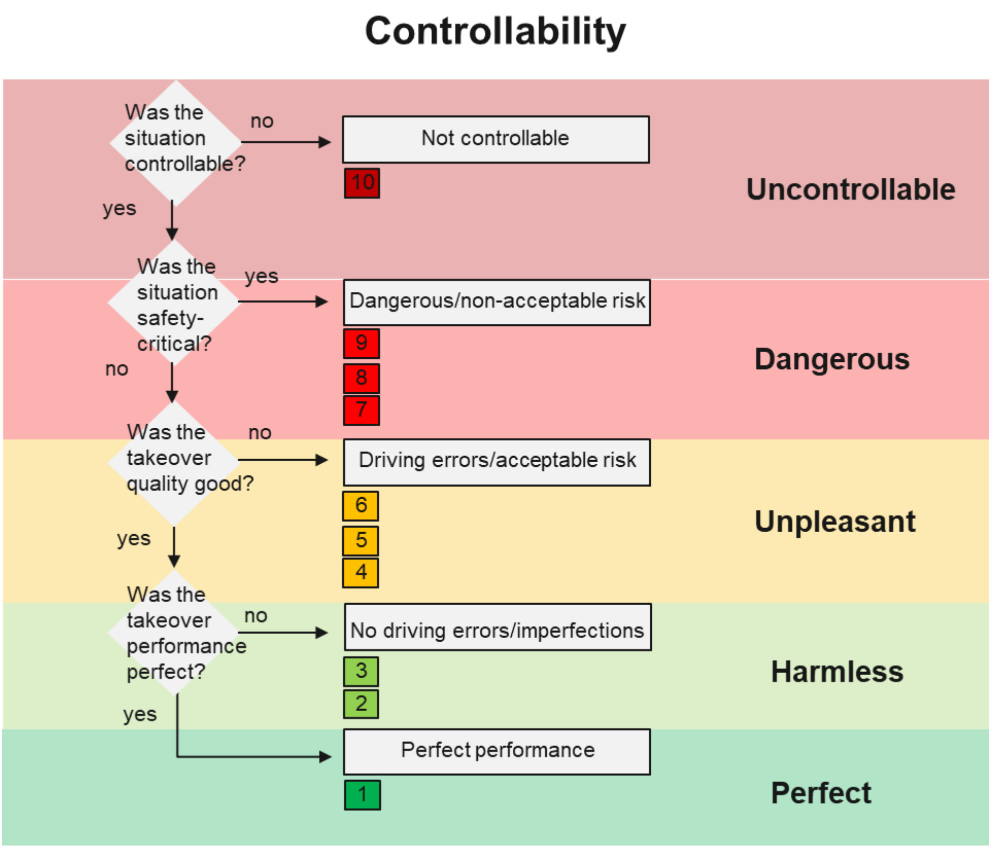


Figure 4-22: Process level 2 of video-based validation and safety criticality assessment, adapted by L3Pilot (Metz et al, 2019)



[…]

Figure 4-23: Process level 2 of controllability assessment and the assessment sheet, adapted by Neukum et al. (2008) and Naujoks et al. (2018)

Self-reported metrics

In order to gain insights into the participants' impressions and mental comprehension of L2, self-reported data are collected (see Table 4-2). During their test rides, participants could comment on specific situations such as critical situations or surprising system behavior via pressing the trigger button. A driver logbook enabled participants of group A during the H-on test phase to further elaborate on their observations. After each interval (3x group A, 1x group B), participants fill out an online survey. Additionally, an interview is conducted by the experimenter after the L2H-off test rides (group A and group B) as well as after participants of group A return to the L2H-on test vehicle.

The online survey comprises a questionnaire on trust (Körber, 2019), an adjusted questionnaire on acceptance (Osswald et al., 2012), and a self-developed questionnaire on the understanding of the system and the role comprising the awareness of the tasks and roles as drivers, ODD understanding, and system functioning. Participants are asked to estimate their engagement in different NDRTs. Also, participants estimate their performance in monitoring the system. A final question addresses the potential influence of the study setting on the participants' behavior. Free text fields are offered for explanations and further comments. Participants are reminded that questions refer to L2 incl. its behavior and its HMI.

Table 4-2: List of self-reported metrics assigned to CQs

Construct	Metric	Description	CQ
Disuse and misuse; Perceived safety	Trust	Trust in automation (Körber, 2019)	3,5
	Acceptance	Car Technology Acceptance Research Model (CTAM, Osswald et al., 2012). Included subscales: Performance expectancy (item PE2 excluded); Effort expectancy; Attitude towards using technology; Facilitating conditions (item FC4 excluded); Behavioral intention to use the system; Perceived safety Excluded subscales: Anxiety; Self-Efficacy; Social Influence Selected and translated by fka, ika, & LfE	
Knowledge-based confusion	System understanding	8 (17) statements Created by fka, ika, & LfE	4
	Role understanding	8 (11) statements Created by fka, ika, & LfE	
Distraction	Subjective NDRT engagement	List of 8 activities + free text field for further activities Created by fka, ika, & LfE	3
Monitoring	Estimated monitoring performance	1 item Created by LfE	1
Other	Estimated influence of test setting	1 item + free text field for elaborations Created by LfE	3
Perceived safety	Preferred L2 system	1 item Created by LfE	5
	L2 intention to use	1 item Created by LfE	

After returning the L2H-on vehicle (after an interval of L2H-on familiar) and after the h-off test rides, an interview is conducted by the experimenter. The experimenter leads through a series of questions focusing on the following aspects: (1) reasons for pressing the trigger button; (2) driver-initiated transitions; (3) system-initiated transitions; (4) DMS; (5) HMI; and the (6) system behavior. Additionally, participants are asked to estimate their intent to use L2H-on/H-off, if this would be available in their private car. In more detail, a follow-up question addresses the estimated intent to use specific L2 components (lateral support, longitudinal support, h-off option (only L2H-off)). Participants of group A are asked for their preference for the L2H-off or L2H-on option after having experienced both functions.

4.4.2.6 Data analysis

Experimenter protocols and driver logbooks are used to support the data analysis. The FOT DE is exploratory and provides insights into driver behavior for vehicles equipped with L2 automated driving systems. Results are interpreted based on a descriptive analysis with a strong focus on qualitative input provided by the participants. The findings are used to answer the five CQs as well as to derive hypotheses for the subsequent simulator studies planned in this project.

The video data are annotated with the event-logging software BORIS (Behavioral Observation Research Interactive Software) by Friard and Gamba (2016). The version used was “BORIS

7.13.9 for Windows". The software differentiates between point events and state events. While point events refer to a single point of time and are e.g., used to describe the beginning and end of an interval, state events refer to a time in which a specific state prevails, e.g., looking at the instrument cluster.

Objective raw datasets from the FOT vehicles are filtered, pre-processed and provided by ika: The vehicle bus and Lidar data are imported through Python (v3.9.12) by the means of the asammdf-packages (v7.0.7) and pre-processed by numpy (v1.22.3) and pandas (v1.4.2) in order to identify the 45 min intervals and transitions as well as calculate the observable metrics. Ultimately, the annotated data as well as the pre-processed vehicle bus and Lidar data are processed further by LfE for aggregation, analysis, and visualization via the software Matlab (R2022a) and Excel 365. Self-reported data are gathered through the survey tool LimeSurvey (v5.4.10) and are processed for analysis and visualization via the software RStudio (v2022.02.1) and R (v4.2.0).

4.4.3 Results

In the following, the results are presented according to each CQ and the respective constructs and metrics. The findings regarding RQ2 and RQ3 are answered as comparing L2H-on vs. L2H-off if there are no other differences and findings with regards to RQ4 collectively refer to groups A and B as L2H-off if there are no differences between these two groups. Results based on the interviews or others are mentioned when appropriate.

4.4.3.1 Hands-off = mind-off?

Visual attention

There is no difference in the number of eyes-off road glances above 2s between L2H-on and L2H-off (see Figure 4-24). Overall, the mean number is relatively small for both systems.

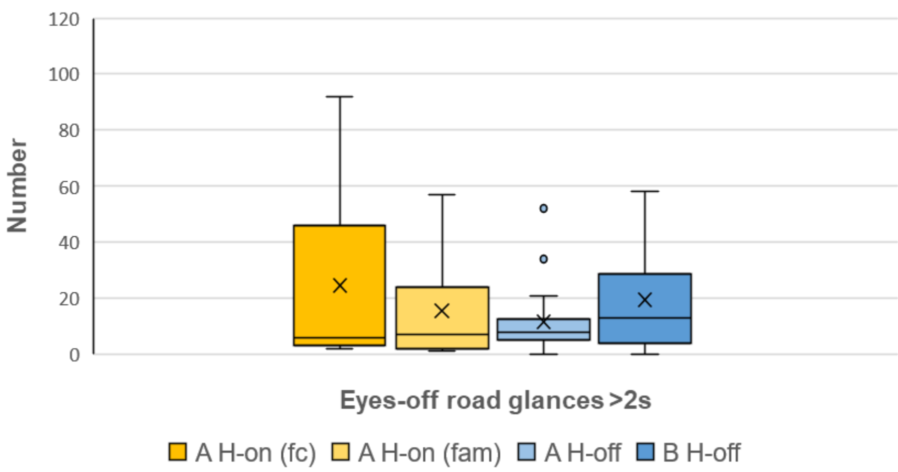


Figure 4-24: Visualization of the distribution of eyes-off road glances above 2s for A H-on (fc) (n = 19), A H-on (fam) (n = 19), A H-off (fc) (n = 19) and B H-off (fc) (n = 26). The sample size indicates the number of participants

L2H-off participants on average show a bit higher visual attention ratio regarding eyes-on road than L2H-on subjects (see Figure 4-25).

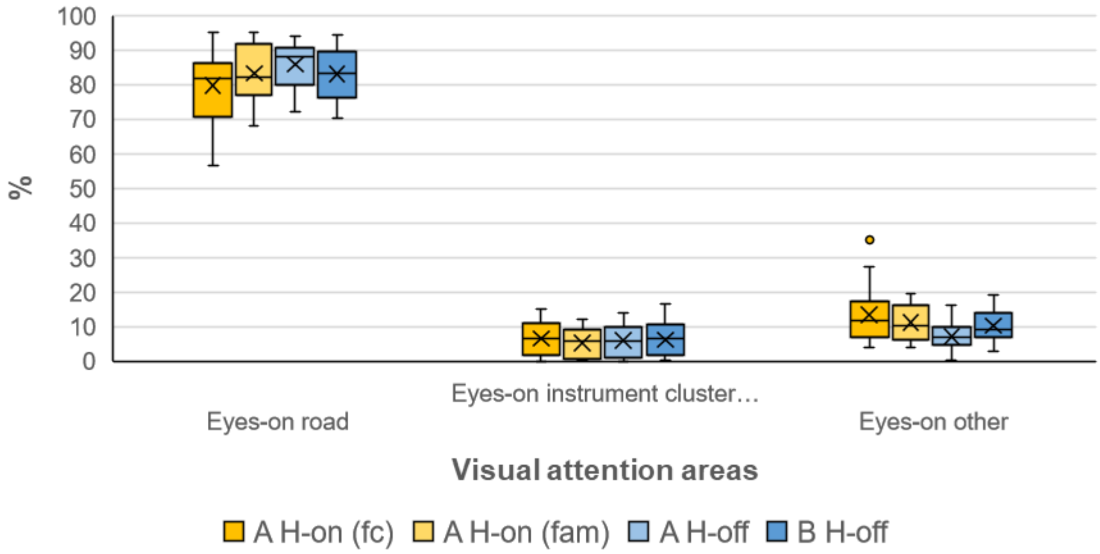


Figure 4-25: Visualization of the distribution of visual attention ratios for A H-on (fc) (n = 19), A H-on (fam) (n = 19), A H-off (fc) (n = 19) and B H-off (fc) (n = 26). The sample size indicates the number of participants

Additionally, the dispersion for L2H-on drivers is a little higher than for L2H-off subjects. In general, visual attention to the road is on a high level for L2H-off participants. The visual attention to the instrument cluster/steering wheel is similar.

Perceptual readiness

Especially before driver-initiated activations, L2H-off drivers show a higher mean and variance in the number of eyes-off-road glances above 2s than L2H-on drivers but still on a low level (see Figure 4-26). Figure 4-27 shows that this is associated with lower visual attention to the road and higher visual attention to the instrument cluster/steering wheel each before the driver-initiated activation for L2H-off subjects than L2H-on subjects. This finding indicates a tendency of higher visual attention regarding checking the automation mode and status in terms of system activation for L2H-off than for L2H-on. However, we have to point out that the sample size of events is rather small making the former mentioned finding a tendency rather than a proof of evidence. The subjects also reported in the interviews that L2H-off is more complex to use than L2H-on which coincides with the former finding.

Before system-initiated deactivations, L2H-off participants show a clearly higher eyes-on-road proportion than L2H-on participants indicating that L2H-off drivers users tend to be more alert in terms of system limits than L2H-on users (see Figure 4-28). However, we have to emphasize that the sample size of events is quite small making the former mentioned finding a tendency rather than a proof of evidence.

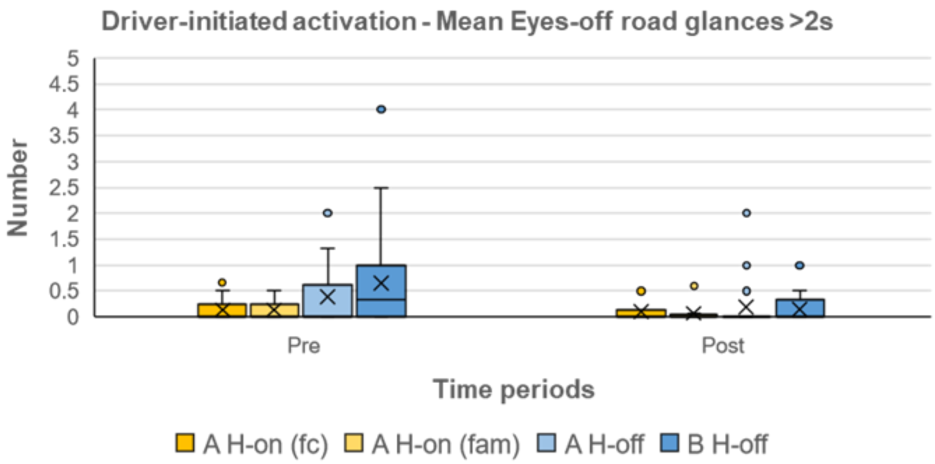


Figure 4-26: Visualization of the distribution of eyes-off road glances above 2s at driver-initiated activations for A H-on (fc) (n = 10), A H-on (fam) (n = 9), A H-off (fc) (n = 18) and B H-off (fc) (n = 23). The sample size indicates the number of events

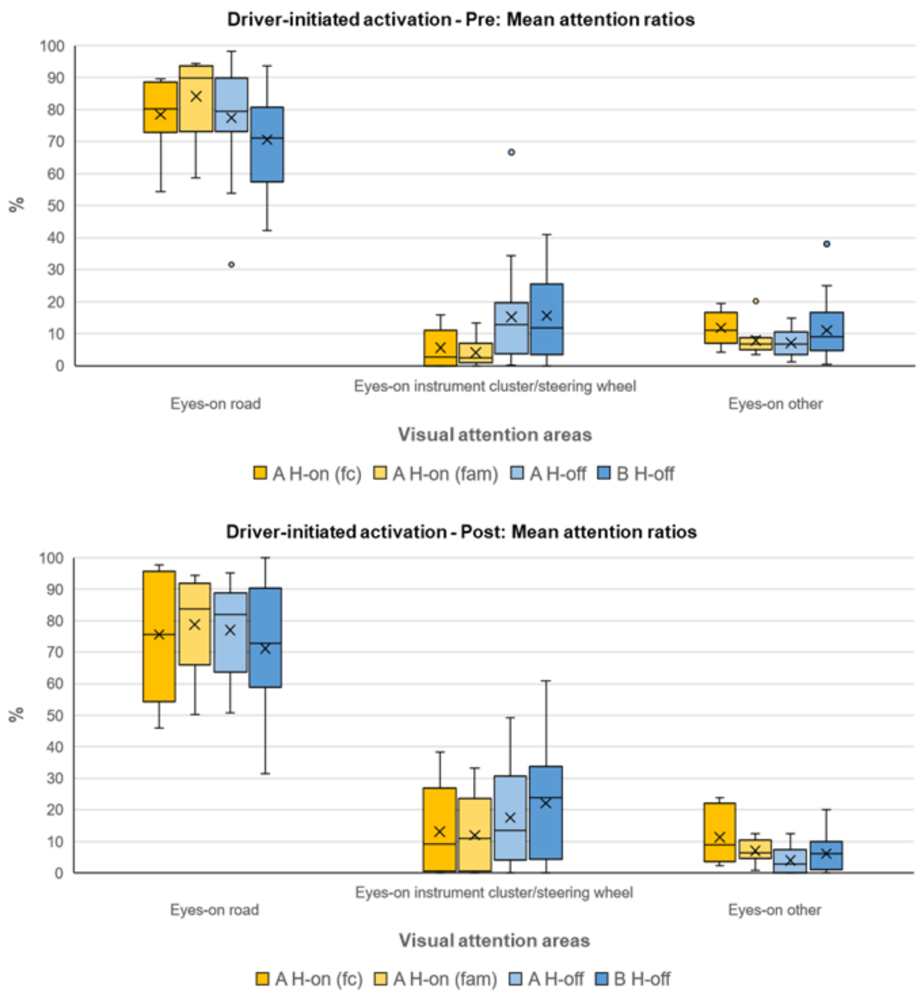


Figure 4-27: Visualization of the distribution of visual attention ratios at driver-initiated activations (pre and post stages) for A H-on (fc) (n = 10), A H-on (fam) (n = 9), A H-off (fc) (n = 18) and B H-off (fc) (n = 23). The sample size indicates the number of events

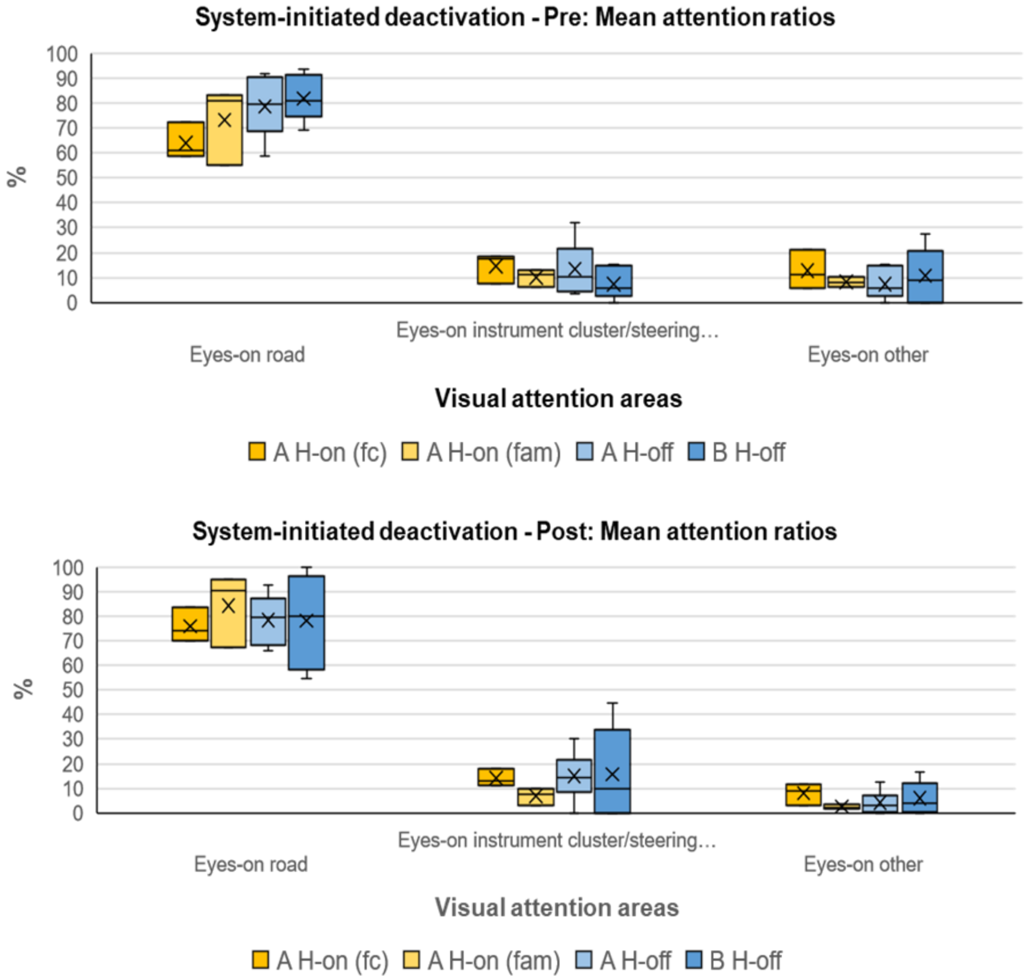


Figure 4-28: Visualization of the distribution of visual attention ratios at system-initiated deactivations (pre and post stages) for A H-on (fc) (*n* = 3), A H-on (fam) (*n* = 3), A H-off (fc) (*n* = 8) and B H-off (fc) (*n* = 6). The sample size indicates the number of events

Monitoring

On average, the number of total hands-off warnings (L2H-on) and total eyes-off warnings (L2H-off) is on a low and similar level (see Figure 4-29). However, the L2H-on users show higher outliers in terms of hands-off warnings.

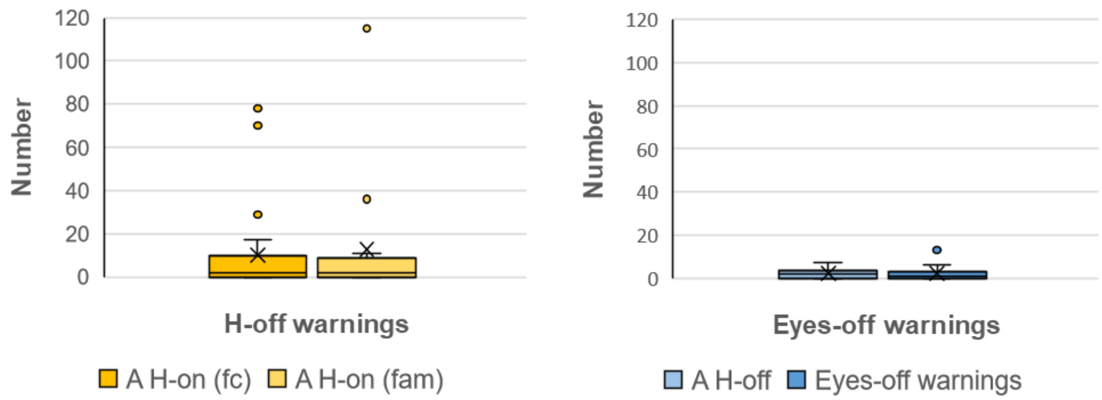


Figure 4-29: Visualization of the distribution of total HOD (L2H-on) and DMS (L2H-off) warnings, respectively, for A H-on (fc) (*n* = 19), A H-on (fam) (*n* = 19), A H-off (fc) (*n* = 19) and B H-off (fc) (*n* = 26). The sample size indicates the number of participants

Figure 4-30 depicts the distribution of hypothetical DMS warnings meaning the eyes-off warnings that would have been emitted by a 3-stage DMS (EOR: 5 s, HOR: 8 s, DDCR: 13 s) in order to make the monitoring performance comparable between L2H-on and L2H-off. On average, the warnings are on a low level for both systems and no difference between both systems can be found. However, L2H-on subjects show very high outliers. Additionally, the warning cascade is predominantly terminated in the first warning stage.

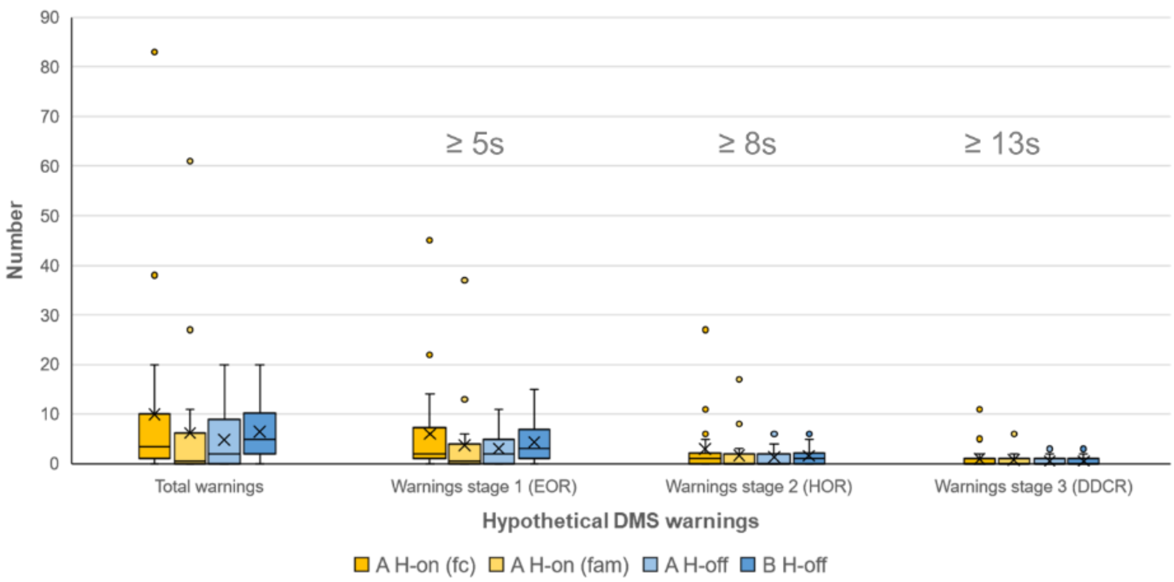


Figure 4-30: Visualization of the distribution of hypothetical DMS warnings for A H-on (fc) (*n* = 19), A H-on (fam) (*n* = 19), A H-off (fc) (*n* = 19) and B H-off (fc) (*n* = 26). The sample size indicates the number of participants

The monitoring performance almost is estimated as always/usually attentive and on average no difference between L2H-on and L2H-off can be identified (see Figure 4-31). This subjective assessment of monitoring coincides with the objective results.

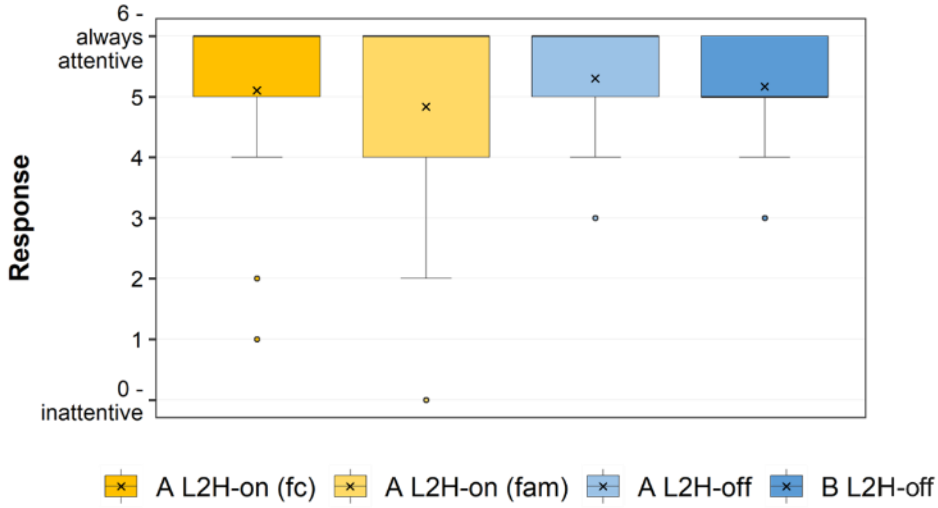


Figure 4-31: Visualization of the distribution of estimated monitoring performance for A H-on (fc) (n = 29), A H-on (fam) (n = 30), A H-off (fc) (n = 30) and B H-off (fc) (n = 30). The sample size indicates the number of participants

In the interviews, the warning cascade by the DMS was evaluated well in terms of time, and warnings are clearly perceived and reasoned.

<u>The motoric ability for safe vehicle guidance</u>

Figure 4-32 illustrates the hands-off proportion. L2H-off users have their hands on the steering wheel on average 45 %. However, a large spread exists. It can be argued that this indicates an awareness of the need for hands on the steering wheel in case of L2H-off use. Further, it is also an indication of balanced trust. During L2H-on use, the drivers mostly have their hands on the steering wheel but some outliers with relatively high hands-off proportion exist indicating the potential for misuse.

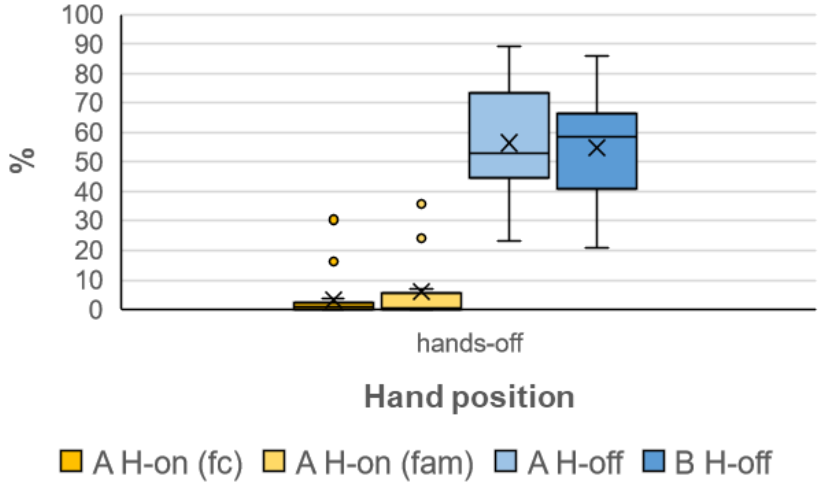


Figure 4-32: Visualization of the distribution of hands-off ratio for A H-on (fc) (n = 19), A H-on (fam) (n = 19), A H-off (fc) (n = 19) and B H-off (fc) (n = 26). The sample size indicates the number of participants

If L2H-off subjects have their hands on the steering wheel, then on average they have rather both hands on than just one hand (see Figure 4-33). Instead, L2H-on subjects on average take both hands/one hand on the steering wheel with a proportion of 60 % and 40 %, respectively. So, there is relatively large period of just one hand control. However, a very large spread exists for both hands or just one hand on the steering wheel indicating a subject-related effect.

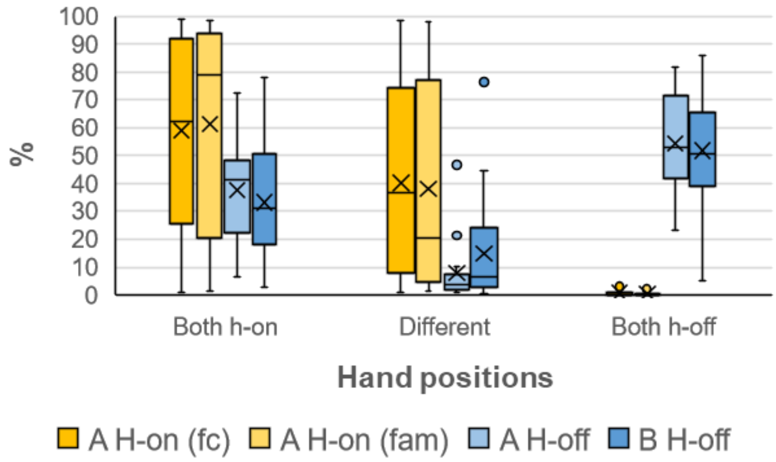
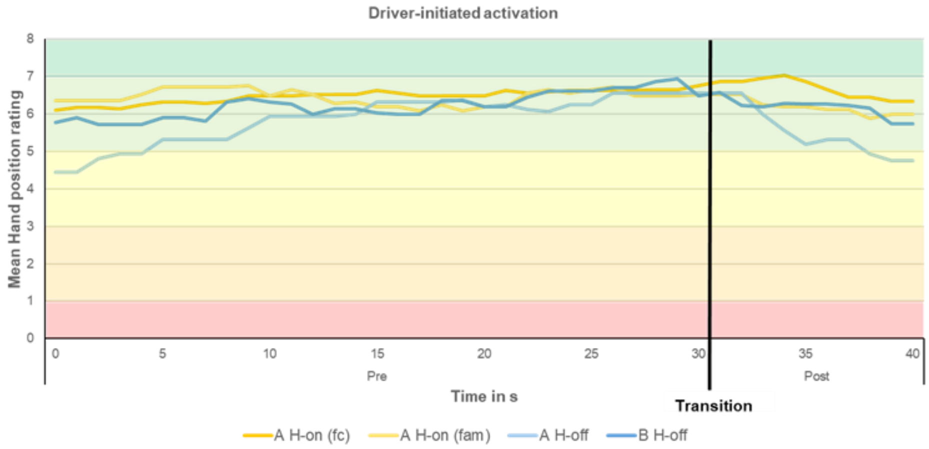


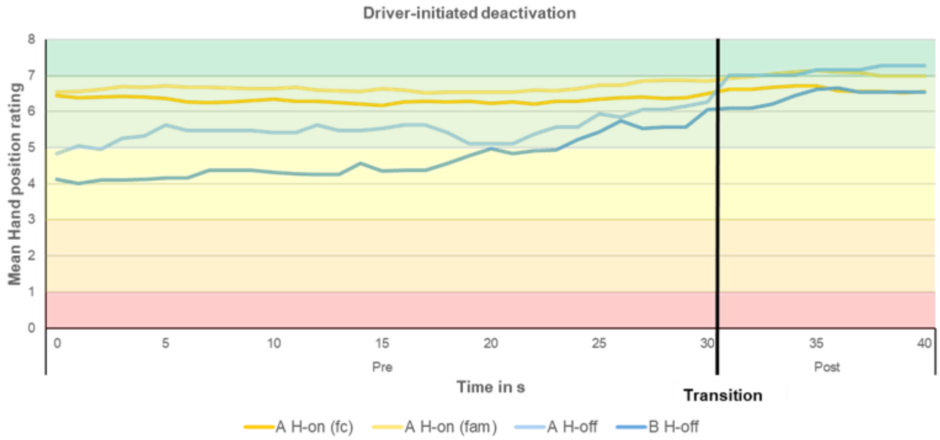
Figure 4-33: Visualization of the distribution of hand position ratios for A H-on (fc) (*n* = 19), A H-on (fam) (*n* = 19), A H-off (fc) (*n* = 19) and B H-off (fc) (*n* = 26). The sample size indicates the number of participants

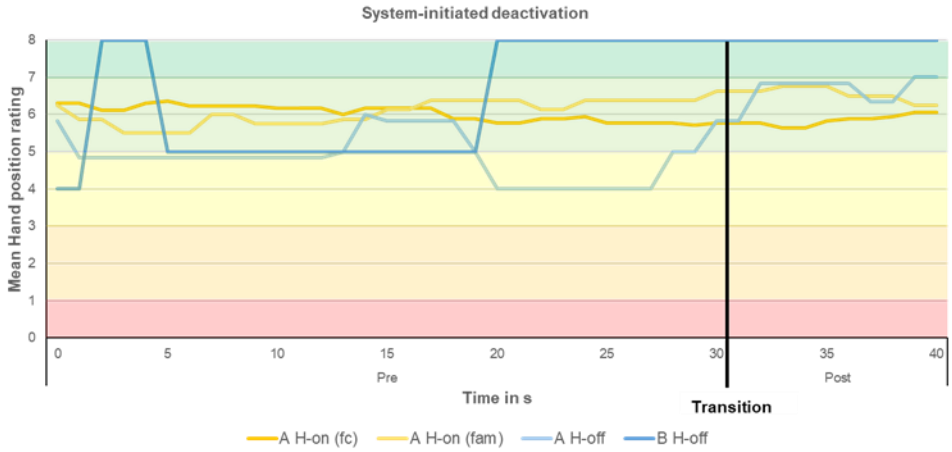
In addition, the most frequent hand positions are analyzed. Here, it can be shown that with L2H-off use, the participants most frequently (30 %) put their hands on their laps and are therefore ready to quickly intervene, and the second (12 %) and third (7 %) most frequently hand positions are both hands or one hand on the upper half of the steering wheel, respectively. Therefore, the hands are mainly placed close to or on the steering wheel. Placing the hands behind the head, far away from the steering wheel (e.g., passenger seat, grabbing for objects), or occupying the hands with objects was not observed during the FOT with the sole exception of infotainment use for a few seconds.

<u>Motoric readiness</u>

Figure 4-34 demonstrates the temporal process of motoric control by the driver at transitions. In all four transition types, it can be seen that the closer the transition, the closer the L2H-off drivers get to the ready-to-drive hand position in the way of taking both hands on the upper steering wheel. Direct to the transition, the motoric control level for L2H-off users are on a similar level to L2H-on users. However, it should be pointed out that the sample size of system-initiated deactivations is quite small making the respective finding a tendency rather than a proof of evidence.


Driver-initiated activation
Mean Hand position rating
8
7
6
5
4
3
2
1
0
0
5
10
15
20
25
30
35
40
Pre
Post
Time in s
Transition
A H-on (fc)
A H-on (fam)
A H-off
B H-off


Driver-initiated deactivation
Mean Hand position rating
8
7
6
5
4
3
2
1
0
0
5
10
15
20
25
30
35
40
Pre
Post
Time in s
Transition
A H-on (fc)
A H-on (fam)
A H-off
B H-off


System-initiated deactivation
Mean Hand position rating
8
7
6
5
4
3
2
1
0
0
5
10
15
20
25
30
35
40
Pre
Post
Time in s
Transition
A H-on (fc)
A H-on (fam)
A H-off
B H-off

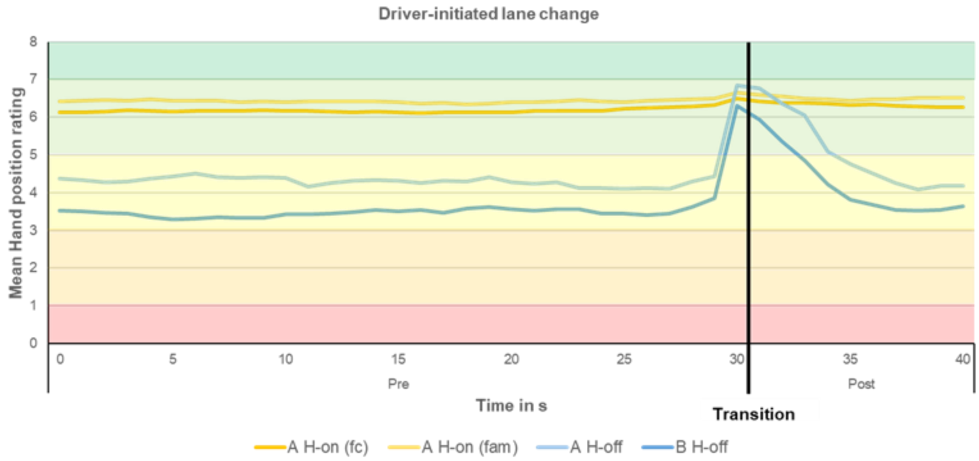


Figure 4-34: Visualization of the temporal process of hand position control rating at transitions for A H-on (fc) (*n* = 28(DA), 96(DD), 16(SD), 282(DL)), A H-on (fam) (*n* = 24, 82, 7, 214), A H-off (fc) (*n* = 15, 18, 5, 192) and B H-off (fc) (*n* = 30, 31, 2, 282). The sample size indicates the number of events

4.4.3.2 Prolonged transition times

Reaction time

Both L2H-on drivers and L2H-off drivers react relatively fast to hands-off or eyes-off warnings considering the meantime (see Figure 4-35). However, L2H-on users show clearly higher maximum reaction times to hands-off warnings than L2H-off users to eyes-off warnings. In addition, the maximum reaction times of L2H-on subjects are mainly high, instead, the maximum reaction times of L2H-off subjects are still at a low level except for one outlier.

Figure 4-36 depicts the hands-on reaction time as well as direct control time at system-initiated deactivations. These system deactivations almost occurred by FDCRs without pre-warning time. A few FDCRs occurred with pre-warning time and just one system deactivation occurred without any FDCR at all (silent system failure/limit). For L2H-off participants, the hands-on reaction time and direct control time are 1.7 s and 1.85 s on average.

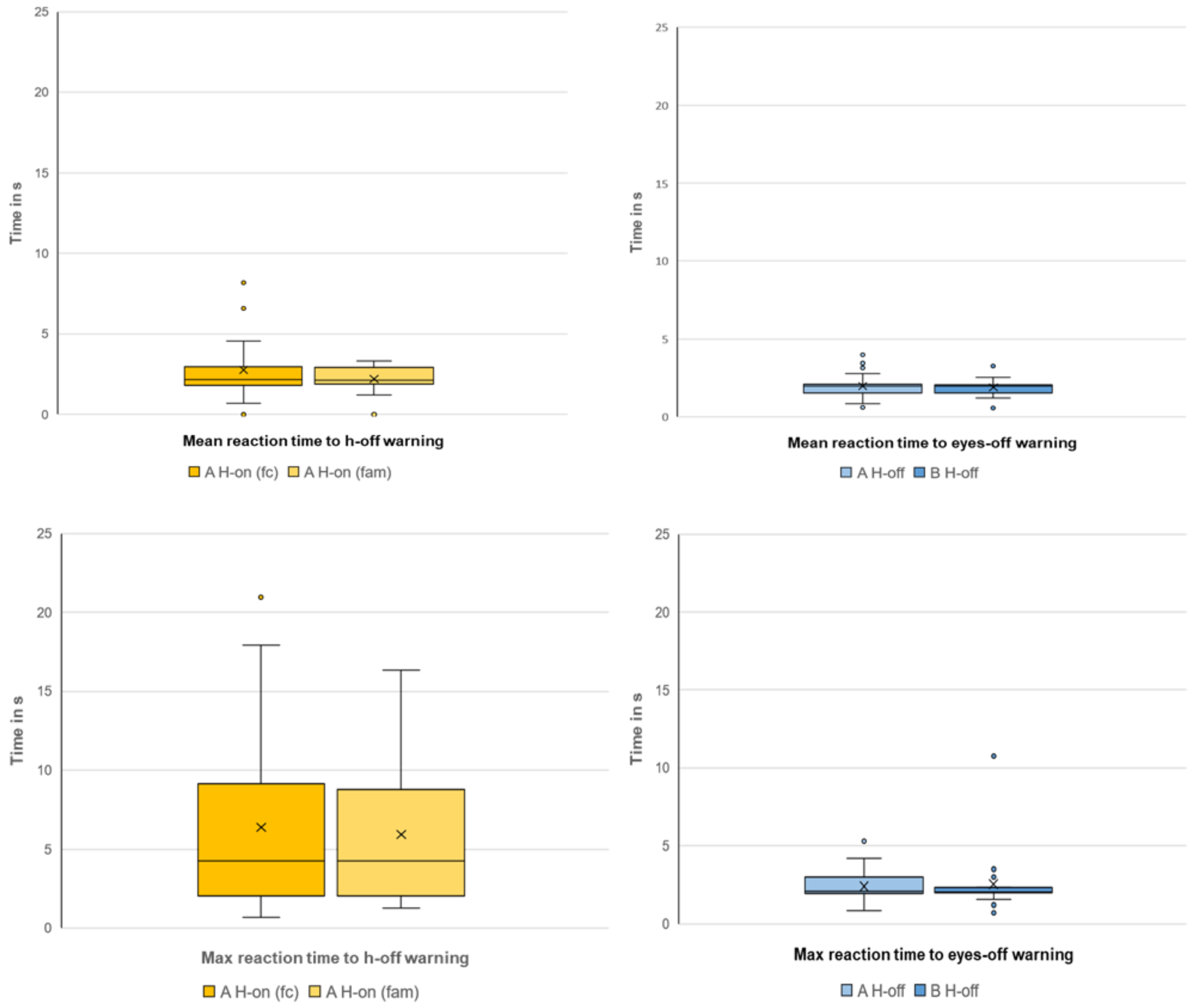


Figure 4-35: Visualization of the distribution of mean and max reaction times to hands-off/eyes-off warnings for A H-on (fc) (*n* = 19), A H-on (fam) (*n* = 19), A H-off (fc) (*n* = 24) and B H-off (fc) (*n* = 20). The sample size indicates the number of participants

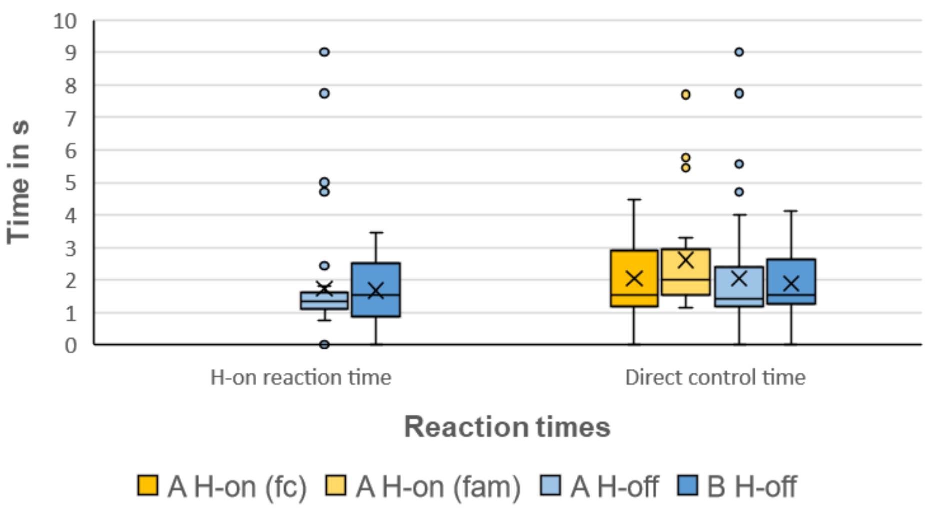


Figure 4-36: Visualization of the distribution of hands-on reaction time (only L2H-off) and direct control time concerning system-initiated deactivations for A H-on (fc) (*n* = 47), A H-on (fam) (*n* = 44), A H-off (fc) (*n* = 63) and B H-off (fc) (*n* = 44). The sample size indicates the number of events

Overall, these are fast reaction times. With regards to the direct control time, no relevant difference can be found between L2H-on and L2H-off. A few high outliers exist for both L2H-on and L2H-off. If these outliers result in safety-critcal events, then they are included in the incident candidate analysis in Section 4.4.3.5. Also, the deviation between hands-on reaction time and direct control time mainly is 0s when considering the median (see Figure 4-37).

From the results, it can be concluded that for L2H-off drivers the physical disadvantage of the lack of steering wheel contact is compensated and does not lead to higher direct control times than L2H-on drivers. It is assumed that the physical disadvantage is compensated by the fact that L2H-off users already make an action decision while putting their hands on the steering wheel and then perform an action immediately with the steering wheel contact. This results in a similar direct control time in comparison to the L2H-on users.

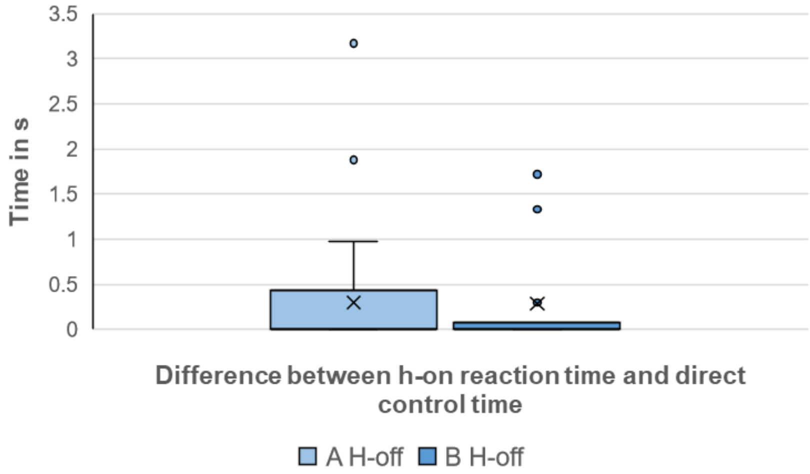


Figure 4-37: Visualization of the distribution of the deviation between hands-on reaction time and direct control time concerning system-initiated deactivations for A H-off (fc) (n = 63) and B H-off (fc) (n = 44). The sample size indicates the number of events

<u>Type of driver interventions</u>

Figure 4-38 shows the distribution of the number of driver-/system-initiated deactivations. In total, 127 driver deactivations occurred for A H-on (fc), 96 for A H-on (fam), 118 for A H-off, and 232 for B H-off. Instead, 47 system deactivations occurred for A H-on (fc), 44 for A H-on (fam), 63 for A H-off, and 44 for B H-off. It can be seen that L2H-off subjects tend to deactivate more often than L2H-on subjects. However, system deactivations are on a low and similar level.

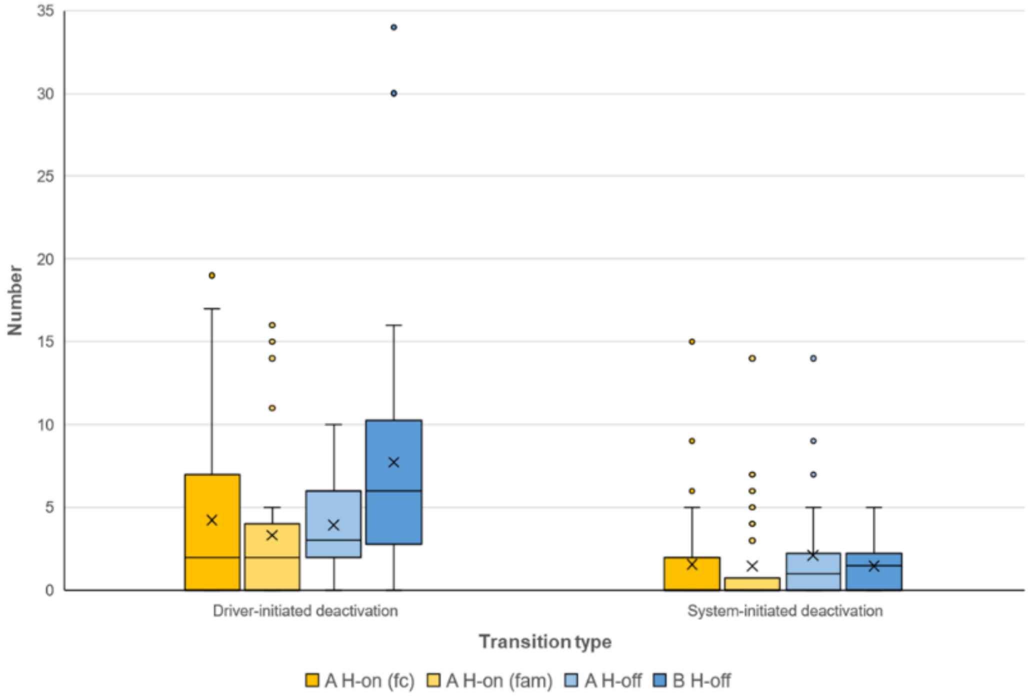
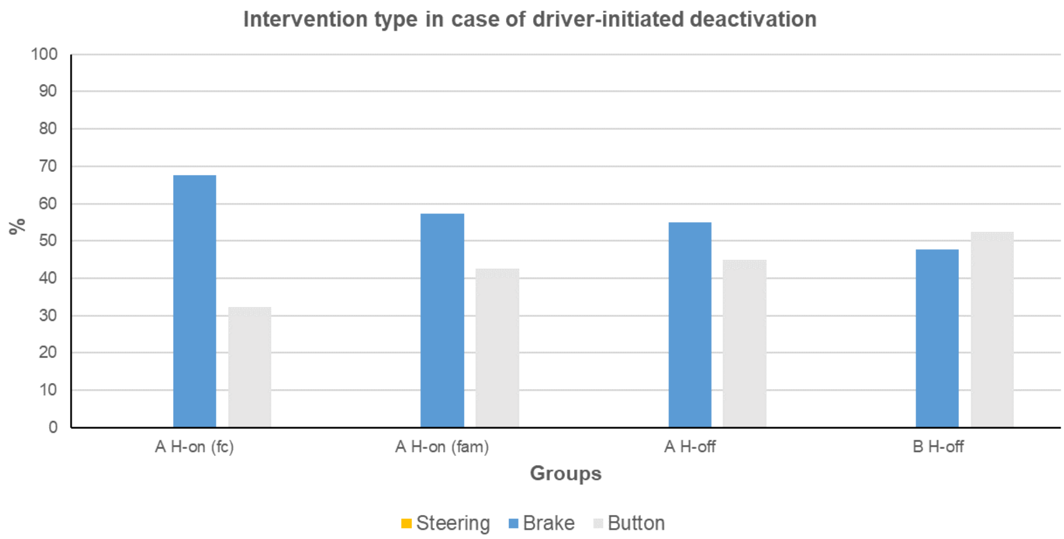


Figure 4-38: Visualization of the distribution of the number of driver-/system-initiated deactivations for A H-on (fc) (n = 30), A H-on (fam) (n = 29), A H-off (fc) (n = 30) and B H-off (fc) (n = 30). The sample size indicates the number of participants



[…]

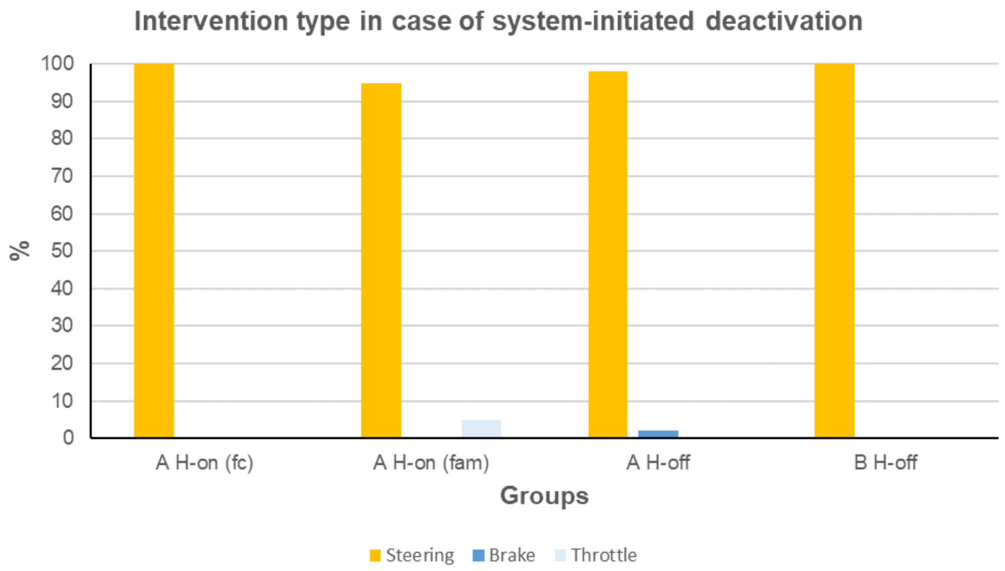


Figure 4-39: Visualization of the relative frequency of intervention types (first driver reaction) in case of driver-/system-initiated deactivations for A H-on (fc) (*n* = 30), A H-on (fam) (*n* = 29), A H-off (fc) (*n* = 30) and B H-off (fc) (*n* = 30). The sample size indicates the number of participants

In the case of driver deactivations, L2H-on participants mainly directly control (first reaction) by braking followed to a less proportion through the button (see Figure 4-39). Instead, for L2H-off participants, this ratio changes towards a balance between braking and button. However, for both systems steering as the first reaction does not happen at all. In the case of system deactivations, the pattern is different. For both systems steering is the predominant first reaction, a few direct controls are accomplished by the gas pedal and only one first reaction is braking. No differences between L2H-on and L2H-off can be found. It should be emphasized that the contextual situations between the driver- and system-initiated deactivations are not necessarily comparable.

<u>Outcome of transitions</u>

In general, when considering the TOC rating, both driver- as well as system-initiated deactivations are rated well-controllable (see Figure 4-40). There are no clear differences between L2H-on drivers and L2H-off drivers. However, two driver-initiated deactivations are dangerous (7-9 TOC-rating) for L2H-off users whereas in the case of system-initiated deactivations two are dangerous for L2H-on users and four for L2H-off drivers. It can be seen that more dangerous events occurred for L2H-off subjects. However, we have to make clear that the sample size of L2H-on participants is considerably smaller than L2H-off participants which is why a final comparison is difficult. Hence, there is rather no difference between both systems concerning dangerous events. In addition, the two dangerous driver-initiated deactivations are associated with a level of service of B and C whereas the six dangerous system-initiated deactivations are associated with a level of service of A-C (two events on each level) (see Figure 4-70 and Figure 4-71). Due to the low sample size, it is impossible to identify a clear relationship between TOC rating and level of service. However, a tendency exists in the way that dangerous events rather happen at free flow up to stable flow in terms of traffic density.

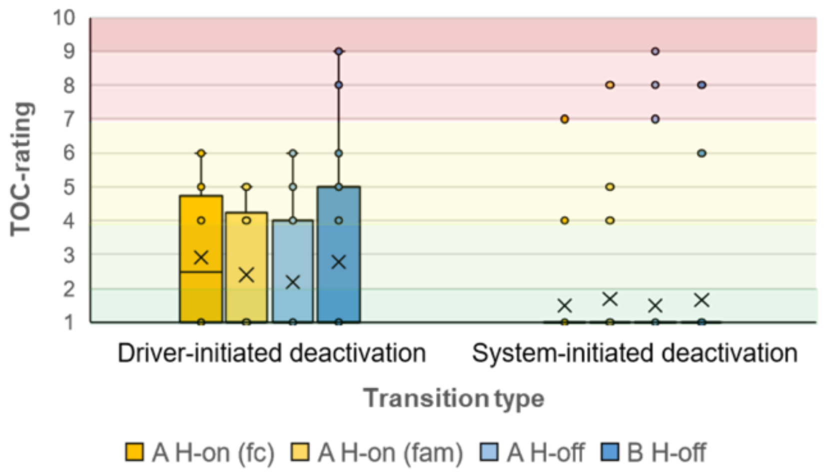


Figure 4-40: Visualization of the distribution of the TOC-rating for A H-on (fc) (*n* = 12(DD), 18(SD)), A H-on (fam) (*n* = 10, 20), A H-off (fc) (*n* = 61, 43) and B H-off (fc) (*n* = 119, 18). The sample size indicates the number of events

Crash-relevant

L2H-on

TOC	Type	Description	Issue
7	SD	Crossing of lane but subject intervend well	End of **roadwork**, **lane detection**
8	SD	Crossing of lane but subject intervend well	**Roadwork**, **bump**

Legend:
DD = Driver-initiated deactivation
SD = System-initiated deactivation

L2H-off

TOC	Type	Description	Issue
7	SD	Crossing of lane but subject intervend well	**Roadwork**, **lane detection**
8	SD	Subject intervened fast and well otherwise lane would have been crossed	**Roadwork**, **lane detection**
8	SD	Little lateral distance	**Roadwork**
8	DD	Crossing of shoulder but subject intervend well	**Roadwork**, **lane detection**
9	DD	Truck pulled over, necessary to leave the lane to avoid truck invading it, subject brakes rapidly and steers quickly to left lane	**Very close cut-in (truck)**
9	SD	FDCR: participant played around with L2H-off control buttons and took over only after 9s	**Mode confusion**

Figure 4-41: Visualization of the crash-relevant TOC-ratings and their transition type, description, and issue for A H-on (fc) (*n* = 12(DD), 18(SD)), A H-on (fam) (*n* = 10, 20), A H-off (fc) (*n* = 61, 43) and B H-off (fc) (*n* = 119, 18). The sample size indicates the number of events (icon source: Flaticon.com)

The eight dangerous events are described in Figure 4-41. It can be seen that for L2H-on users as well as L2H-off users the main issue is roadworks due to lane detection, bumps, or little lateral distance to the crash barrier. One dangerous driver-initiated deactivation is due to a close cutting-in truck and one dangerous system-initiated deactivation is due to mode confusion. In most events, the driver reacts or intervened well but overall the situations have to be determined as crash relevant in terms of safety criticality.

Ultimately, the four transition types are assessed in terms of criticality metrics (TTC, longitudinal and lateral acceleration) (see Appendix C – Taxonomy hand positions). In general, it can be seen that driver- and system-initiated deactivations are not critical and no differences exist between L2H-on or L2H-off drivers. Just a few events can be identified as incident candidates which will be presented in more detail in Section 4.4.3.5. It should be pointed out that the sample size for L2H-on related events with regards to the TTC is small due to a loss of data which is why both systems are difficult to compare in terms of TTC. In principle, the lateral distance should also be evaluated. However, in most transitions (deactivations) no object was close to the ego vehicle which is why visualization of this metric regarding transitions does not make sense. Nevertheless, this metric is represented for the analysis of incident candidates in Section 4.4.3.5.

4.4.3.3 Foreseeable misuse

<u>Distraction</u>

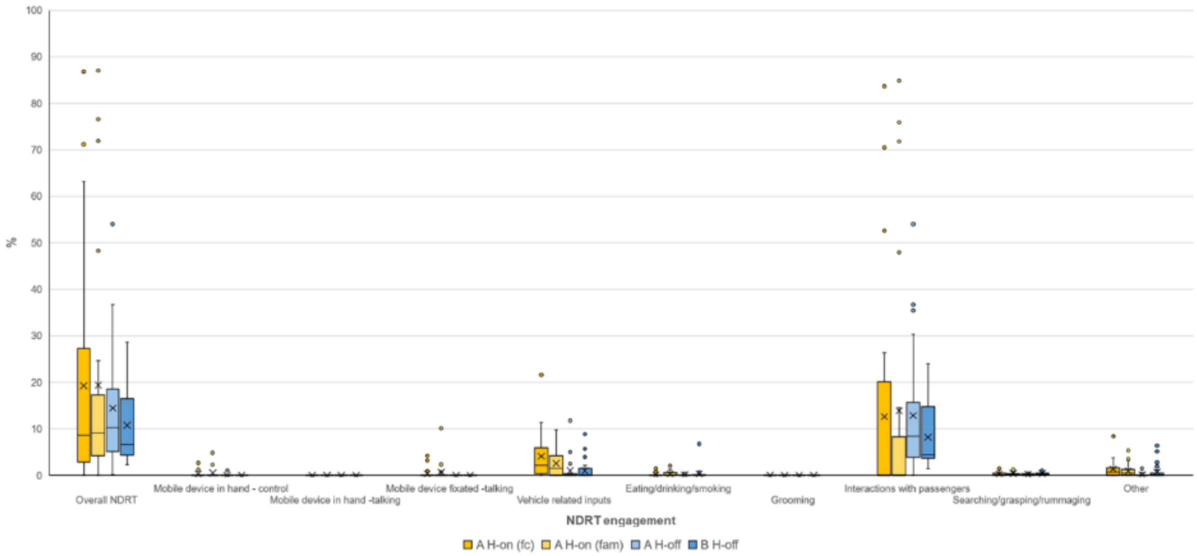


Figure 4-42: Visualization of the distribution of NDRT engagement for A H-on (fc) (*n* = 22), A H-on (fam) (*n* = 22), A H-off (fc) (*n* = 22) and B H-off (fc) (*n* = 28). The sample size indicates the number of participants

Figure 4-42 shows the distribution of NDRT engagement based on nine different NDRT categories adapted by Metz et al. (2014) and Pfleging et al. (2016). Overall, for both systems, the NDRT engagement is at a low level and no significant differences can be found. By far the two most frequent NDRTs are vehicle-related inputs/infotainment use and interaction with passengers (talking). L2H-on drivers tend to do/use more vehicle-related inputs/infotainment than L2H-off drivers. Instead, L2H-off users tend to interact/talk with passengers more frequently than L2H-on users. The use of mobile devices is really low. However, there is a slight tendency that L2H-on drivers to use their mobile phones more often than L2H-off drivers. The three categories of eating/drinking/smoking, grooming, and searching/grasping/rummaging only take place by a very small proportion, and differences between L2H-on subjects or L2H-off subjects exist. We can conclude that free hands do not necessarily lead or motivate to engage more frequently in NDRTs that are characterized by motoric actions in the way of using the hands. It should be pointed out that the NDRT engagement for L2H-off drivers could be influenced by the safety co-driver (study leader).

When taking a closer look at the NDRT engagement at transitions (see Figure 4-43), before all four transition types, the mean NDRT engagement is on a low and similar level for both systems except the system initiated deactivations where L2H-on subjects clearly show higher NDRT engagement than L2H-off subjects. However, in all four transition types, the dispersion is significantly higher for L2H-on participants than for L2H-off participants.

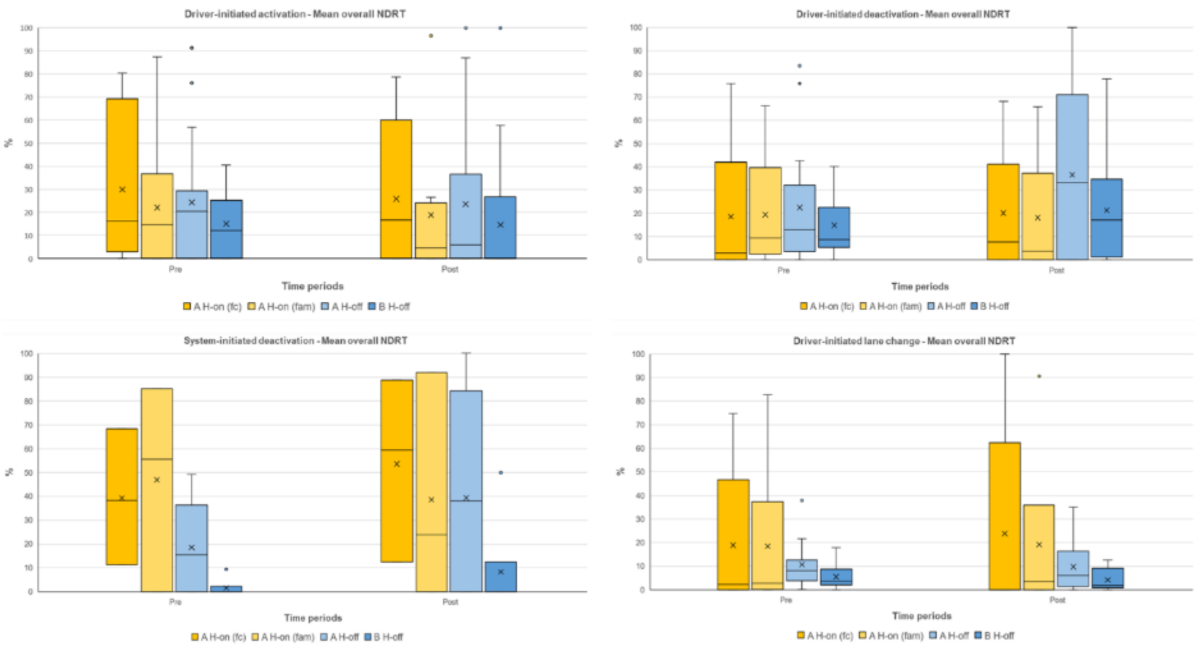


Figure 4-43: Visualization of the distribution of NDRT engagement at transitions for A H-on (fc) (*n* = 10(DA), 12(DD), 3(SD), 11(DL)), A H-on (fam) (*n* = 9, 12, 3, 6), A H-off (fc) (*n* = 18, 17, 8, 11) and B H-off (fc) (*n* = 23, 23, 6, 14). The sample size indicates the number of events

Figure 4-44 visualizes the subjective NDRT engagement differentiating between the three following types of NDRTs: visual and motoric (yellow), primary motoric and other modalities subordinated (green), and primary auditory and other modalities subordinated (red). The results of the subjective NDRT engagement coincide with objective NDRT engagement.

There are no differences between L2H-on subjects and L2H-off subjects. The most frequent NDRTs are the use of a fixed mobile device for talking, vehicle-related inputs/infotainment, eating/drinking/smoking, and interaction with a passenger. The four tasks are predominantly associated with less visual distraction. However, the use of a fixed mobile device for talking and eating/drinking/smoking stands in contrast to the objective NDRT engagement findings. Ultimately, we have to bear in mind that the NDRT engagement can only be evaluated on a short-term base and potential long-term effects cannot be assessed.

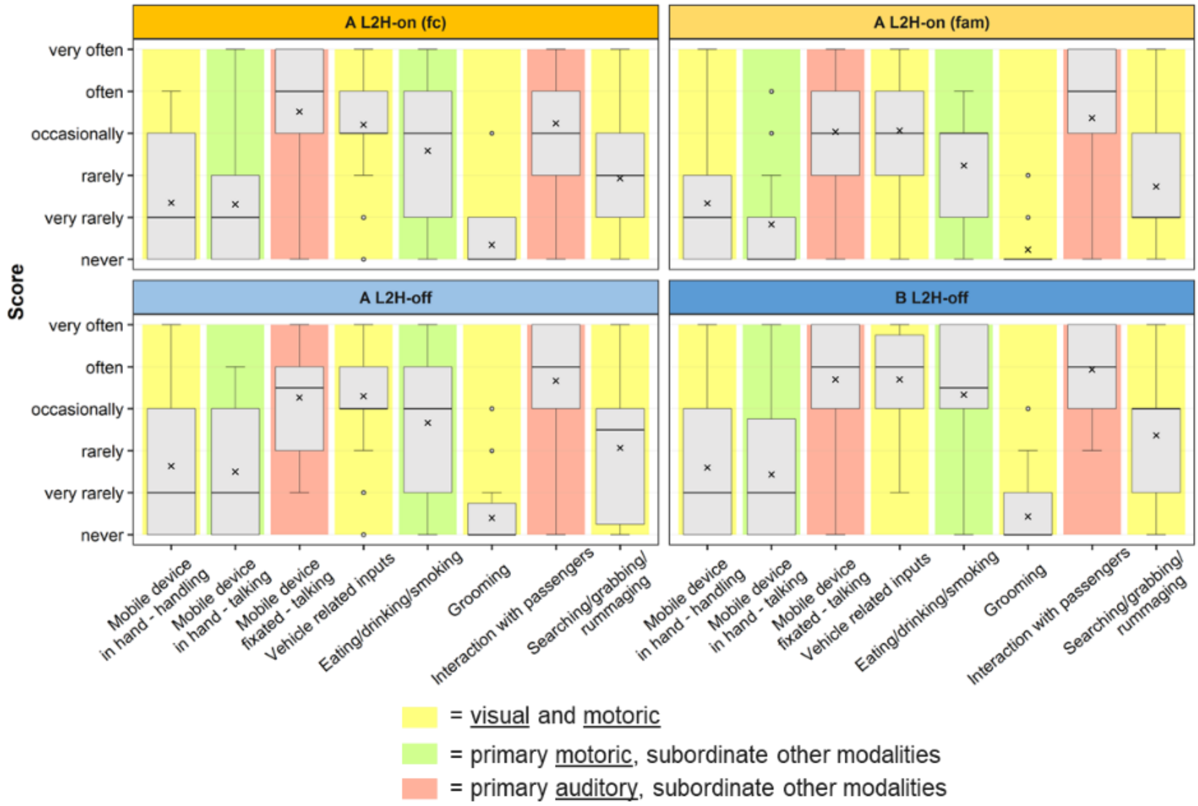


Figure 4-44: Visualization of the distribution of subjective NDRT engagement for A H-on (fc) (*n* = 29), A H-on (fam) (*n* = 30), A H-off (fc) (*n* = 30) and B H-off (fc) (*n* = 30). The sample size indicates the number of participants

<u>Disuse and misuse</u>

L2H-on drivers mostly have their hands on the steering wheel but some outliers with relatively high hands-off proportion exist indicating a potential for misuse (see Figure 4-45).

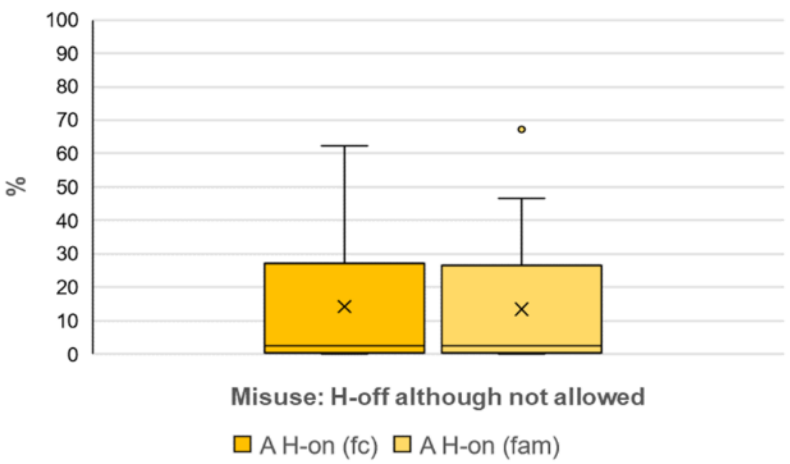


Figure 4-45: Visualization of the distribution of misuse in terms of hands-off in L2H-on vehicles for A H-on (fc) (*n* = 20), A H-on (fam) (*n* = 22). The sample size indicates the number of participants

Overall, trust is given for both systems but not strong and rather balanced which makes mind-off and misuse or disuse less likely (see Figure 4-46). No substantial differences exist between

L2H-on drivers and L2H-off drivers. However, the former experience with L2H-on leads to a slightly increased trust and propensity to trust in L2H-off rides. The understanding and predictability component shows a larger variance but for the most part, is very good which makes mode confusion less likely.

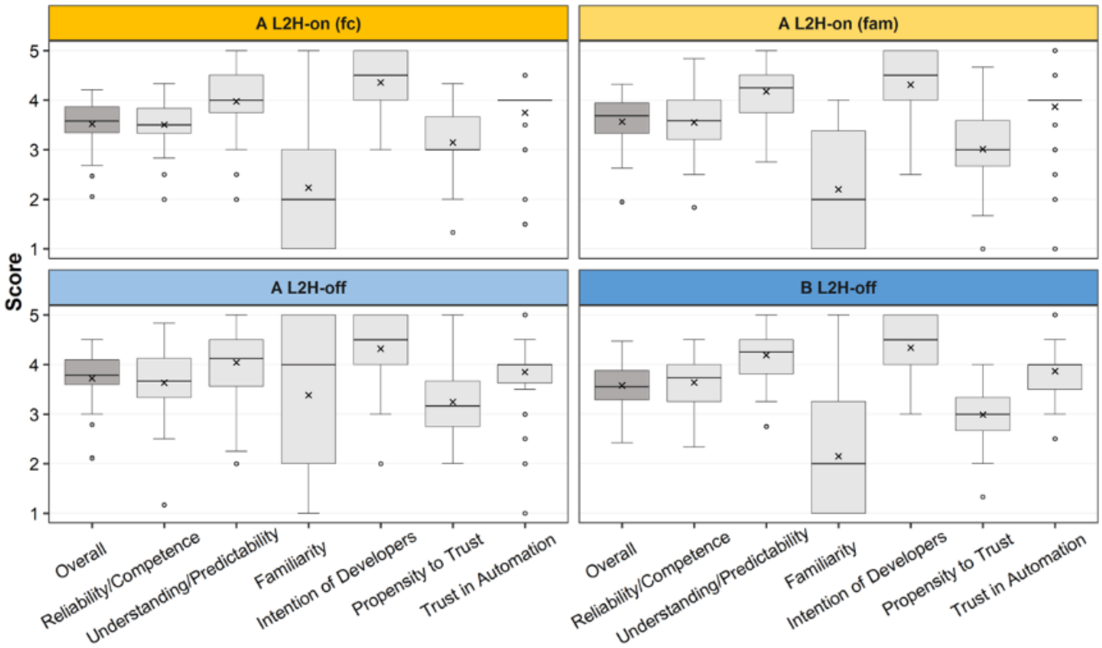


Figure 4-46: Visualization of the distribution of the trust rating for A H-on (fc) (n = 30), A H-on (fam) (n = 30), A H-off (fc) (n = 30) and B H-off (fc) (n = 30). The sample size indicates the number of participants

Acceptance (attitude, behavioral intention) is given and balanced but also large dispersion can be identified (see Figure 4-47).

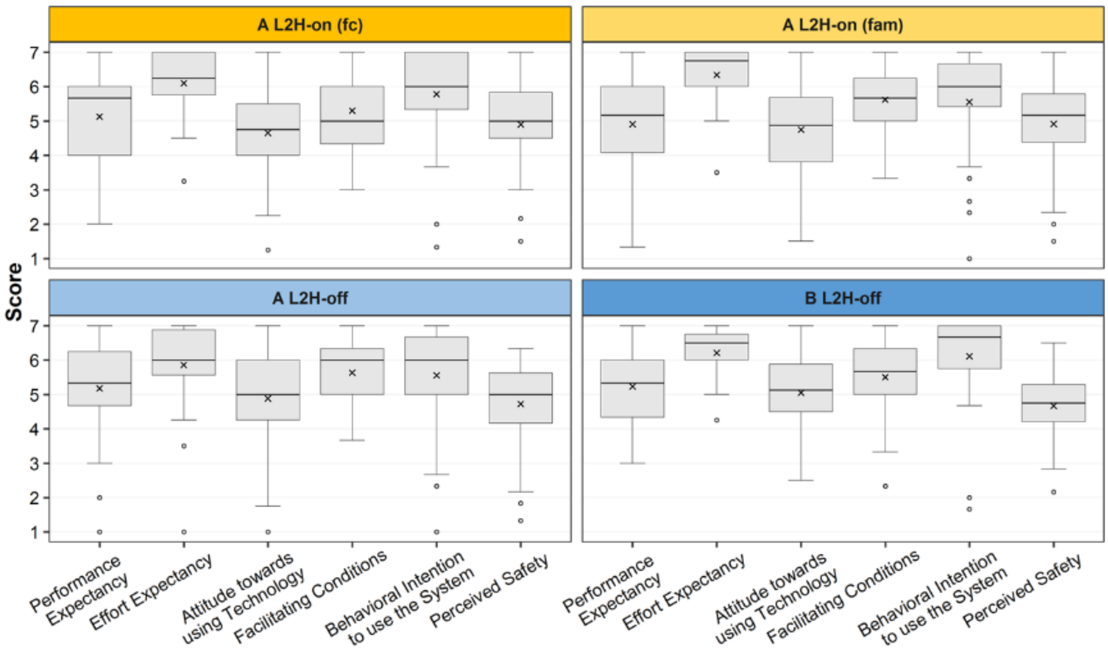


Figure 4-47: Visualization of the distribution of the acceptance rating for A H-on (fc) (n = 29), A H-on (fam) (n = 30), A H-off (fc) (n = 30) and B H-off (fc) (n = 30). The sample size indicates the number of participants

This makes misuse or disuse less likely. No clear differences exist between L2H-on drivers and L2H-off drivers.

Finally, the influence of the experimental setting was estimated by the participants. The results show that the experimental setting rather has less influence but a tendency for an increased influence by the experimental setting can be found for L2H-off rides. It can be argued that the subjects predominantly showed their natural driving behavior (short-term) but the safety co-driver during the L2H-off rides seems to have an impact. Therefore, the comparability between both systems is questionable for at least visual attention and NDRT engagement.

4.4.3.4 Mode confusion

Knowledge-based confusion

Figure 4-48 depicts the overall score of the system- and role understanding. Overall, the system and role understanding are good for both systems and no differences exist.

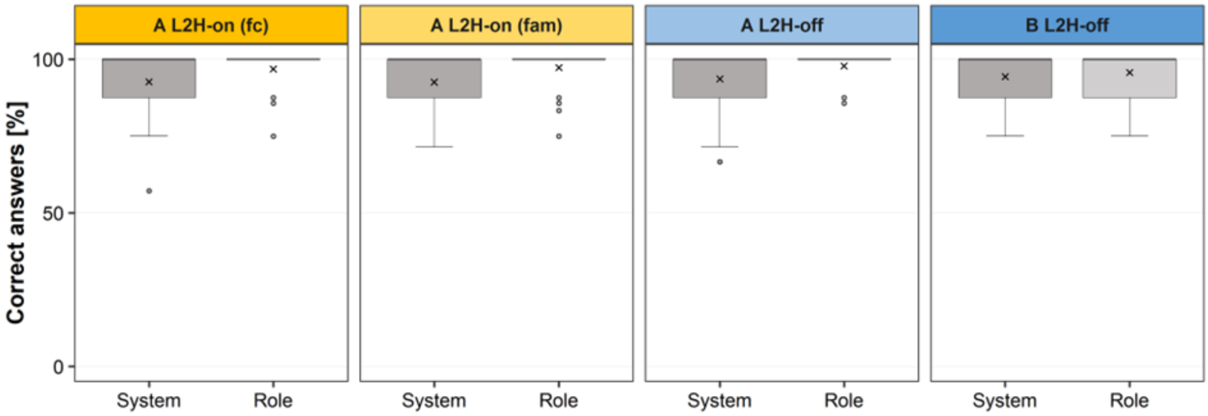


Figure 4-48: Visualization of the distribution of the overall score of the system- and role understanding for A H-on (fc) (*n* = 29), A H-on (fam) (*n* = 30), A H-off (fc) (*n* = 30) and B H-off (fc) (*n* = 30). The sample size indicates the number of participants

When considering the single items of the system and role understanding, then the following three minor issues can be identified (see Figure 4-49):

L2H-on and L2H-off:

- the system always detects if it is not able to handle a situation (system 05)
- permanent monitoring by the driver when the system is active (role 01)

L2H-on:

- the system automatically steers (system 06)

The three issues can be evaluated as minor issues because scores between 70% and 90% can be assessed as acceptable. The issue with „system 06“ only occurs for the first contact rides. So, there is a learning effect over time. Overall, the hands-on wheel requirement during

L2H-on use seems to lead to confusion about whether the driver even has to actively steer or not which is clear for L2H-off drivers.

Thus, the assumption is that the requirement to permanently take the hands on the steering wheel correlates with the driver rather less monitoring the system which provides supportive evidence for the findings regarding visual attention in monitoring that the hands-free option does not necessarily lead to mind-off.

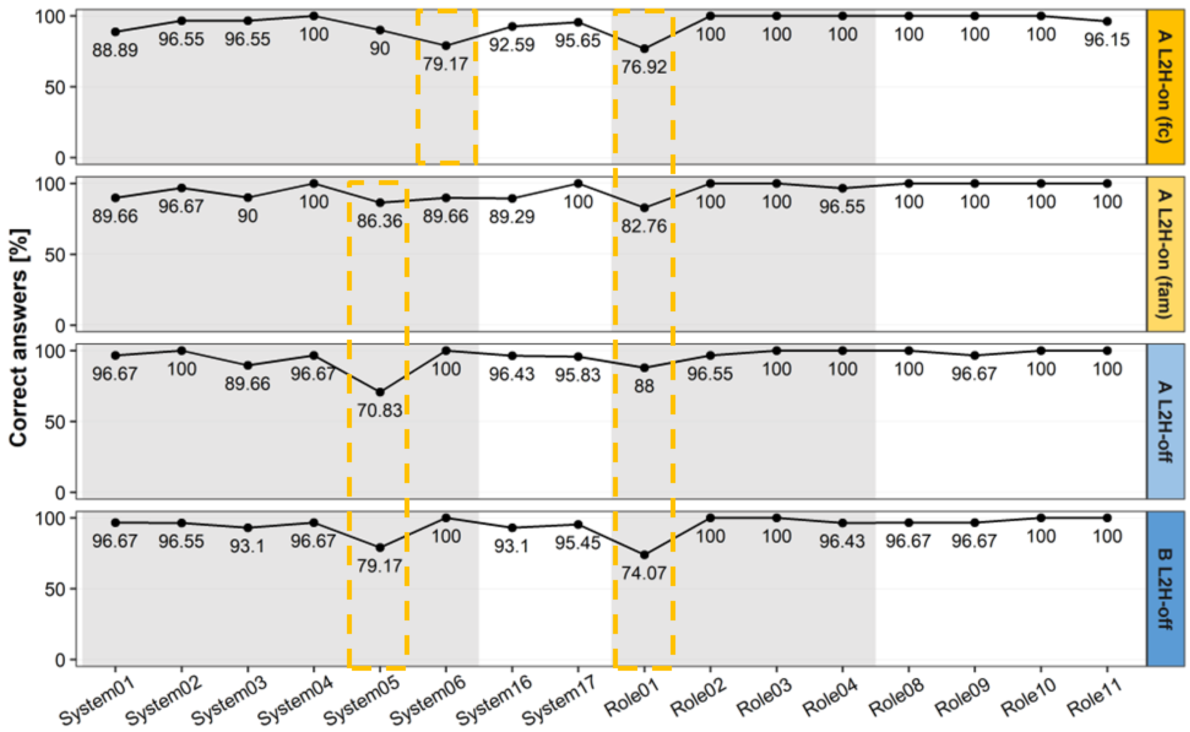


Figure 4-49: Visualization of the relative frequency of the single-item scores of the system- and role understanding for A H-on (fc) (n = 29), A H-on (fam) (n = 30), A H-off (fc) (n = 30) and B H-off (fc) (n = 30). The sample size indicates the number of participants

Behavior-based confusion

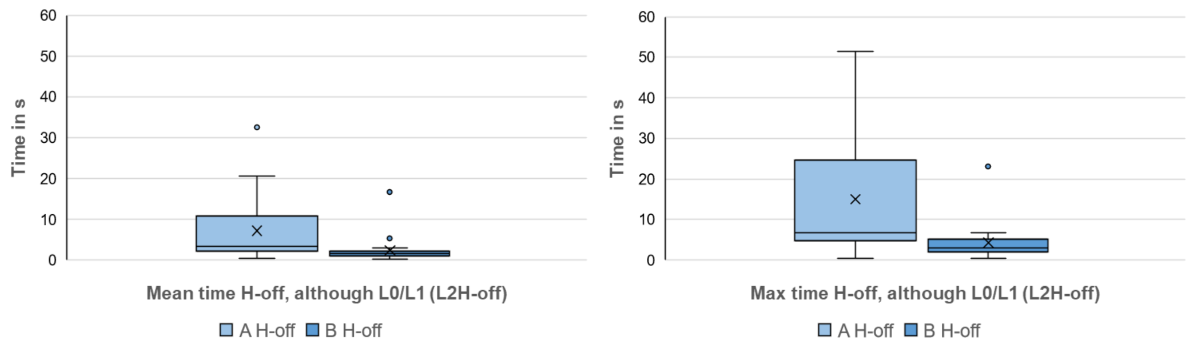


Figure 4-50: Visualization of the distribution of the mean and max times of hands-off in L2H-off vehicles although driving in L0/L1 mode for A H-off (fc) (n = 17) and B H-off (fc) (n = 20). The sample size indicates the number of participants

On average, L2H-off drivers show low times of hands-off the steering wheel in L0/L1 mode as well as in L2H-on mode (see Figure 4-50 and Figure 4-51). However, a few high outliers exist for both cases. Thus, a minor issue regarding behavior-based confusion in terms of mode switches as downgrades from L2H-off to L2H-on or L0/L1 exists for L2H-off participants. This is supported by participants' reports that switches between L2H-off and L2H-on must be clearly communicated and that changes from L2 to L0/L1 must be displayed by extensive and clear color changes.

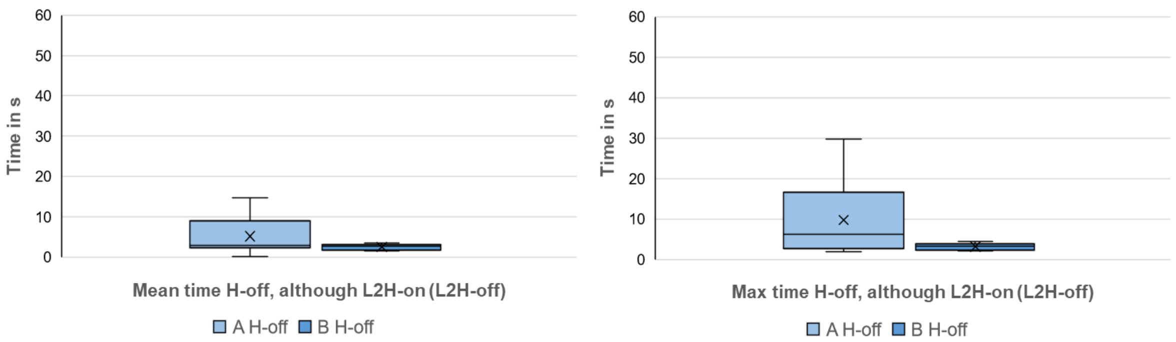


Figure 4-51: Visualization of the distribution of the mean and max times of hands-off in L2H-off vehicles although driving in L2H-on mode for A H-off (fc) (*n* = 11) and B H-off (fc) (*n* = 5). The sample size indicates the number of participants

The attempts to activate the L2 system although it is not available, are on a low level for both systems on average. However, L2H-off users tend to activate more often although the system is not available, also high outliers exist.

In total, the behavior-based results indicate that the L2H-off system is more complex than the L2H-on system but this complexity does not lead to over-proportional mode confusion.

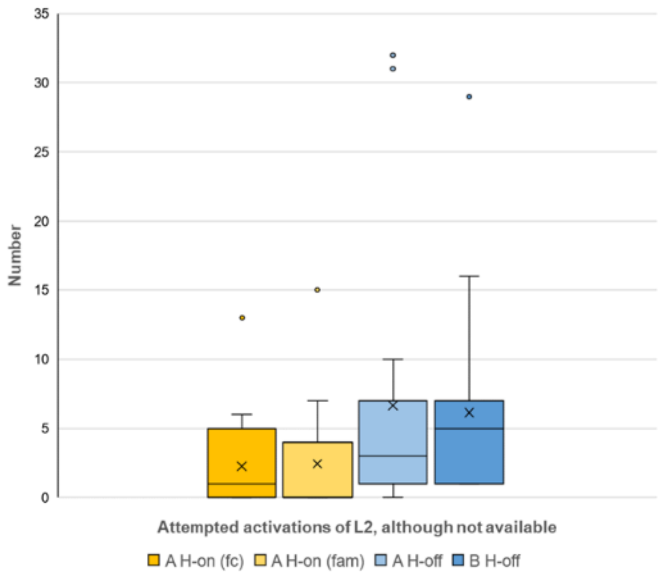


Figure 4-52: Visualization of the distribution of attempted activations of L2 although not available for A H-on (fc) (*n* = 15), A H-on (fam) (*n* = 16), A H-off (fc) (*n* = 15) and B H-off (fc) (*n* = 15). The sample size indicates the number of participants

4.4.3.5 Safety level

Objective safety (Safety-I)

Fortunately, no accidents happened. However, a few incidents are identified which will be presented in more detail below.

All transition types were predominantly uncritical in terms of criticality metrics (TTC, THW, longitudinal and lateral distance/acceleration) and no differences between L2H-on and L2H-off could be found (see Figure 4-62 – Figure 4-69). Just a few events can be identified as incident candidates which will be presented in more detail below. It should be pointed out that the sample size for L2H-on related events with regards to the TTC is small which is why both systems are difficult to compare in terms of TTC. In principle, the lateral distance should also be evaluated. However, in most transitions, no object was close to the ego vehicle which is why visualization of this metric regarding transitions does not make sense. Nevertheless, this metric is represented for the analysis of incident candidates below.

In total, we have 64 incident candidates for L2H-on subjects and 59 incident candidates for L2H-off subjects (see Figure 4-53). The sample size is comparable and thus, it could be said that there is no difference between both systems regarding the number of incident candidates.

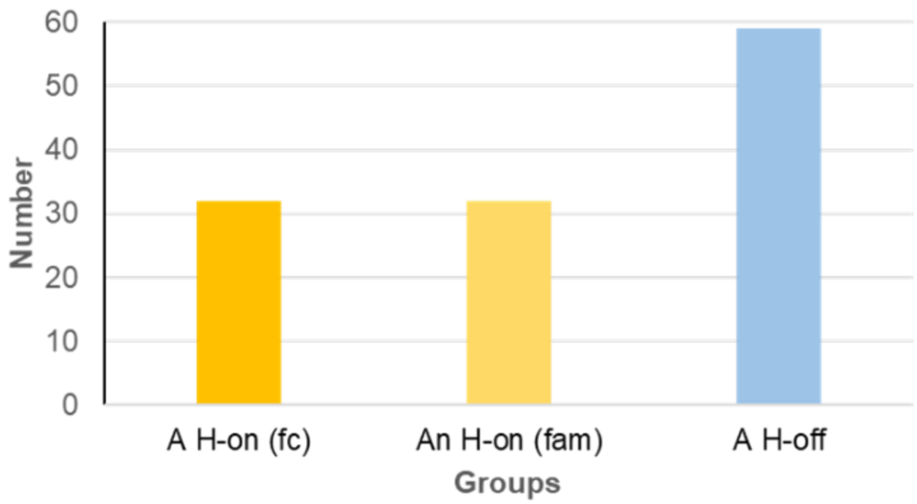


Figure 4-53: Visualization of the absolute frequency of incident candidates for A H-on (fc) (n = 9(DB), 29(VB)), A H-on (fam) (n = 19, 27), A H-off (fc) (n = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates

Most of the incident candidates are normal driving and thus not safety-critical (see Figure 4-54). Only 14 incidents exist as increased risk or crash-relevant whereas the proportion between L2H-on and L2H-off is balanced. Here, it is noticeable that vehicle dynamics-based longitudinal incidents predominate compared to distance-based incidents but vehicle dynamics-based lateral incidents did not occur at all (see Figure 4-55). The incidents at increased risk were mainly unpleasant (acceptable risk) in terms of controllability whereas one incident was perfect (see Figure 4-56). The incidents which are crash-relevant were mainly dangerous (non-acceptable risk) in terms of controllability whereas one incident was unpleasant (see Figure 4-56). No clear differences in controllability rating can be found between L2H-on participants and L2H-off participants. In particular, major issues for L2H-on are cut-in scenarios (especially in traffic jams) (see Figure 4-57 and Figure 4-58) and major issues for L2H-off are roadworks (lane detection, bumps) (see Figure 4-57 and Figure 4-58). Minor issues are a lane change in the traffic jam and to fall short of distance in roadwork for L2H-on drivers, passing on the right, unnecessary intervention or inappropriate braking by the driver, and overtrust during a lane change for L2H-of drivers.

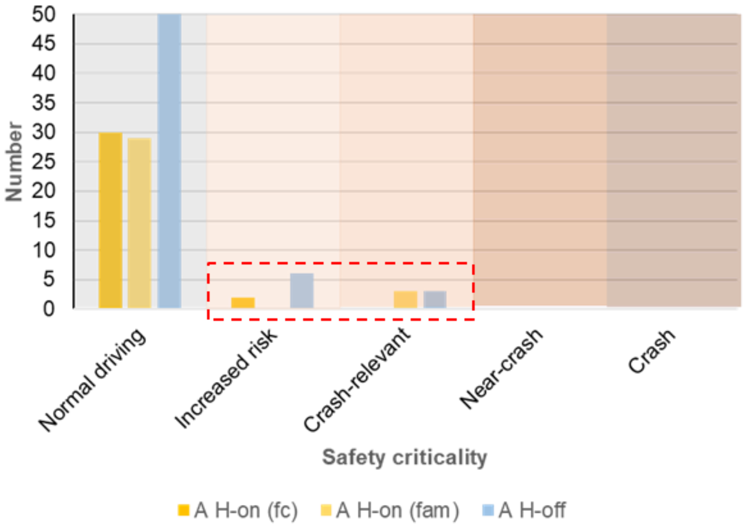


Figure 4-54: Visualization of the absolute frequency safety-criticality levels of the incident candidates for A H-on (fc) (*n* = 9(DB), 29(VB)), A H-on (fam) (*n* = 19, 27), A H-off (fc) (*n* = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates

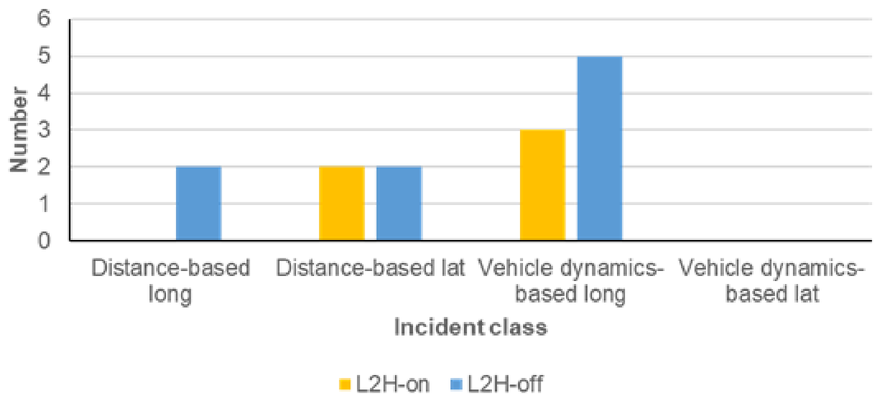


Figure 4-55: Visualization of the absolute frequency of incident classes of safety-critical incidents in terms of increased risk and crash-relevant for A H-on (fc) (*n* = 9(DB), 29(VB)), A H-on (fam) (*n* = 19, 27), A H-off (fc) (*n* = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates

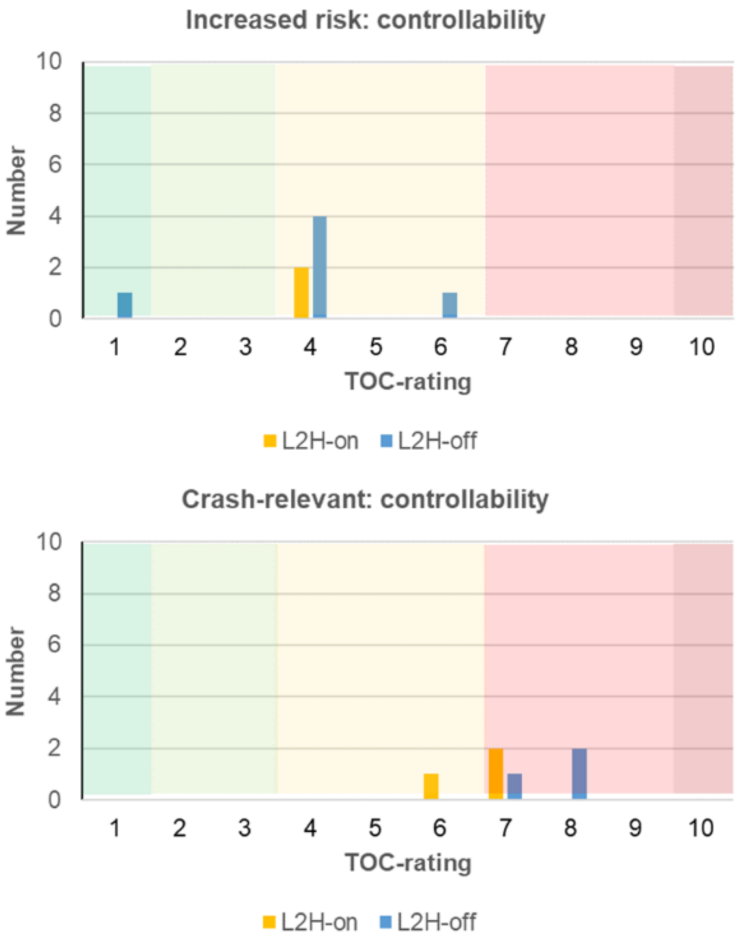
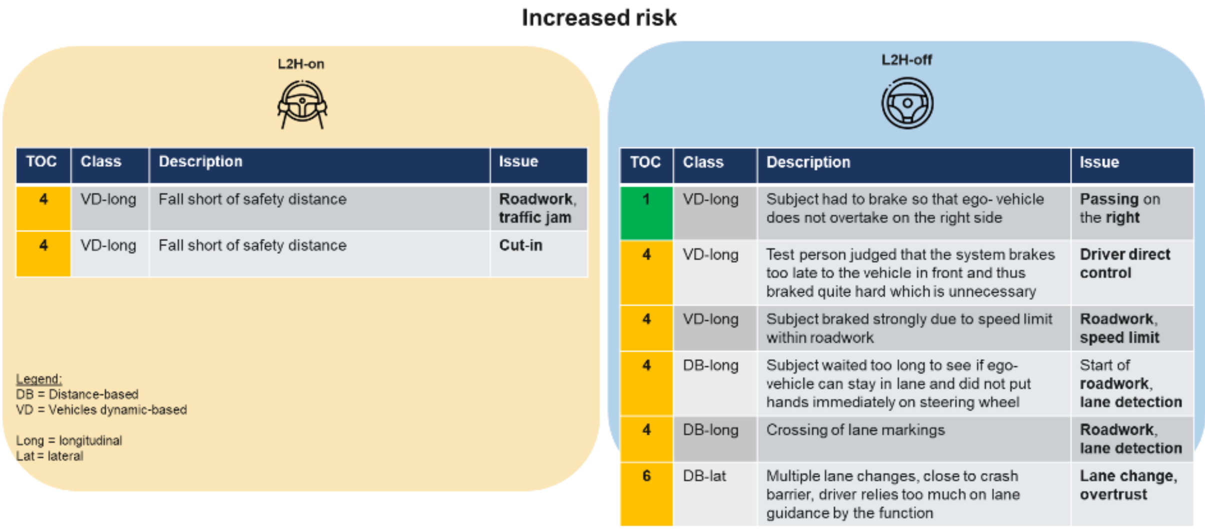


Figure 4-56: Visualization of the absolute frequency of TOC-ratings of safety-critical incidents in terms of increased risk and crash-relevant for A H-on (fc) (n = 9(DB), 29(VB)), A H-on (fam) (n = 19, 27), A H-off (fc) (n = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates



L2H-on

TOC	Class	Description	Issue
4	VD-long	Fall short of safety distance	**Roadwork**, **traffic jam**
4	VD-long	Fall short of safety distance	**Cut-in**

L2H-off

TOC	Class	Description	Issue
1	VD-long	Subject had to brake so that ego- vehicle does not overtake on the right side	**Passing** on the **right**
4	VD-long	Test person judged that the system brakes too late to the vehicle in front and thus braked quite hard which is unnecessary	**Driver direct control**
4	VD-long	Subject braked strongly due to speed limit within roadwork	**Roadwork**, **speed limit**
4	DB-long	Subject waited too long to see if ego-vehicle can stay in lane and did not put hands immediately on steering wheel	Start of **roadwork**, **lane detection**
4	DB-long	Crossing of lane markings	**Roadwork**, **lane detection**
6	DB-lat	Multiple lane changes, close to crash barrier, driver relies too much on lane guidance by the function	**Lane change**, **overtrust**

Figure 4-57: Visualization of the increased risk TOC-ratings and their incident class, description, and issue for A H-on (fc) (n = 9(DB), 29(VB)), A H-on (fam) (n = 19, 27), A H-off (fc) (n = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates (icon source: Flaticon.com)

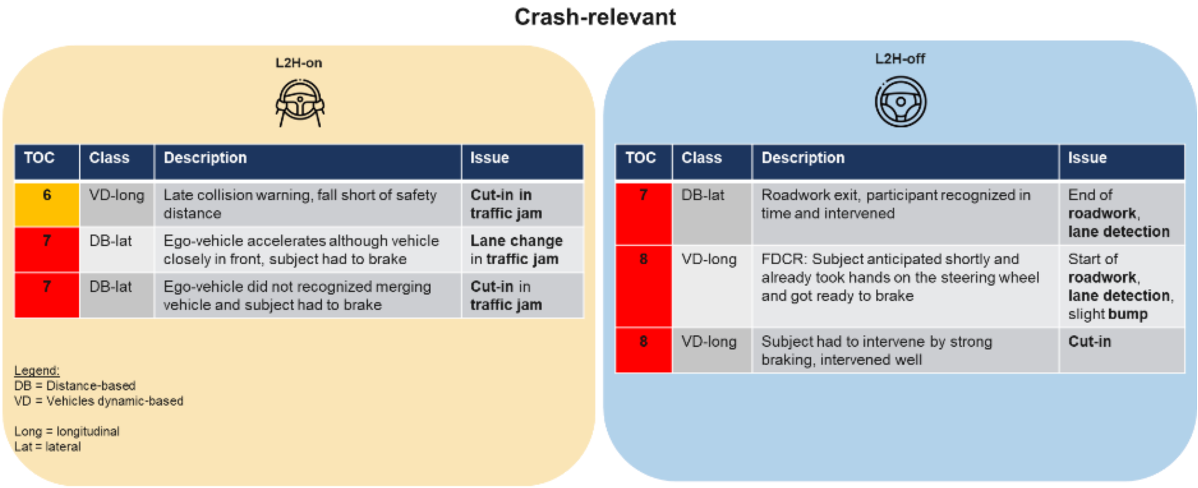


Figure 4-58: Visualization of the crash-relevant TOC-ratings and their incident class, description, and issue for A H-on (fc) (*n* = 9(DB), 29(VB)), A H-on (fam) (*n* = 19, 27), A H-off (fc) (*n* = 27, 30). The sample size indicates the number of participants in terms of distance-based (DB) and vehicle-dynamics-based (VB) incident candidates
(icon source: Flaticon.com)

Perceived safety

During L2H-off mode, L2H-off users have their hands on the steering wheel approximately 30% on average (Figure 4-59). However, a large spread exists.

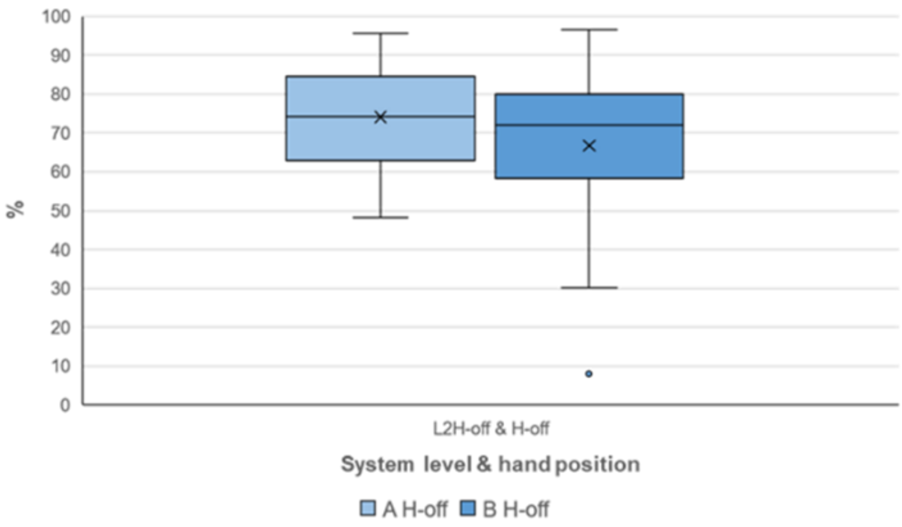


Figure 4-59: Visualization of the distribution of the hands-off ratio while driving in L2H-off mode for A H-off (fc) (*n* = 13) and B H-off (*n* = 14). The sample size indicates the number of participants

In terms of perceived safety, the overall trust score (see Figure 4-46) as well as the perceived safety component of the acceptance rating (see Figure 4-47) must be evaluated. We can observe a balanced trust and acceptance for both L2H-on subjects and L2H-off subjects indicating no over-/undertrust.

The participants rate the L2H-on & L2H-off systems with a good overall intention to use (see Figure 4-61). In particular, L2H-off was preferred over L2H-on (Figure 4-60) as it is more comfortable and participants had a clear desire to use L2H-off although a large dispersion can be observed (see Figure 4-61). Especially, in the FOT, participants reported that L2H-off drives smoother and more stable than L2H-on. Additionally, in rain and spray, L2H-off sometimes experienced problems/frequent system drops, which is annoying that in turn can lead to decreased perceived safety.

It can be argued that the formerly mentioned findings indicate a good and balanced perceived safety that strengthens the cognitive component of the driving task, which in turn as a precondition (information processing) promotes the ability to safely guide the vehicle.

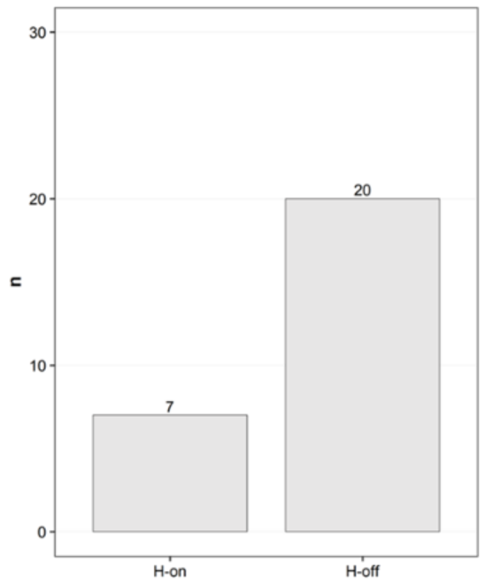


Figure 4-60: Visualization of the absolute frequency of L2H-on/-off system preference for A H-off (fc) (*n* = 27). The sample size indicates the number of participants

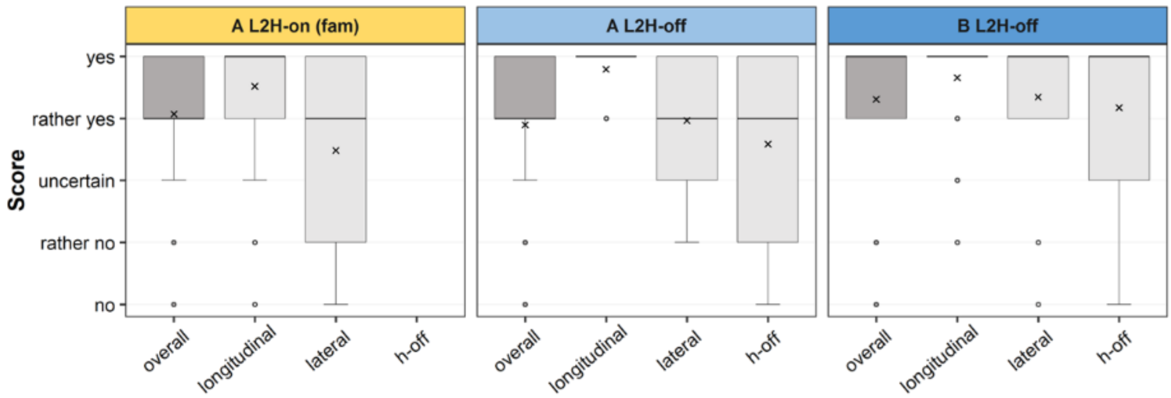


Figure 4-61: Visualization of the distribution of the overall and specific component scores of L2 intention to use for A H-on (fam) (*n* = 29), A H-off (fc) (*n* = 29) and B H-off (*n* = 29). The sample size indicates the number of participants

4.4.4 Discussion

In the following, the former mentioned specific results are discussed on a more abstract level to provide insights regarding the five CQs. Additionally, potential limitations are discussed.

4.4.4.1 Conclusions

<u>CQ1: Hands-off = mind-off?</u>

We have evidence that L2H-off does not lead to mind-off when using a proper DMS because visual attention is improved, the monitoring behavior is good, perceptual and motoric readiness is established to actively prepare for transitions (anticipation) as well as the sufficient motoric ability for safe vehicle guidance could be observed in general. The DMS clearly showed positive effects resulting in better visual attention and good monitoring behavior. In order to completely assess the proposed challenge of mind-off (CQ1), we refer to the conclusions regarding the cognitive component analysis in CQ 3 and 4.

<u>CQ2: Prolonged transition times</u>

The physical disadvantage of hands-free driving is compensated by taking the hands on the steering wheel and making a decision to act simultaneously (and not sequentially), transition times are not prolonged, and direct controls are successfully accomplished on average.

<u>CQ3: Foreseeable misuse</u>

Distraction and potential of misuse and disuse are low and even not more than L2H-on drivers. It has to be taken into account that foreseeable misuse is not evaluable in terms of long-term effects. In particular, the NDRT engagement for L2H-off drivers could be influenced by the safety co-driver.

<u>CQ4: Mode confusion</u>

Mode confusion is low and even not more than for L2H-on users. Overall, a good system and role understanding can be acknowledged. However, L2H-off is rated as more complex to use than L2H-on which can be seen by means of minor behavior-based confusion when it comes to mode switches from L2H-off to L2H-on and L0/L1.

<u>CQ5: Safety level</u>

No differences between L2H-on and L2H-off exist in terms of safety-critical outcomes. If considering the findings in CQ’s1-4, then it can be argued that a slight safety improvement is achieved by L2H-off mainly by the increased visual attention performance through the implemented DMS. With regards to the environment, some issues can be identified that are almost equal for both systems except for weather & lighting conditions. In terms of the infrastructure, the main issue is roadwork due to failed lane detection or bumps, and false speed limit detection at interchanges. The interaction with other road users is seldomly hazardous mainly in form of closely cutting-in vehicles and in the way that L2H-on and L2H-off rather impair the traffic flow in heavy traffic due to unnatural driving behavior (slow accelerating, abrupt braking). In rain and spray, many mode switches happened in the case of L2H-off use which increases the probability of disuse and mode confusion. This could not be observed for L2H-on.

4.4.4.2 Limitations

Ultimately, one have to bear in mind that the mentioned conclusions are only valid under the given driver, vehicle/system, and environmental characteristics. For example, the results are sensitive to the specification of the warning cascades for the HOD and DMS as well as the design of HMIs. The environmental characteristics are broad and rather realistic and naturalistic whereas the driver population is generally representative but has some limitations regarding the technology affinity, familiarity with L2 automation, and duration of use. This means that the results are valid in terms of drivers who are open-minded to new technologies, have no experience with L2, and use the system on a short-term base. In contrast, effects through older drivers tending to struggle with new technology, drivers with L2 experience, and long-term use cannot be assessed with this FOT. In addition, the comparison of safety effects between manual drivers and L2H-off drivers cannot be answered. Hence, we refer to findings in the literature in general or to findings in the simulator study 1 in Section 5.2 where manual drivers are included.

4.4.5 References

Benmimoun, M., Fahrenkrog, F., Zlocki, A., & Eckstein, L. (2011). Incident detection based on vehicle CAN-data within the large scale field operational test “euroFOT”. *22nd Enhanced Safety of Vehicles Conference* (ESV 2011).

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., et al. (2006). The 100-Car Naturalistic Driving Study Phase II – Results of the 100-Car Field Experiment. Report No. DOT HS 810 593. *Washington: National Highway Traffic Safety Admin*. (NHTSA)

Faber, F., Jonkers, E., Aust, M. L., Benmimoun, M., Regan, M., Jamson, S., & Dobberstein, J. (2011). Analysis methods for user related aspects and impact assessment on traffic safety, traffic efficiency and environment. *Deliverable 6.2 of European Large-Scale Field Operational Test on Active Safety Systems (euroFOT) project*. Retrieved from http://www.eurofotip.eu/en/library/deliverables/sp6_d62_analysis_methods_for_user_related_aspects_and_impact_assessment.htm.

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction*, *35*(6), 456-467.

Fleischer, M., & Chen, S. (2020, July). How Do We Sit When Our Car Drives for Us?. In *International Conference on Human-Computer Interaction* (pp. 33-49). Springer, Cham.

French, D. J., West, R. J., Elander, J., & Wilding, J. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics*, *36*(6), 627-644.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in ecology and evolution*, *7*(11), 1325-1330.

Götz, M. (2007). *Die Gestaltung von Bedienelementen unter dem Aspekt ihrer kommunikativen Funktion* (Dissertation, Technische Universität München).

Hickman, J. S., Hanowski, R. J., & Bocanegra, J. (2010). Distraction in commercial trucks and buses: Assessing prevalence and risk in conjunction with crashes and near-crashes. Report No. FMCSA-RRR-10-049. Washington D.C*.: Federal Motor Carrier Safety Administration*. Retrieved from http://www.fmcsa.dot.gov/facts-research/researchtechnology/report/Distraction-in-Commercial-Trucks-and-Buses-report.pdf.

Jonsson, B. (2011). Hand position on steering wheel during driving. *Traffic injury prevention*, 12(2), 187-190.

Körber, M. (2019). Theoretical considerations and development of a questionnaire to measure trust in automation. *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (1st ed., pp.13–30). Springer.

Transportation Research Board (2000). *Highway capacity manual*. Washington, DC, 2(1)

Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., et al. (2019). Deliverable 3.3: Evaluation methods. *L3Pilot Driving automation consortium*. Retrieved from https://l3pilot.eu/fileadmin/user_upload/Downloads/Deliverables/Update_07102021/L3Pilot-SP3-D3.3_Evaluation_Methods-v1.0_for_website.pdf

Metz, B., Landau, A., & Just, M. (2014). Frequency of secondary tasks in driving–Results from naturalistic driving data. *Safety science*, 68, 195-203.

Neukum, A., Lübbeke, T., Krüger, H. P., Mayser, C., & Steinle, J. (2008). ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In *5. Workshop Fahrerassistenzsysteme-FAS*(pp. 141-150). fmrt Karlsruhe.

Naujoks, F., Wiedemann, K., Schömig, N., Jarosch, O., & Gold, C. (2018). Expert-based controllability assessment of control transitions from automated to manual driving. *MethodsX*, *5*, 579-592.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations, final report.Report No. FMCSA-RRR-09-042. Washington, D.C.: *Federal Motor Carrier Safety Administration*. Retrieved from http://www.fmcsa.dot.gov/facts-research/research-technology/report/FMCSA-RRR-09-042.pdf

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012, October). Predicting information technology usage in the car: towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 51-58).

Pfleging, B., Rang, M., & Broy, N. (2016, December). Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th international conference on mobile and ubiquitous multimedia* (pp. 91-99).

Verordnung über die Zulassung von Personen zum Straßenverkehr (Fahrerlaubnis-Verordnung - FeV), § 48a et seq. (2010). https://www.gesetze-im-internet.de/fev_2010/BJNR198000010.html#BJNR198000010BJNG001200000

Risser, R., & Brandstätter, C. (1985). Die Wiener Fahrprobe. Freie Beobachtung. *KL FACHBUCHR KFV*, *21*.

Walton, D., & Thomas, J. A. (2005). Naturalistic observations of driver hand positions. *Transportation research part F: traffic psychology and behaviour*, 8(3), 229-238.

4.4.6 Appendix

4.4.6.1 Appendix A – Criticality metrics at transitions and TOC-rating combined with level of service

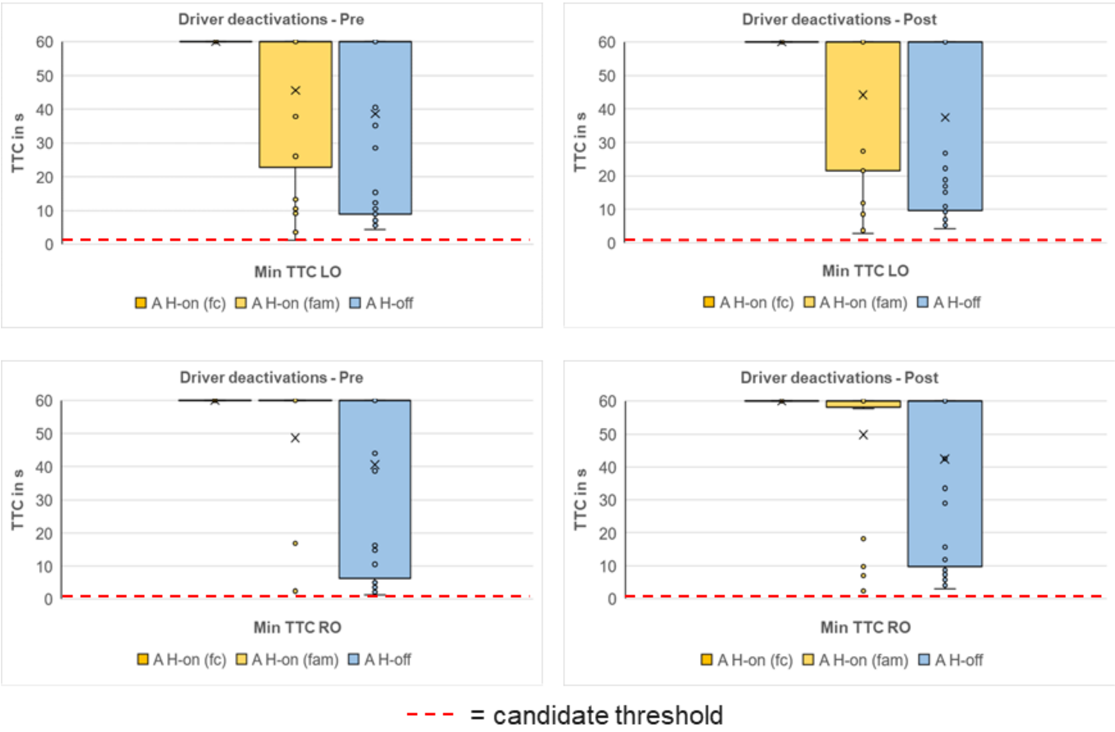


Figure 4-62: Visualization of the distribution of the minimum TTC (LO, RO) at driver-initiated deactivations divided into pre and post stage for A H-on (fc) (*n* = 6), A H-on (fam) (*n* = 22), A H-off (fc) (*n* = 47). The sample size indicates the number of events. The candidate threshold (TTC = 1.75s) defines the threshold for incident candidates

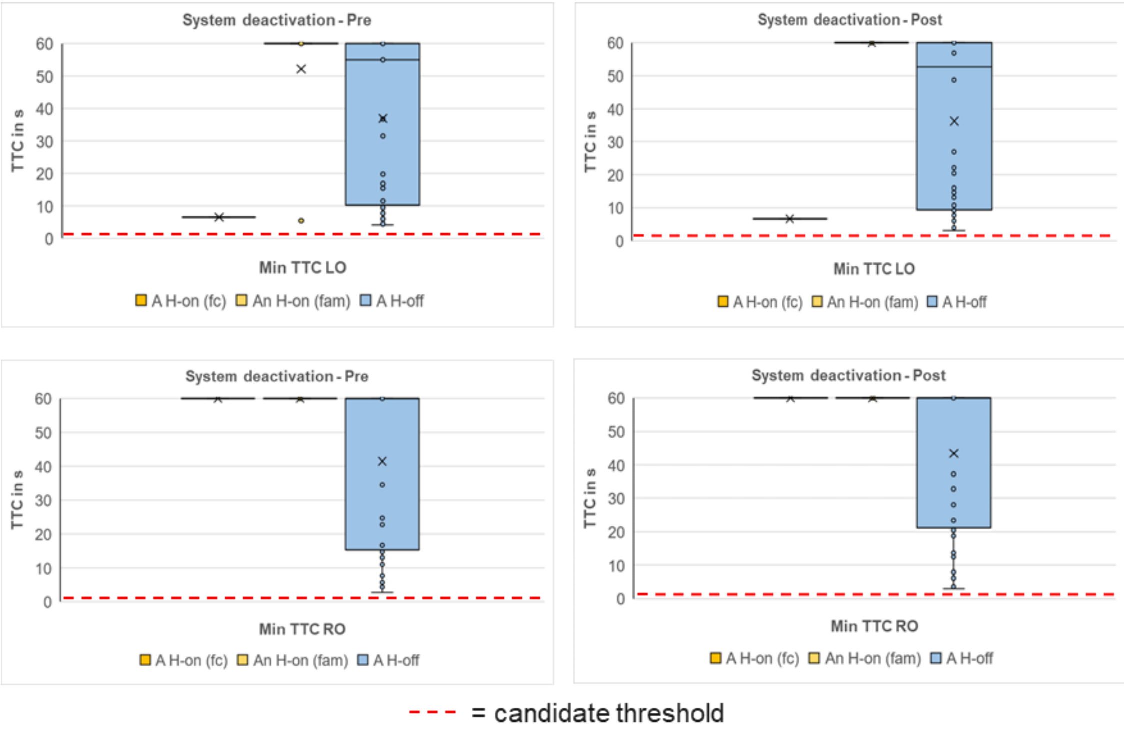


Figure 4-63: Visualization of the distribution of the minimum TTC (LO, RO) at system-initiated deactivations divided into pre and post stage for A H-on (fc) (*n* = 1), A H-on (fam) (*n* = 7), A H-off (fc) (*n* = 45). The sample size indicates the number of events. The candidate threshold (TTC = 1.75s) defines the threshold for incident candidates

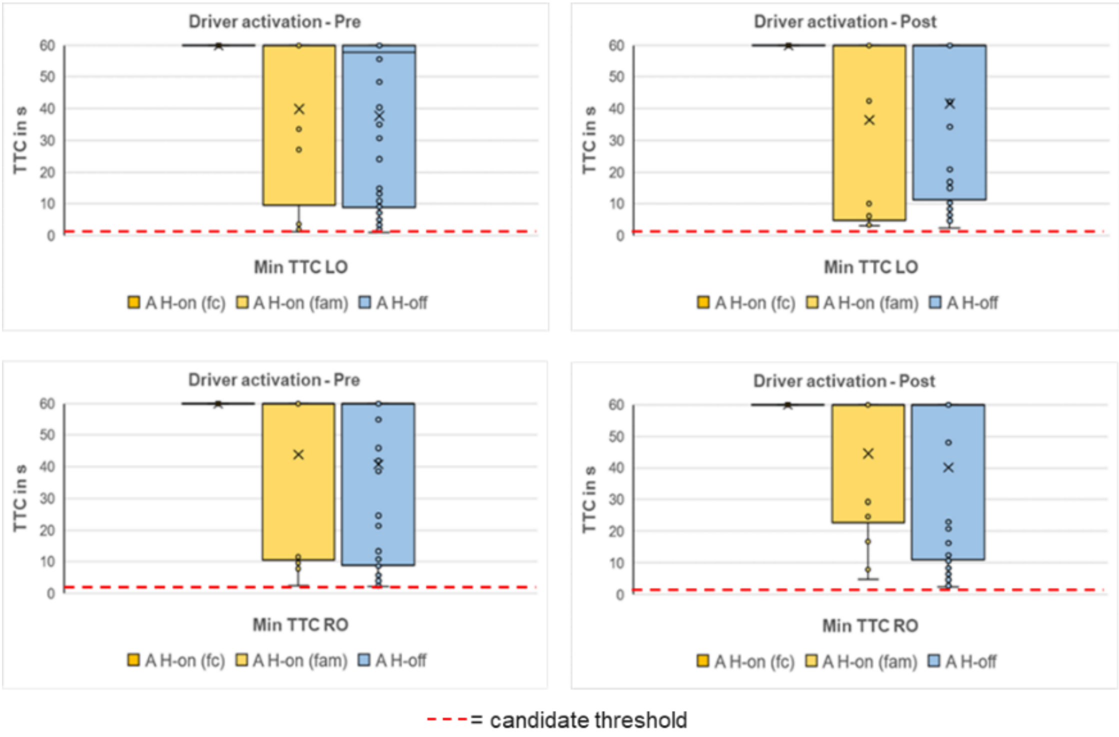


Figure 4-64: Visualization of the distribution of the minimum TTC (LO, RO) at driver-initiated activations divided into pre and post stage for A H-on (fc) (n = 9), A H-on (fam) (n = 16), A H-off (fc) (n = 58). The sample size indicates the number of events. The candidate threshold (TTC = 1.75s) defines the threshold for incident candidates

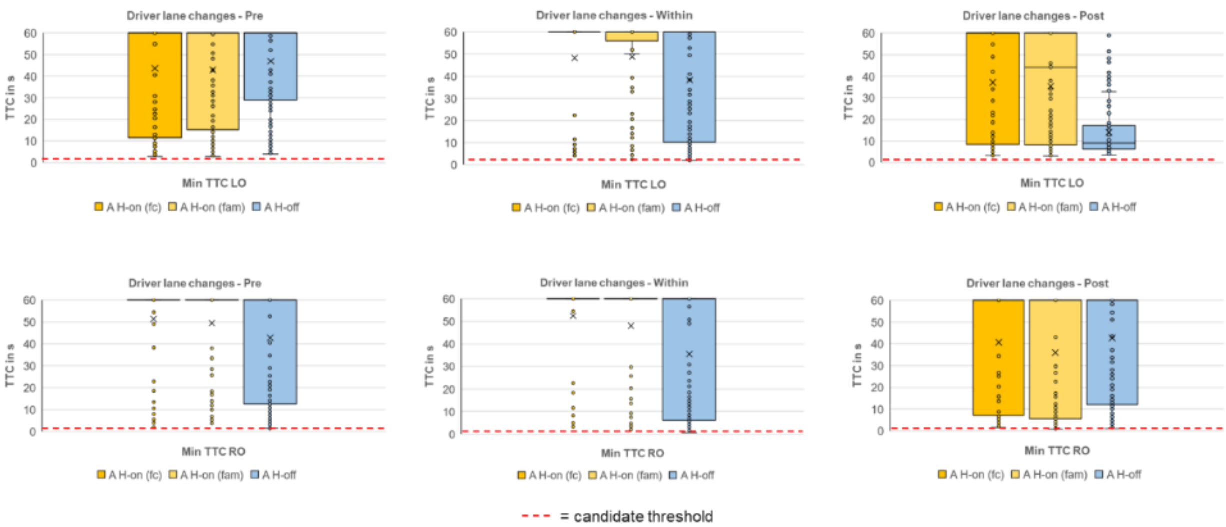


Figure 4-65: Visualization of the distribution of the minimum TTC (LO, RO) at driver-initiated deactivations divided into pre and post stage for A H-on (fc) (n = 120), A H-on (fam) (n = 188), A H-off (fc) (n = 208). The sample size indicates the number of events. The candidate threshold (TTC = 1.75s) defines the threshold for incident candidates.

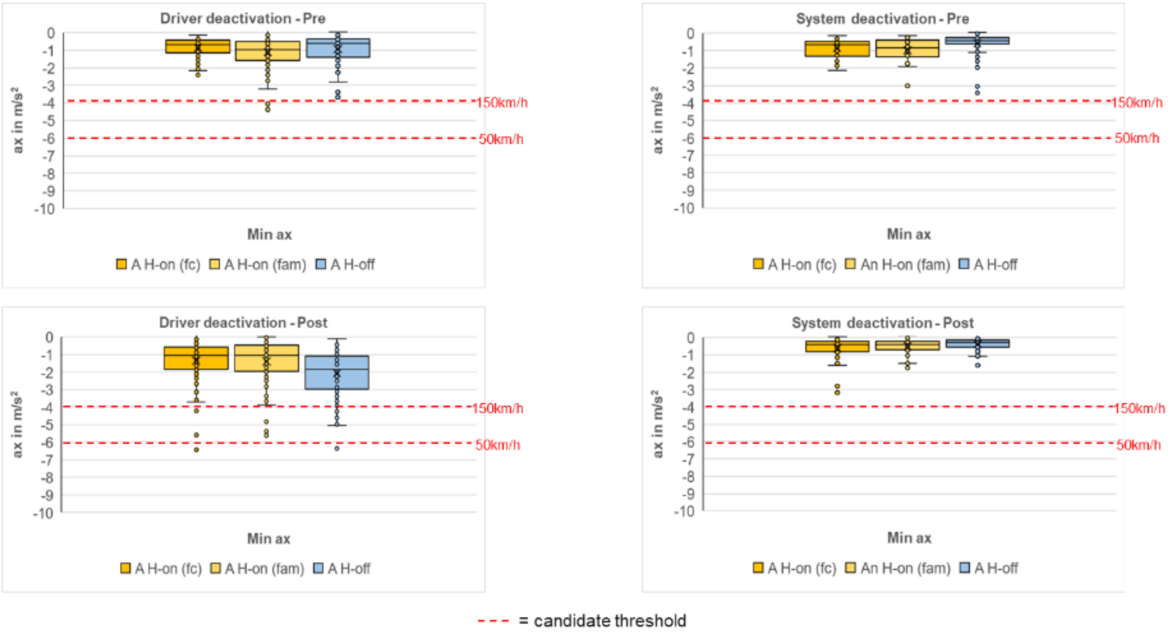


Figure 4-66: Visualization of the distribution of the minimum longitudinal acceleration at driver-/system-initiated deactivations divided into pre and post-stage for A H-on (fc) (*n* = 110(DD), 58(SD)), A H-on (fam) (*n* = 97, 46), A H-off (fc) (*n* = 58, 62). The sample size indicates the number of events. The candidate thresholds (ax = -4/-6m/s²) define the thresholds for incident candidates in dependency on the speed (50/150km/h)

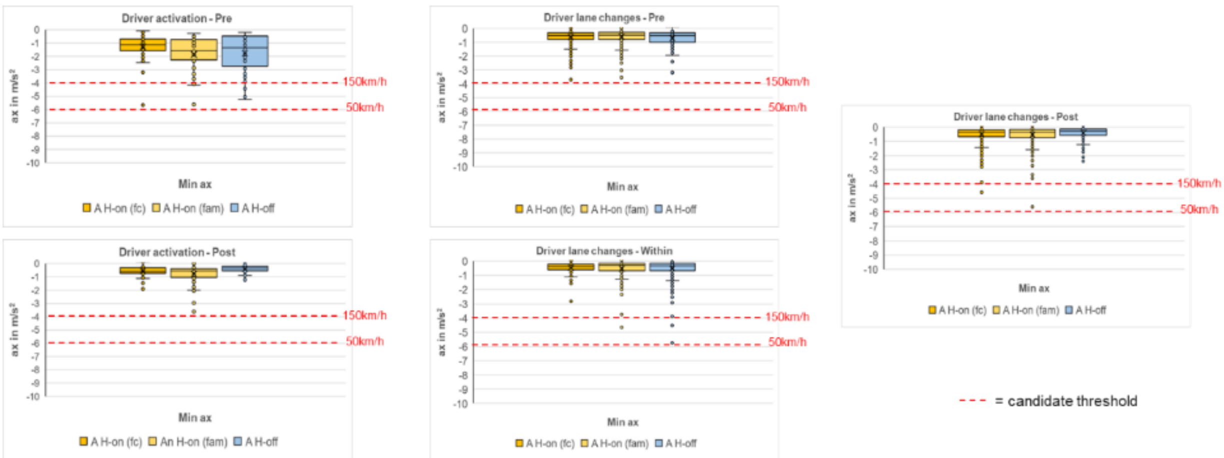


Figure 4-67: Visualization of the distribution of the minimum longitudinal acceleration at driver-initiated activations/lane changes divided into pre and post-stage for A H-on (fc) (*n* = 47(DA), 624(DL)), A H-on (fam) (*n* = 53, 668), A H-off (fc) (*n* = 60, 281). The sample size indicates the number of events. The candidate thresholds (ax = -4/-6m/s²) define the thresholds for incident candidates in dependency on the speed (50/150km/h)

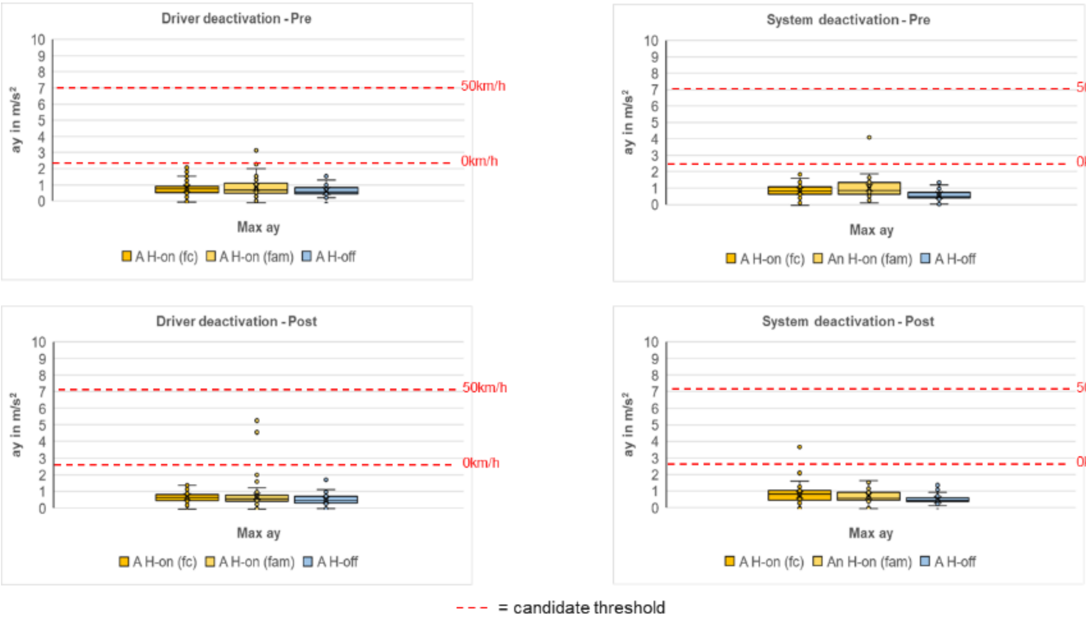


Figure 4-68: Visualization of the distribution of the maximum lateral acceleration at driver-/system-initiated deactivations divided into pre and post-stage for A H-on (fc) (n = 110(DD), 58(SD)), A H-on (fam) (n = 97, 46), A H-off (fc) (n = 58, 62). The sample size indicates the number of events. The candidate thresholds (ay = 2.5/7m/s^2) define the thresholds for incident candidates in dependency on the speed (0/50km/h)

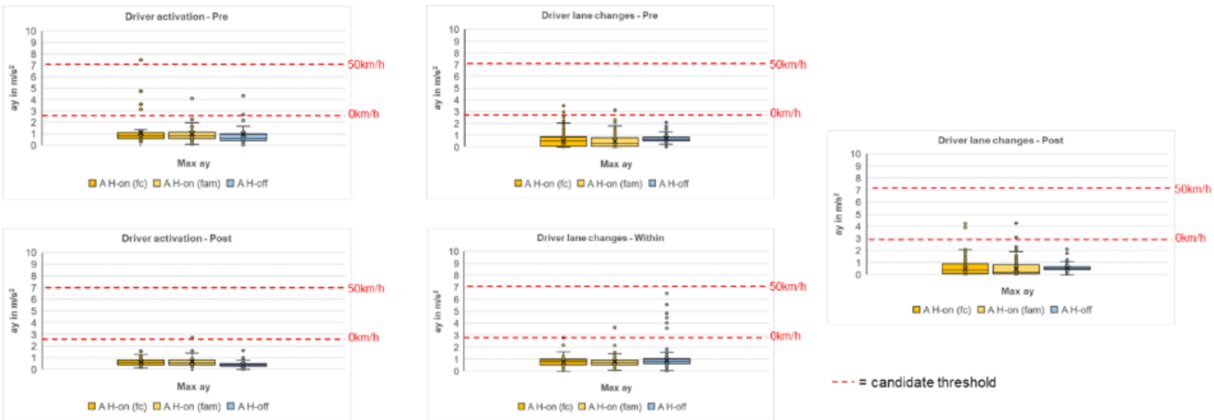


Figure 4-69: Visualization of the distribution of the maximum lateral acceleration at driver-initiated activations/lane changes divided into pre and post stage for A H-on (fc) (n = 47(DA), 624(DL)), A H-on (fam) (n = 53, 668), A H-off (fc) (n = 60, 281). The sample size indicates the number of events. The candidate thresholds (ay = 2.5/7m/s^2) define the thresholds for incident candidates in dependency on the speed (0/50km/h)

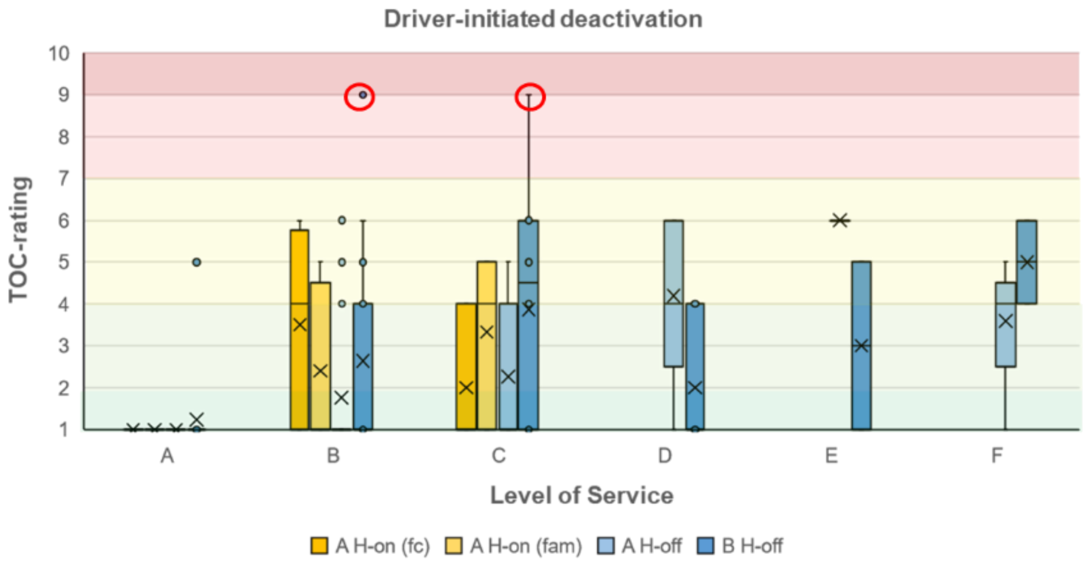


Figure 4-70: Visualization of the distribution of the TOC-rating in case of driver-initiated deactivations for A H-on (fc) (n = 12), A H-on (fam) (n = 10), A H-off (fc) (n = 61) and B H-off (fc) (n = 119). The sample size indicates the number of events

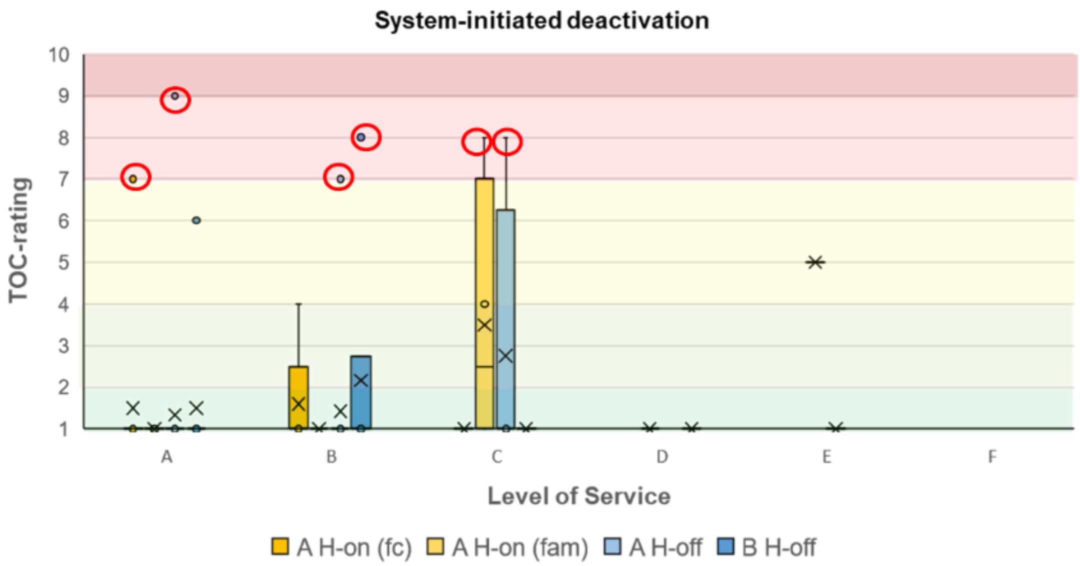


Figure 4-71: Visualization of the distribution of the TOC-rating in case of driver-initiated deactivations for A H-on (fc) (n = 18), A H-on (fam) (n = 20), A H-off (fc) (n = 43) and B H-off (fc) (n = 18). The sample size indicates the number of events

4.4.6.2 Appendix B – Questionnaires and interview guidance material

Pre-Questionnaire

[all groups]

Beschreibung	**Frage**	**Antwortformat**
Datenzuordnung		
Probanden-code	Bitte generieren Sie Ihren persönlichen Versuchs-personen-Code für die Studie. Dieser Code besitzt den Vorteil, dass Sie den Code mittels der Fragen jederzeit neu generieren können, außenstehende Dritte jedoch kaum. Wir benötigen diesen Code, um Ihre Daten der Vorbefragung mit den Daten der Versuchsfahrt zu verknüpfen.	[…]
Name	Bitte geben Sie Ihre Kontaktdaten an. Diese Daten dienen ausschließlich der Kontaktauf-nahme nach Zuordnung zu einer Versuchsgruppe (A oder B). Die Daten werden getrennt von den weiteren im Fragebogen erhobenen Daten aufbe-wahrt und mit Abschluss der Datenerhebung ge-löscht.	[…]
Organisatorische Fragen		
Gruppe	Bitte geben Sie an, in welcher/n Gruppe/n Sie teil-nehmen möchten. Eine Übersicht über die Tätigkeiten innerhalb der Gruppe ist im Folgenden dargestellt.	• Nur A • Nur B • Egal • Lieber A, aber B wäre auch in Ordnung • Lieber B, aber A wäre auch in Ordnung
Parkoptionen	Bitte geben Sie an, ob Sie eines der Versuchsfahr-zeuge für den Zeitraum der Überlassung (ca. 5 Tage) den Abstellregeln entsprechend unterstellen können.	• Nein • Eher nein, individuelle Klärung mit Stu-dienteam gewünscht • Eher ja, individuelle Klärung mit Stu-dienteam gewünscht • Ja
ParkenMasse	Hier sind die voraussichtlichen Maße der Versuchs-fahrzeuge inkl. Messequipment angegeben. Bitte kreuzen Sie an, welche(s) der Autos Sie un-terstellen können, also welches in Ihre Garage, durch Ihr Tor etc. passt. • Länge: 4,4m x Breite: 1,9m x Höhe: 1,8m • Länge: 4,5m x Breite: 1,9m x Höhe: 1,7m	Multiple Choice

	• Länge: 4,9m x Breite: 1,9m x Höhe: 1,9m • Länge: 5m x Breite: 2m x Höhe: 2,1m • Keines der genannten • Unsicher, individuelle Klärung mit Studienteam gewünscht	
Lademoeglichkeiten	Eines der potenziellen Versuchsfahrzeuge ist ein reines Elektrofahrzeug. Bitte geben Sie an, ob Sie eine oder mehrere Möglichkeiten haben, um ein solches Fahrzeug zu laden. • Laden auf Privatgelände, z.B. eigene Ladebox in der Garage • Laden im öffentlichen Raum, z.B. öffentliche (Schnell-)Ladesäulen. ACHTUNG: das Fahrzeug darf dort aufgrund des Messaufbaus nicht länger als 1h ohne Sichtkontakt abgestellt werden. • Keine Lademöglichkeit vorhanden (-> kein Elektrofahrzeug möglich) • Sonstiges	Multiple Choice
ZeitraumAllg	Haben Sie innerhalb des Zeitraums Februar bis Juni 2022 bestimmte Wochen, die Ihnen für die Studie besonders gut passen bzw. Wochen, die Sie bereits ausschließen können?	• Ja • Nein
ZeitraumFilterJa	Hier sind die Kalenderwochen des Versuchszeitraumes angegeben. Bitte geben Sie Ihre Verfügbarkeiten an. • KW6 • … • KW22	• Woche wird präferiert • Woche könnte generell passen • Woche ist ausgeschlossen
ZeitraumSonst	Haben Sie sonst allgemeine Einschränkungen Ihrer zeitlichen Verfügbarkeit wie z.B. „nie vormittags“ oder „nie freitags“. Geben Sie diese bitte stichpunktartig an.	[…]
Kommentare	Haben Sie zum aktuellen Zeitpunkt Wünsche, Fragen oder Kommentare?	[…]
Soziodemographisch		
Alter	Bitte geben Sie Ihr Alter in Jahren an.	[…]
Geschlecht	Bitte geben Sie Ihr Geschlecht an.	• Männlich • Weiblich • Divers • k.A.
Haendigkeit	Welche Hand bevorzugen Sie bei alltäglichen Verrichtungen (z.B. eine Schere benutzen)?	• Rechts • Links • Kein Unterschied

Sehschwaec he	Benutzen Sie beim Autofahren eine Sehhilfe?	• Ja, ich benutze Sie auch während des Versuchs (Brille/Kontaktlinsen). • Ja, ich benutze Sie jedoch nicht während des Versuchs. • Nein
Far-bfehlsichtigke it	Liegt bei Ihnen eine Farbfehlsichtigkeit vor?	• Ja, rot-grün Sehschwäche • Ja, blau-gelb Sehschwäche • Nein
Ho-erschwaeche	Liegt bei Ihnen eine Hörschwäche vor? Wenn ja, ist diese korrigiert?	• Ja, sie wird auch während des Versuchs korrigiert. • Ja, sie wird während des Versuchs jedoch nicht korrigiert. • Nein
Kontext Fahren		
Fueh-rerschein	In welchem Jahr haben Sie Ihren Pkw-Führerschein gemacht?	[…]
Fahrtfre-quenz	Wie oft sind Sie in den letzten zwölf Monaten im Durc	• Täglich • Mehrmals pro Woche • Mehrmals pro Monat • Weniger als einmal pro Monat • (Selten bis) Nie
Fahrtstrecke	Wie viele Kilometer sind Sie in den letzten zwölf Monaten circa mit dem Auto gefahren?	• 0 km (keine Fahrt) • 1 km - 5.000 km • 5.001 km - 20.000 km • 20.001 km - 50.000 km • 50.000 km - 100.000 km • Mehr als 100.000 km
FrequenzAu-tobahn	Wie oft sind Sie in den letzten zwölf Monaten im Durchschnitt auf Autobahnen Auto gefahren?	s. Fahrtfrequenz
StreckeAuto-bahn	Wie viele Kilometer sind Sie in den letzten zwölf Monaten circa mit dem Auto auf Autobahnen gefahren?	s. Fahrtstrecke
KenntnisAS	Welche Erfahrungen haben Sie persönlich gesammelt mit:	• Unbekannt • bekannt, aber nie benutzt • selten genutzt

	1. Tempomat (CC) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit.] 2. Abstandsregeltempomat (ACC) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Abstand zum Vorderfahrzeug ein.] 3. Aktiver Spurhalteassistent [Dieses System erkennt die Fahrstreifenbegrenzungen und hält das Fahrzeug in den Begrenzungen.] 4. Stauassistent [Dieses System regelt die Geschwindigkeit und den Abstand zum Vorderfahrzeug im Stau und hält dabei das Fahrzeug auf dem Fahrstreifen] 5. Park Assist [Dieses System übernimmt während des Einparkvorgangs nur die Lenkbewegung.] 6. Teilautomation (L2) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Abstand zum	• regelmäßig genutzt
FiltHersteller CC	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Tempomat (CC)	• BMW • VW • Mercedes • Audi • Tesla • Weitere Marken • […]
FiltHerstellerACC	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Abstandsregeltempomat (ACC)	s. FiltHerstellerCC
FiltHersteller Spurha	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Aktiver Spurhalteassistent	s. FiltHerstellerCC
FiltHersteller StauAs	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Stauassistent	s. FiltHerstellerCC
FiltHersteller-ParkAs	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Park Assist	s. FiltHerstellerCC
FiltHersteller Teilau	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Teilautomation (L2)	s. FiltHerstellerCC
Fahrstil	DSQ [15 Items; Uebersetzung durch fka+ika+LfE]	French et al., 1993
Technikaffinitaet	ATI-S [9 Items]	Franke et al., 2019

H-on Follow-up-Questionnaire

[group A (H-on part, filled out twice)]

Beschreibung	Frage	Antwortformat
Metadaten		
Probanden-code	Bitte generieren Sie Ihren persönlichen Versuchspersonen-Code für die Studie. Dieser Code besitzt den Vorteil, dass Sie den Code mittels der Fragen jederzeit neu generieren können, außenstehende Dritte jedoch kaum. Wir benötigen diesen Code, um Ihre Daten der Vorbefragung mit den Daten der Versuchsfahrt zu verknüpfen.	[…]
AutoFahrt	Bitte geben Sie die Automarke an, mit der Sie den Fahrtblock absolviert haben.	• BMW • VW • Ford • Porsche
Fahrtblock	Bitte geben Sie an, ob Sie den Fragebogen für Fahrtblock 1 (zu Beginn der Überlassung) oder für Fahrtblock 2 (Ende der Überlassung) ausfüllen.	• Fahrtblock 1 • Fahrtblock 2
DatumFahrt	An welchem Tag haben Sie den Fahrtblock absolviert?	[…]
DauerFahrt	Bitte schätzen Sie, wie lange Sie am Stück auf der Autobahn gefahren sind. Geben Sie Ihre Schätzung in Minuten an. [Eine kurze Unterbrechung <5 min z.B. zum Wenden kann vernachlässigt werden.]	[…]
L2AnteilFahrt	Bitte schätzen Sie, wie lange Sie am Stück mit aktivem L2 gefahren sind. Geben Sie Ihre Schätzung in Minuten an. [Kurze Unterbrechungen können vernachlässigt werden.]	[…]
KmVorFahrt	Bitte schätzen Sie, wie viele Kilometer Sie insgesamt vor Beginn des Fahrtblocks 1 bzw. des Fahrtblocks 2 auf der Autobahn mit L2 gefahren sind. [Es gilt die Gesamtstrecke an Kilometern auf der Autobahn, die Sie vor Beginn des jeweiligen Fahrtblocks auf der Autobahn bei überwiegender Nutzung von L2 absolviert haben. Zufahrtswege, also Stadt- und Landstraßen zählen nicht dazu.]	[…]
Subjektive Metriken		
Vertrauen	TiA Körber [19 Items]	Körber, 2019
Akzeptanz	CTAM [Subskalen: Performance expectancy (-PE2); Effort expectancy; Attitude towards using technology; Facilitating conditions (-FC4); Behavioral intention to use the system; Perceived safety ausgeschlossen: Subskalen: Anxiety; Self-Efficacy; Social Influence; Auswahl und Uebersetzung durch fka+ika+LfE]	Osswald et al., 2012

Sys- temverstaend nis	16 Items [Zusammenstellung von LfE] • Das System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer. • Das System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden. • Ich muss das system stets überwachen, wenn das System aktiviert ist. • Wenn das System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit. • Ich darf mich mit fahrfremden Tätigkeiten wie z.B. E-Mails schreiben beschäftigen, wenn das System aktiviert ist. • Der Fahrer muss das System bewusst aktivieren. • Das System passt die Geschwindigkeit an die des vorausfahrenden Fahrzeuges an. • Das System kann Fahruntauglichkeit durch Müdigkeit des Fahrers ausgleichen. • Der Fahrer darf sich von der Überwachung des Verkehrsraums abwenden, wenn das System aktiviert ist und eine andere Person im Fahrzeug diese Aufgabe für mich übernimmt. • Der Fahrer darf sich von der Überwachung des Verkehrsraums abwenden, wenn das System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind. • Das System erkennt immer, wenn es eine Situation nicht meistern kann. • Das System lenkt automatisch. • Der Fahrer muss innerhalb von Sekunden die Fahraufgabe übernehmen können. • Ich muss auch bei aktivierter Automation zu jedem Zeitpunkt wach bleiben. • Ich sollte niemals unaufgefordert in die Automation eingreifen. • In der folgenden Situation kann es passieren, dass das System die Situation nicht richtig einschätzen kann und der Fahrer eingreifen muss: Es sind Schlaglöcher auf der Straße, die das Erkennen der Fahrbahnmarkierung erschweren.	• Nicht zutreffend • Zutreffend • Unsicher
NDRTs	8 Items [Zusammenstellung von fka+ika+LfE] • NDRT1 Handy oder ähnliches Gerät (Laptop, externes Navi, Tablet, …) in der Hand – Bedienung SMS/WhatsApp Nachrichten verfassen oder lesen; Browsing; …	• Nie • Sehr selten • Selten • Oft • Sehr oft

	• NDRT2 Handy oder ähnliches Gerät (Tablet, …) in der Hand - Sprechen Telefonieren ohne Freisprechanlage; Aufnehmen von Sprachnachrichten; ... • NDRT3 Handy oder ähnliches Gerät (Tablet, …) fest installiert bzw. mit Freisprechanlage verbunden - Sprechen Telefonieren mit Freisprechanlage; aufnehmen von Sprachnachrichten über Sprachbefehle; ... nicht: Unterhaltung mit Beifahrern, also Personen im Fahrzeug befinden • NDRT4 Bedienung von Systemen im Fahrzeug (nicht direkt relevant für die Fahraufgabe) Bedienen der integrierten Navigation; Einstellungen im Infotainmentsystem vornehmen; Verstellen des Sitzes; Einstellen der Klimaanlage; … • NDRT5 Essen/Trinken/Rauchen Öffnen einer Dose; Essen eines Apfels; Anzünden einer Zigarette; … • NDRT6 Körperpflege/ Make-Up/... Frisieren; Make-Up; Nagelpflege; … Nicht: kurze, unbewusste Handlungen (z.B. kratzen) • NDRT7 Interaktionen mit Beifahrern Unterhalten mit Beifahrer; Gestikulieren vor Beifahrer; Blicke zum Beifahrer; … • NDRT8 Suchen; Greifen; Kramen; … Suchen nach Objekt(en) und hingreifen, z.B. in einer Tasche	
MatrixNDRTs-Freitext	Freitextfeld	[…]
SubjUeberwachungsguete	Wie aufmerksam haben Sie die Systemleistung überwacht, wenn Sie L2 aktiviert hatten?	• 0 Unaufmerksam • 1 • 2 • 3 • 4 • 5 • 6 Stets aufmerksam
SubjEinflussSetting	Hätten Sie sich anders verhalten, wenn Sie die letzte Fahrt nicht im Rahmen einer Studie durchgeführt hätten? Z.B. Beschäftigung mit fahrfremden Tätigkeiten oder Ähnliches?	• Ja • Nein

SubjEinflussSetting-Komm	Sie haben die vorige Frage mit "ja" beantwortet. Bitte beschreiben Sie kurz, inwiefern Sie sich anders verhalten hätten.	[…]
Kommentare	Haben Sie Kommentare zu der heutigen Fahrt bzw. dem erlebten System?	[…]
SonstKommentare	Haben Sie sonstige Kommentare oder Anmerkungen zur Studie?	[…]

H-off Follow-up-Questionnaire

[group A (H-off part) + group B]

Beschreibung	Frage	Antwortformat
Metadaten		
VP	VP-Nummer	[…]
AutoFahrt	Mit welchem Auto wurde die Fahrt absolviert?	• BMW • Ford
DatumFahrt	An welchem Tag wurde der Fahrtblock absolviert?	[…]
DauerFahrt	Schätzung, wie lange der Proband am Stück auf der Autobahn gefahren ist. Schätzung in Minuten. [Kurze Unterbrechungen können vernachlässigt werden.]	[…]
L2AnteilFahrt	Schätzung, wie lange der Proband am Stück mit aktivem L2 gefahren ist. Schätzung in Minuten.	[…]
Probanden-code	Bitte generieren Sie Ihren persönlichen Versuchspersonen-Code für die Studie. Dieser Code besitzt den Vorteil, dass Sie den Code mittels der Fragen jederzeit neu generieren können, außenstehende Dritte jedoch kaum. Wir benötigen diesen Code, um Ihre Daten der Vorbefragung mit den Daten der Versuchsfahrt zu verknüpfen.	[…]
Subjektive Metriken – s. H-on Follow-up-Questionnaire		

Interview

[group A (filled out twice) + group B]

Beschreibung	**Frage**	**Antwortformat**
Metadaten		
VP	VP-Nummer	[…]
Fahrt	Gib an, zu welcher Fahrt dieses Interview geführt wird	• Gruppe A - Termin 2: H-on • Gruppe A - Termin 2: H-off • Gruppe B (H-off)
VL-Protokoll		
AOnVP	Gruppe A - Termin 2: H-on Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt (VL-Protokoll Termin A-1)	[…]
AOnFzg	Gruppe A - Termin 2: H-on Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, ... (VL-Protokoll Termin A-1)	[…]
AOffVP	Gruppe A - Termin 2: H-off Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt	[…]
AOffFzg	Gruppe A - Termin 2: H-off Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, ...	[…]
BOffVP	Gruppe B (H-off) Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt	[…]
BOffFzg	Gruppe B (H-off) Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, …	[…]
VLProt-Sonst	Sonstige Anmerkungen	[…]
Interview-Leitfaden		
Trigger	Trigger Haben Sie während der Fahrt den Triggerknopf gedrückt? Bitte führen Sie noch einmal aus, was Ihnen in diesem Moment aufgefallen ist, bzw. was sie uns mitteilen möchten.	[…]
FahrerTrans	Fahrerinitiierte Transitionen Während der Autobahnfahrten sind Sie einen Großteil der Zeit mit L2 gefahren. In verschiedenen Situationen - zum Beispiel beim Auffahren auf oder Abfahren von Autobahnen - haben Sie das System	[…]

	aktiviert oder deaktiviert, um damit in einen anderen Fahrmodus wechseln. Gab es hierbei Situationen oder Aspekte, über die Sie uns gerne mehr erzählen würden?	
SystemTrans	Systeminitiierte Transitionen Während der Autobahnfahrten sind Sie einen Großteil der Zeit mit L2 gefahren. In verschiedenen Situationen hat das System von sich L2 teilweise oder komplett deaktiviert oder Sie zu einer Übernahme aufgefordert. Gab es hierbei Situationen, über die Sie uns gerne mehr erzählen würden?	[…]
DMS	DMS Wenn L2 aktiv war, gab es ein Fahrerbeobachtungssystem, das gewarnt hat, wenn Sie ON --> die Hände zu lange vom Lenkrad genommen haben. oder OFF --> zu lange von der Straße weggesehen haben. Möchten Sie dieses Fahrerbeobachtungssystem noch einmal genauer kommentieren oder bewerten oder Ihre Erlebnisse mit dem Fahrerbeobachtungssystem beschreiben?	[…]
HMI	HMI Für die Bedienung von L2 haben Sie verschiedene Tasten am Lenkrad verwendet. Je nach Systemzustand wurden Ihnen im Anschluss unterschiedliche Anzeigen im Kombi-Display (oder dem Head-up Display) angezeigt. Möchten Sie allgemein die Bedienung des L2 Systems oder die verwendeten Anzeigen noch einmal genauer kommentieren bzw. bewerten?	[…]
Systemverhalten	Systemverhalten Während der Nutzung von L2 haben Sie die aktive Spurführung, die Geschwindigkeitskontrolle, das Halten des Abstandes und die Interaktion mit anderen Verkehrsteilnehmern in verschiedenen Situationen erlebt. Möchten Sie dieses Fahrverhalten des Systems noch einmal genauer kommentieren bzw. bewerten oder Ihre Erlebnisse beschreiben?	[…]
L2PrivNutzung	Bewertung L2 In diesem Versuch konnten Sie Erfahrungen mit L2	• Nein • Eher nein • Unsicher

	sammeln. Würden Sie dies auch gern in Ihrem privaten Fahrzeug nutzen?	• Eher ja • Ja
L2Komponenten	Bewertung L2 Welche Komponente von L2 würden Sie auch gern in Ihrem privaten Fahrzeug nutzen? [3 Items] • Längsführung/ACC • Querführung/Spurhaltung/Lenken • *nur für A-2 & B beantwortbar:* H-off/Hände frei nutzen können	• Nein • Eher nein • Unsicher • Eher ja • Ja • Nicht anwendbar [Spalte nur für A-2 H-on Interview & Frage 3]
Ranking	Bewertung L2 Stellen Sie sich vor, Ihr privates Fahrzeug wird mit einem L2 System ausgestattet. Hätten Sie lieber die Variante H-on oder die Variante H-off?	• H-on • H-off
L2BewertungKomm	Bewertung L2 Anmerkungen zu den vorigen Fragen/Antworten vom Probanden zur L2-Bewertung.	[…]
Sonstiges	Haben Sie noch weitere Kommentare oder Anmerkungen zu der erlebten Fahrt mit L2?	[…]
Studie	Haben Sie noch weitere Kommentare oder Anmerkungen zu der Studie?	[…]

4.4.6.3 Appendix C – Taxonomy hand positions

Scheme:

Handedness: L and R

- **Hand-on**
  - **Position steering wheel:** 1, 2, 3, 0, 10, 11, 12
  - **Types of grip:** CG, G, GG; O
    - 1F, 2F, 3F, 4F, 5F, H, T; Tb, W, K, O
- **Hand-off**
  - **Grasp space:** A, B, C
  - **Activity:** working (w), resting (r)

Examples:

Both H-on L:11 L:G L:H R:2 R:G R:H

Both H-off L:B L:r R:B R:w

Different L:11 L:G L:H R:no R:no R:no; L:no L:no R:B R:r

Short-term unconscious actions like scratching or short keystrokes are not annotated!

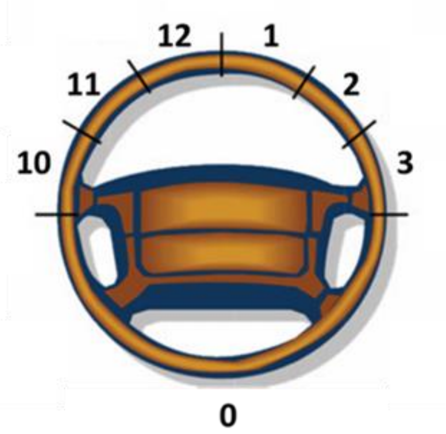


Figure 4-72

Contact grip (CG)	1 Finger (1F)	 
	2 Finger (2F)	 

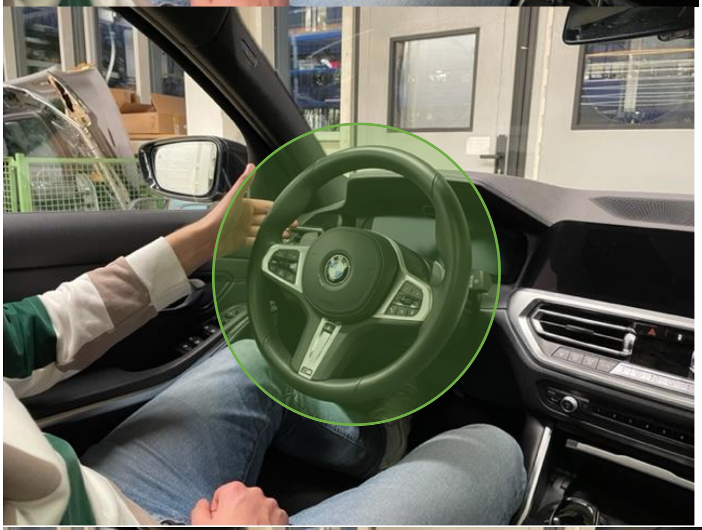
	3 Finger (3F)	 
	4 Finger (4F)	 

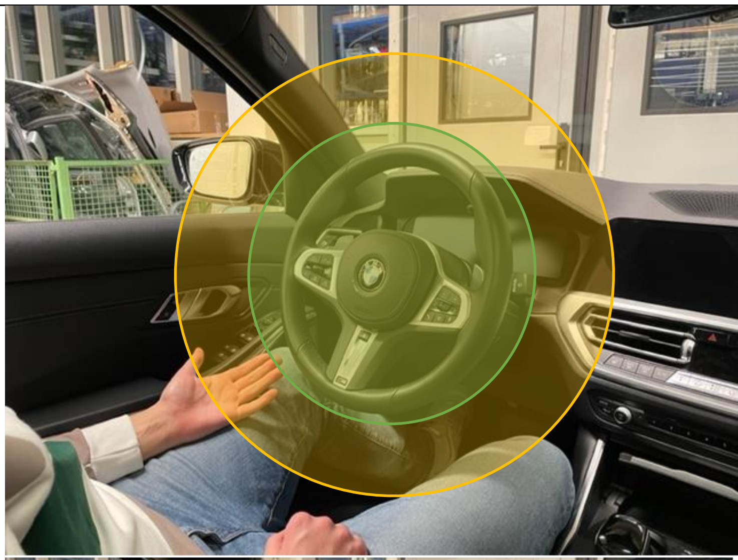
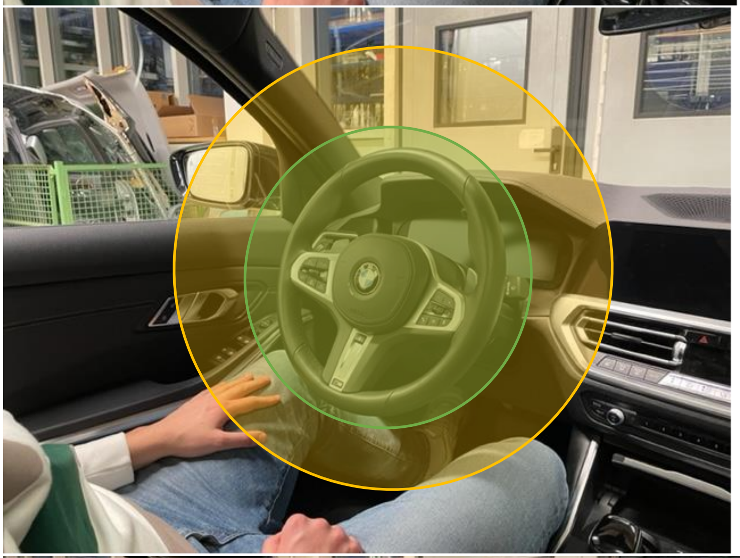
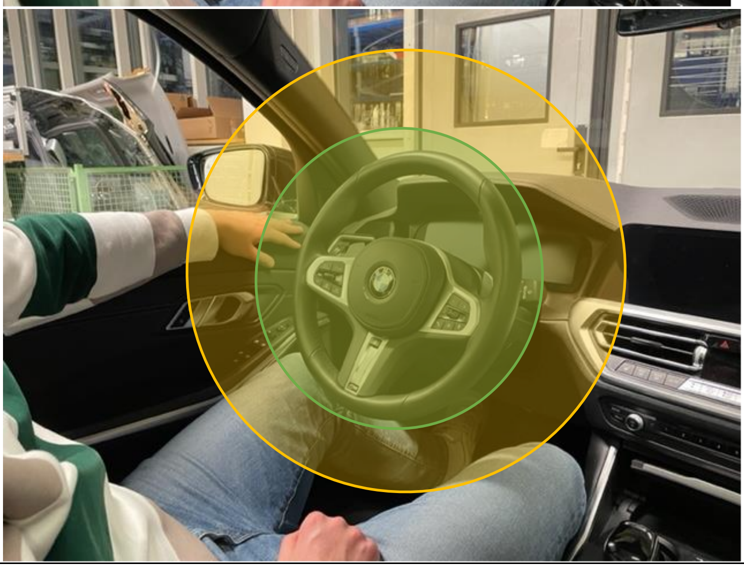
	Hand (H)	
	Daumen (D)	
Grasp (G)	2 Finger (2F)	
	3 Finger (3F)	

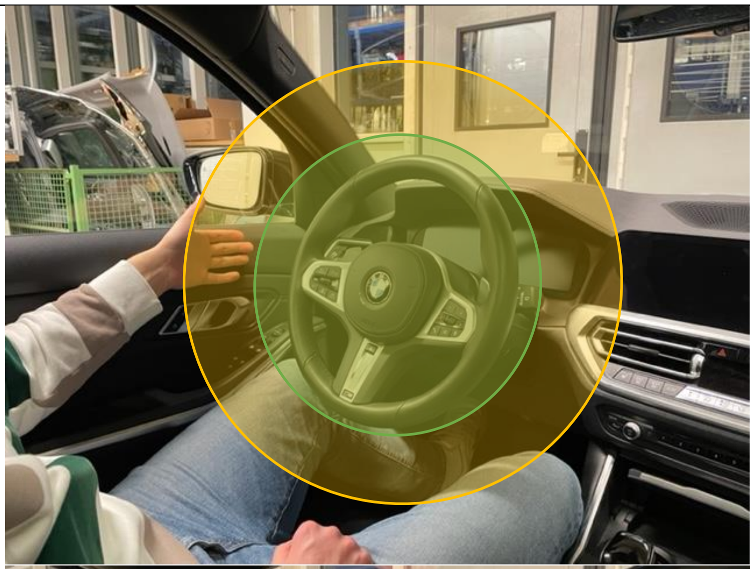
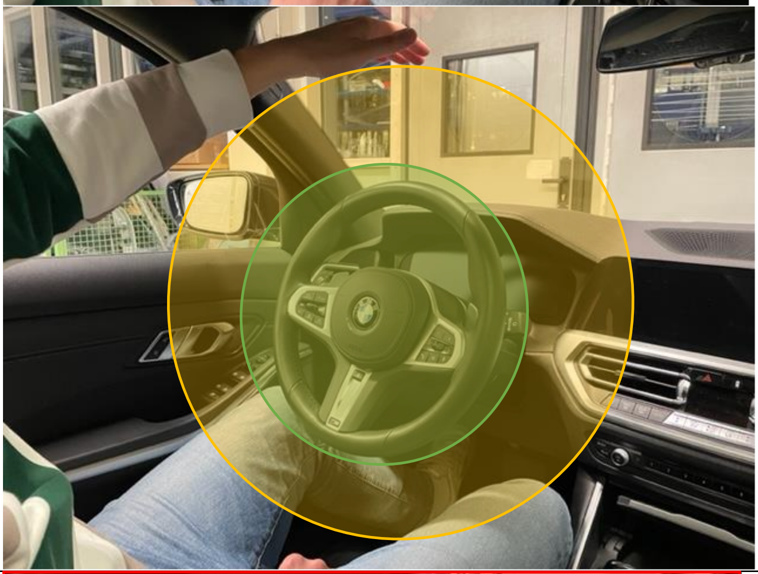
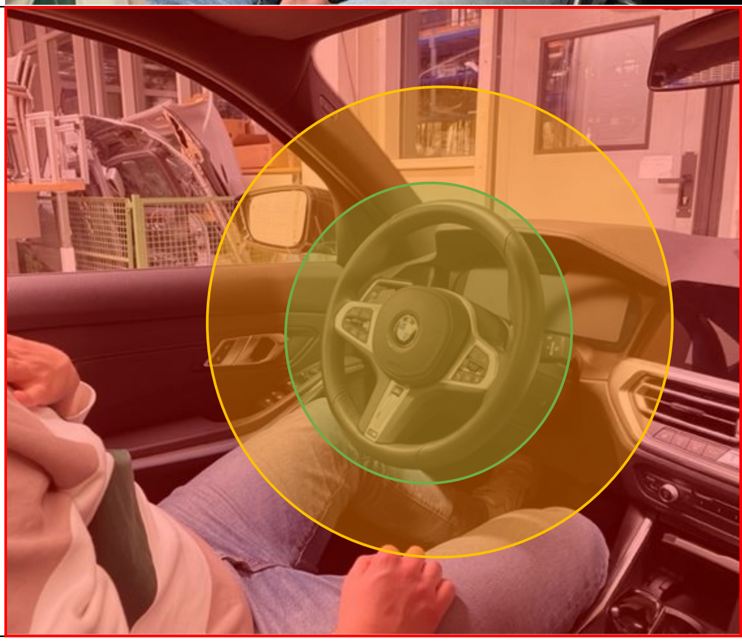
	4 Finger (4F)	
	5 Finger (5F)	
	Hand (H)	
Grasp grip (GG)	2 Finger (2F)	

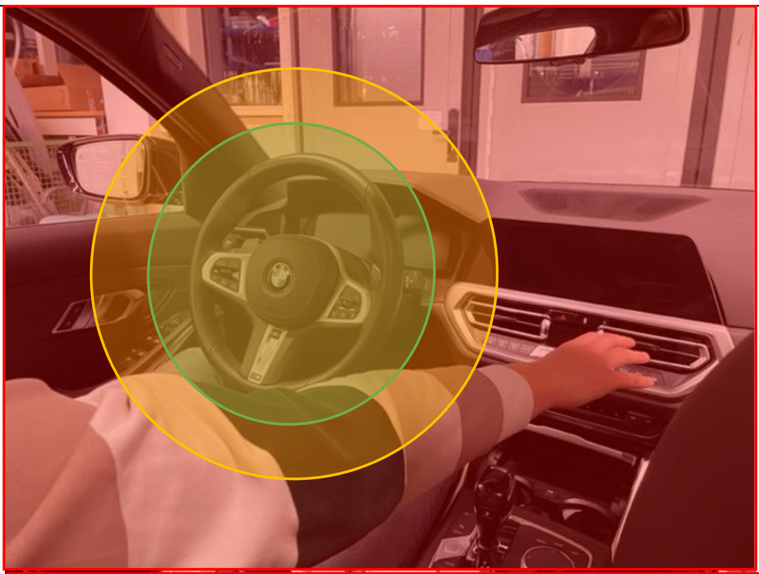
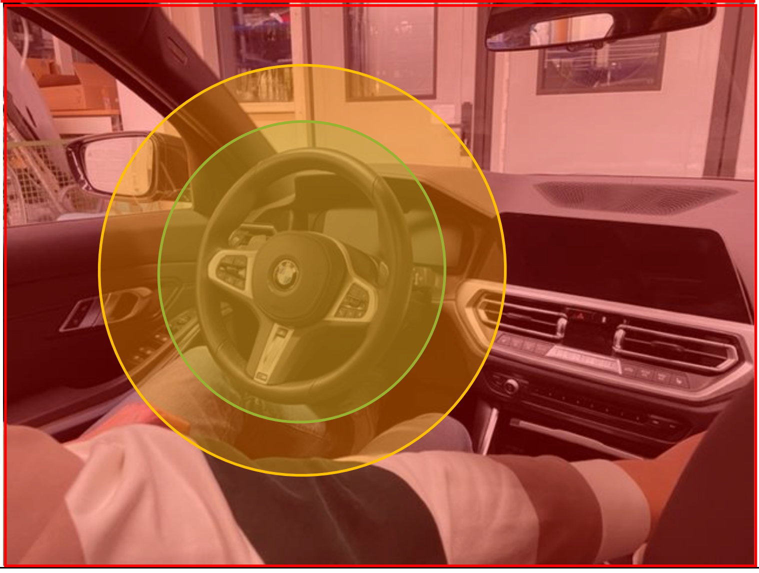
	3 Finger (3F)	
	4 Finger (4F)	
	Hand (H)	

<table>
<tr><td rowspan="4">Other (O)</td><td>Ball of the thumb(Tb)</td><td></td></tr>
<tr><td>Wrist (W)</td><td></td></tr>
<tr><td>Knee (K)</td><td></td></tr>
<tr><td>Other (O)</td><td></td></tr>
</table>

Grasp spaces Hands-off	**Examples**
A (maximum of 5cm)	  

B (maximum of 20cm)	  

	 
C	

4.4.6.4 Appendix D – Instructions and methodological guidelines

5_H-on_InstruktionL2

Teilautomation Level 2 [L2]

Im Rahmen dieses Projektes sollen technische und ergonomische Anforderungen an Level 2 automatisiertes Fahren ermittelt werden.

Level 2 automatisiertes Fahren (abgekürzt L2) bedeutet, dass Fahrerassistenzsysteme die Längs- und Querführung des Autos in bestimmten Anwendungsbereichen unterstützen. Das Fahrzeug übernimmt also sowohl die Längsführung (Beschleunigen/Bremsen) als auch die Querführung (Lenken). Das System passt sich dabei an die Geschwindigkeit des umgebenden Verkehrs an, bremst bei Bedarf eigenständig und folgt der Fahrspur.

Auch während das L2 System aktiv ist, tragen Sie die volle Verantwortung für die Fahrweise des Fahrzeugs, das System dient lediglich der Unterstützung. Sie müssen das System und die Umgebung jederzeit überwachen und sind in der Verantwortung, die Verkehrssituation richtig einzuschätzen. Aufgrund von Systemgrenzen kann das System nicht in allen Verkehrssituationen selbsttätig angemessen reagieren. Bei Vorliegen eines Systemfehlers, bei einer Übernahmeaufforderung durch das System oder wenn die Voraussetzungen für eine sichere Fahrt mit dem L2 System nicht mehr gegeben sind, müssen Sie die Fahrweise den Verkehrsverhältnissen anpassen und die Fahraufgabe unverzüglich sowie ggf. eigenständig wieder übernehmen.

Während L2 aktiv ist, dürfen Sie die Füße von den Pedalen nehmen. Die Hände dürfen nicht vom Lenkrad genommen werden. Wenn Sie die Hände vom Lenkrad nehmen, werden Sie vom System aufgefordert, die Hände wieder ans Lenkrad zu nehmen. Die Aufmerksamkeit muss jederzeit auf das Verkehrsgeschehen und die Funktionsweise des Automationssystem gerichtet bleiben. Womöglich erleben Sie bei Abwenden der Aufmerksamkeit eine Warnmeldung.

5_H-off_InstruktionL2

Teilautomation Level 2 [L2]

Im Rahmen dieses Projektes sollen technische und ergonomische Anforderungen an Level 2 automatisiertes Fahren ermittelt werden.

Level 2 automatisiertes Fahren (abgekürzt L2) bedeutet, dass Fahrerassistenzsysteme die Längs- und Querführung des Autos in bestimmten Anwendungsbereichen unterstützen. Das Fahrzeug übernimmt also sowohl die Längsführung (Beschleunigen/Bremsen) als auch die Querführung (Lenken). Das System passt sich dabei an die Geschwindigkeit des umgebenden Verkehrs an, bremst bei Bedarf eigenständig und folgt der Fahrspur.

Auch während das L2 System aktiv ist, tragen Sie die volle Verantwortung für die Fahrweise des Fahrzeugs, das System dient lediglich der Unterstützung. Sie müssen das System und die Umgebung jederzeit überwachen und sind in der Verantwortung, die Verkehrssituation richtig einzuschätzen. Aufgrund von Systemgrenzen kann das System nicht in allen Verkehrssituationen selbsttätig angemessen reagieren. Bei Vorliegen eines Systemfehlers, bei einer Übernahmeaufforderung durch das System oder wenn die Voraussetzungen für eine sichere Fahrt mit dem L2 System nicht mehr gegeben sind, müssen Sie die Fahrweise den Verkehrsverhältnissen anpassen und die Fahraufgabe unverzüglich sowie ggf. eigenständig wieder übernehmen.

Während L2 aktiv ist, dürfen Sie die Füße von den Pedalen und die Hände vom Lenkrad nehmen. Die Aufmerksamkeit muss jederzeit auf das Verkehrsgeschehen und die Funktionsweise des Automationssystem gerichtet bleiben. Womöglich erleben Sie bei Abwenden der Aufmerksamkeit eine Warnmeldung.

0_Leitfaden_A1

Übergabe Fahrzeug Gruppe A – Termin 1

Legende:

Gerader, normaler Text: Hinweise

Gerader, fett markierter Text: Ortswechsel/Ortsangabe

Kursiver, Text: Sprechteile

Gelbe Markierung: Verweis auf andere Dokumente

Vorbereitung: kurz zuvor

- Fahrzeug
  - In Innenhof stellen
  - Fahrtenbuch_OEM: Probandentest + VP-Nummer; km-Beginn; Datum Übergabe
  - Klemmbrett Fahrzeug
    - Dok 4_Fahrtenbuch_Proband
- Im Besprechungsraum/0330 (Niklas'Büro)
  - Allgemeines
    - Spender zur Handdesinfektion + Tücher für Flächendesinfektion
    - Selbsttests
    - Stift (2x, VL und Proband)
    - Getränke + Süßigkeiten
  - Versuchsleiter
    - Dok 0_Leitfaden_A1
    - Dok 1_Ueberlassungsvereinbarung
      → Ergänzungen auf #1,#2,#5,#7 vornehmen
    - Fahrzeugschlüssel + Schrankentransponder
    - Klemmbrett VL (für am & in Auto)
      - Dok 3_Uebergabeprotokoll
      - Dok 5_H-on_InstruktionL2
      - Dok 6_VL-Protokoll
  - Proband
    - Dok 1_Ueberlassungsvereinbarung Blanko zum Mitnehmen
    - Dok 2_H-on_MARKE_Handbuch geheftet zum Mitnehmen
- Geplante Route auf Staus/Unfälle checken

0_Leitfaden_A1

Abholen und Begrüßen

Treffpunkt Probanden: Fahrstühle Hof 3, EG

Proband begrüßen und **zu Besprechungsraum/0330 (Niklas' Büro) führen**

Ggf. Privatfahrzeug innerhalb Schranke parken

Im Besprechungsraum/0330 (Niklas'Büro)

> *„Herzlich Willkommen zu der Studie „Automatisiertes Fahren im Straßenverkehr". Zunächst einmal vielen Dank für Ihre Teilnahme an der Studie und dass Sie sich heute Zeit genommen haben für diesen ersten Termin vor Ort. In dem heutigen Termin möchte ich mit Ihnen noch einmal die Rahmenbedingungen der Fahrzeugüberlassung und die Funktionsweise der verwendeten Automationssysteme klären. Im Anschluss sehen wir uns gemeinsam das Fahrzeug an, das Sie nach dem heutigen Termin für einige Tage mit nach Hause nehmen dürfen. In einer gemeinsamen Trainingsfahrt werde ich Ihnen hier alle wichtigen Funktionen zeigen. Der Termin dauert etwa anderthalb Stunden.*
>
> *Für den heutigen Termin wie auch für den zweiten Termin zur Fahrzeugrückgabe gilt 3G. Daher würde ich Sie nun bitten, mir einen entsprechenden Impf- bzw. Testnachweis zu zeigen."*

Impf- /Testnachweis checken

Kopie des Führerscheins machen [Drucker unten → Ausweiskopie]

Überlassungsvereinbarung

Dok 1_Ueberlassungsvereinbarung

> *„Das Dokument vor Ihnen ist die Überlassungsvereinbarung, in der die Rahmenbedingungen der Fahrzeugüberlassung festgehalten sind. Sie haben sie bereits per E-Mail erhalten. Haben Sie dazu noch Fragen?"*

Auf folgende Aspekte noch einmal eingehen:

- Einschränkungen bzgl. Fahrten → keine privaten Fahrten
- Anforderungen bzgl. Parken
- Unfälle bei Polizei und Lehrstuhl melden
- Hinweise bzgl. Messtechnik, z.B. Maximalgeschwindigkeit, Technikcheck, Waschstraße

Auf 1_Ueberlassungsvereinbarung #5 unterschreiben lassen → abheften

2 / 8

Versuchsleitfaden Gruppe A Termin 1

0_Leitfaden_A1

Handbuch

Dok 2_H-on_MARKE_Handbuch

"Dieses Handbuch enthält eine Zusammenfassung der wichtigsten Hinweise für Sie. Sie finden es auch im Fahrzeug. Sie können also jederzeit Informationen nachlesen.
Enthalten ist eine Zusammenfassung Ihrer Fahraufgabe. Außerdem Checklisten für die Zeitpunkte ***vor, während*** *und* ***nach*** *der Fahrt. Zudem sind Informationen zum Rückgabeort, zum Fahrzeug und zum Assistenzsystem Teilautomation L2 enthalten. Sie können sich darüber hinaus bei Fragen und Problem jederzeit per E-Mail oder Telefon an uns wenden."*

3 – Handbuch

- Ziel: 2 lange Fahrten zum Abgleich „kaum Erfahrung mit L2“ vs. „viel Erfahrung mit L2“; dazwischen weitere Fahrten zur weiteren Gewöhnung
- **Wichtig**: Max. 50 km fahren, dann Fahrtblock 1; insgesamt mind. 200km, max. 350km fahren, dann Fahrtblock 2
  ggf. konkrete Pläne für die Fahrten mit Proband durchsprechen
- Markierung Beginn des Blocks (Verweis „wird im Fahrzeug gezeigt“)
- 45min durchgehende Fahrt
- Danach Fragebogen (selber Tag, möglichst zeitnah an Fahrt)

#4 - Handbuch

- Checkliste für vor der Fahrt
- Fahrtankündigung aus Versicherungsgründen (vor Trainingsfahrt nicht notwendig)
- Alles Weitere am Fahrzeug

0_Leitfaden_A1

Fahrzeugerprobung im Stillstand

Zu Fahrzeug führen → dabei auf Toilette hinweisen + Snacks & Getränke mitnehmen lassen + persönliche Gegenstände mitnehmen

Übergabeprotokoll ausfüllen

→ Dok 3_Uebergabeprotokoll

Ggf. Tankfüllstand (nur ¾ voll) ergänzen

Auf Wunsch Kopie anfertigen für Probanden

Erklärungen vor Fahrzeug

Position und Bedienung von

- Tank-/Ladedeckel (inkl. Kabel, Karte und bei Bedarf kurze Erklärung Ladevorgang)
- Warndreieck, Warnweste und Verbandkasten zeigen
- Kofferraum (Hinweis lose Gegenstände)

Zündung an

→ Verweis Dok 2_H-on_MARKE_Handbuch #4
Messsystem zeigen inkl. Messtechnik-Check

- LIDAR (Dach) & RADAR (vorne) nicht stark verschmutzt
- Datenaufzeichnung → LED links sowie unterste LED rechts leuchtet/blink grün
- Kameras in Position / nicht abgefallen (Hinweis: nicht anfassen/verstellen, ggf. Dachhimmel, Sonnenblende & Sonnenbrille)
- Trigger-Knopf leuchtet rot

Erklärungen im Fahrzeug

Grundbedienung

- Sitz/Lenkrad/Spiegel einstellen
- Motor anschalten
- Automatik erklären (D für Drive, R für Rückwärts, P für Handbremse, ggf. zusätzliches)
- Licht: An- und Ausschalten (falls nicht automatisiert)
- Scheibenwischer (falls nicht automatisiert)

Anzeigeelemente (Non-Assistenz)

- Kombi:
  - Kilometerstand (→ hier checken, wie viel schon gefahren + Fahrtenbuch)
  - Uhrzeit (→ hier checken wie lange schon gefahren + Fahrtenbuch)
  - Tank-/Ladeanzeige & Reichweitenanzeige
  - Weitere relevante Icons/Anzeigeelemente (Non-Assistenz)
  - Hinweis: keine dauerhaften Änderungen am Layout

- Head-up Display
  - Relevante Icons/Anzeigeelemente (Non-Assistenz)
- CID: Generelle Struktur + Bedienung (sehr kurzhalten, Hinweis auf Navi und Mediasystem inkl. Radio)
- Ggf. weitere relevante Anzeigeelemente / Bedienelemente (z.B. Touchscreen, iDrive, etc.)

Funktionen (Non-Assistenz)

- Klimaanlage: Allgemeintemperatur und Scheibenlüftung zeigen
- CID: Hinweis: Alles darf geändert werden, **außer Einstellungen Fahrerassistenz**
  - Ausnahmen: s. S.9 in Handbuch
- Hinweis: Im Zweifelsfall Rücksprache mit Studienteam

Allgemeines

- Bedienungsanleitung Fahrzeug + Fahrzeugschein zeigen (im Handschuhfach)
- Studien-Handbuch zeigen
- Fahrtenbuch erklären und ausfüllen lassen → Verweis Dok 2_H-on_MARKE_Handbuch #4
- Trigger-Knopf → Verweis Dok 2_H-on_MARKE_Handbuch #5 → beide Varianten ausprobieren lassen (3x für Markierung Beginn Fahrtblock + 1x für Kommentieren: drücken + sprechen) → wichtig: *„auch wenn ungewohnt, bitte verwenden, da hilfreich für uns"*

Probefahrt auf Gelände

1-2 Male im Kreis fahren, dann Stop vor Lehrstuhl

0_Leitfaden_A1

Instruktion L2H-on

> *„Ich erkläre Ihnen nun das Automationssystem, das Sie in der Studie überwiegend nutzen werden. Bitte lesen Sie sich dafür zunächst diese allgemeine Beschreibung zu Teilautomation L2 durch."*

→ Verweis Dok 5_H-on_InstruktionL2

Fragen klären

Anleitung L2

→ Verweis Dok 2_H-on_MARKE_Handbuch #10 (f.)

> *„In Ihrem Fahrzeug wird das L2 System „XY" genannt. Im Folgenden zeige ich Ihnen die wichtigsten Bedienschritte für das System. Bitte beachten Sie, dass dies nur einen ersten Überblick darstellt. Konsultieren Sie die Bedienungsanleitung im Handschuhfach für wichtige weitere Hinweise zu Anzeige, Bedienung und Grenzen des Automationssystems."*

Hinweise aktives L2:

- Icon/Anzeige und Farbe beschreiben
- Wenn L2 aktiv ist, übernimmt es die Quer- und die Längsführung. Es kann sein, dass aufgrund von Verkehrs- oder Umweltbedingungen Teile des Systems deaktiviert werden. Dann verändert sich das markierte Icon in der Farbe oder in der Form. Beachten Sie, dass Sie als Fahrer immer für die Fahraufgabe verantwortlich sind.
  → *ggf. Änderungen des Icons/der Anzeige bei (Teil-)Deaktivierung kurz beschreiben*

Bedienschritte (durchgehen und auf Tasten hinweisen, aber nicht ausprobieren lassen)

- Aktivieren
- Deaktivieren
- Geschwindigkeit einstellen
- Abstand einstellen
- Spurwechsel durchführen

Falls Tasten ausprobiert werden → Zündung an & aus **und** kontrollieren, ob Systemstatus unverändert (z.B. ACC und/oder Spurhalteassistent noch im Standby aktiv ist)

> *„Haben Sie hierzu aktuell noch Fragen? Wir werden alle Anzeige- und Bedienelement gleich in einer Trainingsfahrt gemeinsam austesten."*

6 / 8

Versuchsleitfaden Gruppe A Termin 1

0_Leitfaden_A1

Fahrzeugerprobung auf Straße

Vor der Fahrt: **Individuelle Kalibrierung** der (auf Augen gerichteten) Kameras

- VP sollte ab jetzt Maske absetzen
- Fahrposition einnehmen → auf Punkte blicken
- **Trigger 2x drücken** (Markierung Kalibrierung in den Daten)

1. Entspannt Geradeaus
2. Rückspiegel
3. CID
4. Mittelkonsole unten, z.B. Drehdrücksteller
5. Kombi
6. Lenkradtasten links
7. Linker Seitenspiegel
8. Rechter Seitenspiegel
9. Ggf. Head-up Display

Start Trainingsfahrt (Fahrt: A9 nach 67 Allershausen oder 66 Pfaffenhofen und wieder zurück)

- Beginn Trainingsfahrt: A9 Richtung Norden & wieder zurück; insgesamt ~15-20 Min
- L2 ausprobieren
- Anweisungen Versuchsleiter jederzeit Folge leisten
- Informieren, falls Unwohlsein oder Pause benötigt
- Proband frühzeitig navigieren und auf Fußgänger hinweisen

Start Autobahn

Protokoll führen→ Verweis Dok 6_VL-Protokoll

- Proband: Beobachtungen, z.B. nimmt Hände sofort weg, Misuse, etc.)
- Auto: Beobachtungen, seltsames Verhalten, z.B. Eyes Off Warning, obwohl Blick auf Straße, ...

Zu beachten

- Sehr individuell nach Bedürfnissen gestalten, ausreichend Zeit lassen, ggf. Pausen einlegen
- Auf Fehlverhalten hinweisen
- Auf Verantwortung des Fahrers, Sicherheit als höchste Prio & Maximalgeschwindigkeit hinweisen
- Fragen klären, proaktiv nach Fragen fragen
- Ggf. Hinweis: Erkennung aktuelle Geschwindigkeit nicht immer zuverlässig

Manöver

- Manuelle Manöver: Bremsen, Beschleunigen, Spurwechsel
- L2
  - Aktivierung
  - Deaktivierung (verschiedene Wege)
  - Zielgeschwindigkeit anpassen
  - Abstand regeln (nicht bei iX, da dort über CID)
  - Spurwechsel durchführen
- Trigger-Knopf
  - Drücken & ausprobieren (sprechen lassen)

7 / 8

Versuchsleitfaden Gruppe A Termin 1

0_Leitfaden_A1

Schranke

Ggf. aussteigen & selbst öffnen

Ende

- **Trigger 2x drücken** (während Zündung aktiv) → dies markiert das Ende der Trainingsfahrt & bedeutet, dass als nächstes die Versuchsfahrt folgt

Im Hof

Fragen klären

Fahrtenbuch ausfüllen → Verweis Dok 2_H-on_MARKE_Handbuch #7 + Hinweis auf Fragebogen nach den Fahrtblöcken (kommt per Mail)

Toilette, Snacks & Getränke vor Weiterfahrt anbieten

Hinweis: Kontaktieren bei Fragen/Unklarheiten

Fahrtankündigung:

- Wie im Dok 2_H-on_MARKE_Handbuch #4
- Alternativ: Notieren im Dok 6_VL-Protokoll: welche Fahrt findet im Anschluss statt (Fahrt nach Hause oder Erprobungsfahrt → bei letzterem grobe Route notieren)

Nachbereitung:

Ausformulieren von Dok 6_VL-Protokoll (Ergänzungen vornehmen)

Abheften von

- Führerschein
- Dok 1_Ueberlassungsvereinbarung
- Dok 3_Uebergabeprotokoll
- Dok 6_VL-Protokoll

Desinfizieren von Tischen, Stiften und Instruktionen

0_Leitfaden_A2

0_Leitfaden_A2

Übergabe Fahrzeug Gruppe A – Termin 2

Legende:

Gerader, normaler Text: Hinweise

Gerader, fett markierter Text: Ortswechsel/Ortsangabe

Kursiver, Text: Sprechteile

Gelbe Markierung: Verweis auf andere Dokumente

Vorbereitung: kurz zuvor

- H-off Fahrzeug
    - In Innenhof stellen
    - Fahrtenbuch_OEM bereitlegen und vorausfüllen: Probandentest + VP-Nummer; km-Beginn; Datum
    - Check: Dok 2_H-off_MARKE_Handbuch im Fahrzeug?
- Im Besprechungsraum/0330 (Niklas'Büro)
    - Allgemeines
        - Spender zur Handdesinfektion + Tücher für Flächendesinfektion
        - Selbsttests
        - Stift (2x, VL und Proband)
        - Getränke + Süßigkeiten
    - Versuchsleiter
        - Dok 0_Leitfaden_A2
        - Dok 7_Aufwandsentschaedigung-Bestaetigung
        - Dok 8_Einverstaendniserklärung VP-Datenbank
        - Dok 6_VL-Protokoll von Termin A1 (H-on) bereitlegen (während Lime H-off Nachbefragung in Lime FOT Interview einpflegen)
        - Laptop für Lime FOT Interview (i.d.R. eigener Laptop)
        - Fahrzeugschlüssel + Schrankentransponder
        - Klemmbrett VL
            - Dok 3_Uebergabeprotokoll für Rückgabe H-on (im Ordner)
            - Dok 5_H-off_InstruktionL2
            - Dok 6_VL-Protokoll
    - Proband
        - Laptop für Lime H-on Nachbefragung und Lime H-off Nachbefragung
- Geplante Route auf Staus/Unfälle checken

0_Leitfaden_A2

Abholen und Begrüßen

Treffpunkt Probanden: Schranke

Proband begrüßen, in den Hof fahren; Fahrzeug abstellen

Prüfen: keine Gegenstände in Fahrzeug vergessen?

Dok 3_Uebergabeprotokoll ausfüllen; ggf. Schäden protokollieren

Fahrtenbuch Proband prüfen → Plausibilitätscheck, ob Fahraufgabe erfüllt. Falls ja, nahtlos weitermachen, falls nicht, direkt zu „Abschied" springen.

Im Besprechungsraum/0330 (Niklas'Büro)

Ggf. Impf- /Testnachweis (nochmal) checken

Abfragen

- Kommentare/Probleme letzte Tage`
- Triggerbutton gedrückt?
- Beide Fahrtblocks absolviert?
- Fragebogen 2x ausgefüllt?
  → falls FB2 noch nicht ausgefüllt: dann nachholen

Erklärung Ablauf Termin 2

- Interview (15min)
- Ggf. Pause
- Einweisung Versuchsfahrzeug + L2
- Trainingsfahrt 15-20min
- Versuchsfahrt 45-60 min
- Nachbefragung – FB (20min)
- Interview (10min)

Interview

Laptop VL: Lime FOT Interview laden → VP & Termin A-On angeben → als Leitfaden nutzen, direkt ausfüllen, selbst sortieren, in welche Felder Antworten einzuordnen sind.

0_Leitfaden_A2

Fahrzeugerprobung im Stillstand

Zu Fahrzeug führen → dabei auf Toilette hinweisen + Snacks & Getränke mitnehmen lassen + persönliche Gegenstände mitnehmen

Erklärungen im Fahrzeug
kurz halten, da für h-off eher nicht relevant,
die Fahrer sollten aber kurz Zeit haben, sich mit dem Fahrzeug und auch mit dem Kombi vertraut zu machen!

Grundbedienung

- Sitz/Lenkrad/Spiegel einstellen
- Motor anschalten
- Automatik erklären (D für Drive, R für Rückwärts, P für Handbremse, ggf. zusätzliches)
- Licht: An- und Ausschalten (falls nicht automatisiert)
- Scheibenwischer (falls nicht automatisiert)

Anzeigeelemente (Non-Assistenz)

- Kombi:
    - Kilometerstand (→ hier checken, wie viel schon gefahren + Fahrtenbuch)
    - Uhrzeit (→ hier checken wie lange schon gefahren + Fahrtenbuch)
    - Tank-/Ladeanzeige & Reichweitenanzeige
    - Weitere relevante Icons/Anzeigeelemente (Non-Assistenz)
    - Hinweis: keine dauerhaften Änderungen am Layout
- Head-up Display
    - Relevante Icons/Anzeigeelemente (Non-Assistenz)
- CID: Generelle Struktur + Bedienung (sehr kurzhalten, Hinweis auf Navi und Mediasystem inkl. Radio)
- Ggf. weitere relevante Anzeigeelemente / Bedienelemente (z.B. Touchscreen, iDrive, etc.)

Funktionen (Non-Assistenz)

- Klimaanlage: Allgemeintemperatur und Scheibenlüftung zeigen
- CID: Hinweis: Alles darf geändert werden, **außer Einstellungen Fahrerassistenz**
    - Ausnahmen: s. Dok 2_H-off_MARKE_Handbuch #2
- Hinweis: Im Zweifelsfall Rücksprache mit Studienteam

Allgemeines

- Kamerapositionen zeigen (ggf. Hinweis auf Sonnenblende)
- Trigger-Knopf → Dok 2_H-off_MARKE_Handbuch #1

Probefahrt auf Gelände

1-2 Male im Kreis fahren, dann Stop vor Lehrstuhl

0_Leitfaden_A2

Instruktion L2H-off

> *„Ich erkläre Ihnen nun das Automationssystem, das Sie in der Studie überwiegend nutzen werden. Bitte lesen Sie sich dafür zunächst diese allgemeine Beschreibung zu Teilautomation L2 durch."*

→ Dok 5_H-off_InstruktionL2

Fragen klären

Anleitung L2

→ Dok 2_H-off_MARKE_Handbuch #3f.

> *„In Ihrem Fahrzeug wird das L2 System „XY" genannt. Im Folgenden zeige ich Ihnen die wichtigsten Bedienschritte für das System. Bitte beachten Sie, dass dies nur einen ersten Überblick darstellt. Konsultieren Sie die Bedienungsanleitung im Handschuhfach für wichtige weitere Hinweise zu Anzeige, Bedienung und Grenzen des Automationssystems."*

Hinweise aktives L2:

- Icon/Anzeige und Farbe beschreiben
- Wenn L2 aktiv ist, übernimmt es die Quer- und die Längsführung. Es kann sein, dass aufgrund von Verkehrs- oder Umweltbedingungen Teile des Systems deaktiviert werden. Dann verändert sich das markierte Icon in der Farbe oder in der Form.
  Beachten Sie, dass Sie als Fahrer immer für die Fahraufgabe verantwortlich sind.
  → ggf. Änderungen des Icons/der Anzeige bei (Teil-)Deaktivierung kurz beschreiben

Bedienschritte (durchgehen und auf Tasten hinweisen, aber nicht ausprobieren lassen)

- Aktivieren
- Deaktivieren
- Geschwindigkeit einstellen
- Abstand einstellen
- Spurwechsel durchführen

Falls Tasten ausprobiert werden → Zündung an & aus **und** kontrollieren, ob Systemstatus unverändert (z.B. ACC und/oder Spurhalteassistent noch im Standby aktiv ist)

> *„Haben Sie hierzu aktuell noch Fragen? Wir werden alle Anzeige- und Bedienelement gleich in einer Trainingsfahrt gemeinsam austesten."*

4 / 7

Versuchsleitfaden Gruppe A Termin 2

0_Leitfaden_A2

Fahrzeugerprobung auf Straße

Vor der Fahrt: **Individuelle Kalibrierung** der (auf Augen gerichteten) Kameras

- VP sollte ab jetzt Maske absetzen
- Fahrposition einnehmen → auf Punkte blicken
- **Trigger 2x drücken** (Markierung Kalibrierung in den Daten)

1. Entspannt Geradeaus
2. Rückspiegel
3. CID
4. Mittelkonsole unten, z.B. Drehdrücksteller
5. Kombi
6. Lenkradtasten links
7. Linker Seitenspiegel
8. Rechter Seitenspiegel
9. Ggf. Head-up Display

Start Trainingsfahrt (Fahrt: A9 nach 66 Pfaffenhofen P&R)

- Beginn Trainingsfahrt: A9 Richtung Norden; insgesamt ~15-20 Min
- L2 ausprobieren
- Anweisungen Versuchsleiter jederzeit Folge leisten
- Informieren, falls Unwohlsein oder Pause benötigt
- Proband frühzeitig navigieren und auf Fußgänger hinweisen

Start Autobahn

Protokoll führen→ Dok 6_VL-Protokoll

- Proband: Beobachtungen, z.B. nimmt Hände sofort weg, Misuse, etc.)
- Auto: Beobachtungen, seltsames Verhalten, z.B. Eyes Off Warning, obwohl Blick auf Straße, ...

Zu beachten

- Sehr individuell nach Bedürfnissen gestalten, ausreichend Zeit lassen, ggf. Pausen einlegen
- Auf Fehlverhalten hinweisen
- Auf Verantwortung des Fahrers, Sicherheit als höchste Prio & Maximalgeschwindigkeit hinweisen
- Fragen klären, proaktiv nach Fragen fragen
- Ggf. Hinweis: Erkennung aktuelle Geschwindigkeit nicht immer zuverlässig

Manöver

- Manuelle Manöver: Bremsen, Beschleunigen, Spurwechsel
- L2
  - Aktivierung
  - Deaktivierung (verschiedene Wege)
  - Zielgeschwindigkeit anpassen
  - Abstand regeln (nicht bei iX, da dort über CID)
  - Spurwechsel durchführen
- Trigger-Knopf
  - Drücken & ausprobieren (sprechen lassen) → wichtig: *„auch wenn ungewohnt, bitte verwenden, da hilfreich für uns"*

Ende

- **Trigger 2x drücken** (während Zündung aktiv) → dies markiert das Ende der Trainingsfahrt & bedeutet, dass als nächstes die Versuchsfahrt folgt

Versuchsfahrt

Dauer >45min

Route 1: A9: Von 66-Pfaffenhofen nach 63-Manching, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~17 min sein; falls deutlich kürzer, dann weiterfahren nach 62-Ingolstadt-Süd (statt 63-Manching).

Alternativen:

Route 2: A9+A92: Von 66-Pfaffenhofen nach 10-Moosburg-Süd, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~28min sein; falls deutlich kürzer, dann weiterfahren nach 12-Landshut-West (statt 10-Moosburg-Süd).

Route 3: A9+A92+A99+A8: Von 66-Pfaffenhofen nach 77-Sulzemoos, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~33min sein; falls deutlich kürzer, dann weiterfahren nach 76-Odelzhausen (statt 77-Sulzemoos).

Hinweise:

- während der Fahrt Auffälligkeiten über Trigger kommentieren, nicht direkt mit VL kommunizieren, da lediglich Funktion Sicherheitsfahrer
- bei Baustellen: bei An- und Abfahrt in Baustelle Hände ans Lenkrad + rechte Spur

Markierung Beginn der Fahrt (auf Autobahn) durch **3x Druck des Triggers (<1s)** – ACHTUNG VL macht das – nicht der Proband

Protokoll führen → Dok 6_VL-Protokoll

- Proband: Beobachtungen, z.B. nimmt Hände sofort weg, Misuse, etc.)
- Auto: Beobachtungen, seltsames Verhalten, z.B. Eyes Off Warning, obwohl Blick auf Straße, ...
- Weiteres: Auffälligkeiten, Probleme, Kommentare

Schranke

Ggf. aussteigen & selbst öffnen

Fahrzeug parken

Im Hof

Fahrtenbuch ausfüllen

Toilette, Snacks & Getränke vor Befragung anbieten

Nachbefragung

Im Besprechungsraum/0330 (Niklas' Büro)

Laptop Proband: Lime H-off Nachbefragung laden → ausfüllen lassen

Währenddessen

- ggf. Datensicherung starten
- mit Lime FOT Interview beginnen (Übertragung → Dok 6_VL-Protokoll)

Interview

Laptop VL: Lime FOT Interview laden → VP & Termin A-Off angeben
→ Notizen aus Dok 6_VL-Protokoll Termin A-On **&** Termin A-Off übertragen
→ als Leitfaden nutzen, direkt ausfüllen, selbst sortieren, in welche Felder Antworten einzuordnen sind.

Abschied

Auszahlung: Dok 7_Aufwandsentschaedigung-Bestaetigung
→ i.d.R. 220,- €; falls Anforderungen an Fahraufgabe nicht erfüllt (nur Plausibilitätscheck), dann nur 120,- € (70,- € Tankgeld + 50,-€ Aufwand) oder anteilsmäßig mehr, falls späterer Abbruch

Bei Interesse: Dok 8_Einverstaendniserklaerung VP-Datenbank

Ggf. Privatfahrzeug abholen

Nachbereitung:

Dokumente

- Führerschein vernichten
- Dok 3_Uebergabeprotokoll abheften
- Dok 6_VL-Protokoll abheften
- Dok 4_Fahrtenbuch_Proband abheften
- Dok 7_Aufwandsentschaedigung-Bestaetigung abheften
- Dok 8_Einverstaendniserklaerung VP-Datenbank abheften

Desinfizieren von Tischen, Stiften und Instruktionen

H-on Fahrzeug

- Datensicherung
- Messequipment-Check
- Fahrtenbuch prüfen + ggf. ergänzen (km-Zahl, Probandentest VP-Nr)
- Desinfizieren
- Tanken/Laden
- Dok 2_H-on_MARKE_Handbuch noch im Fahrzeug?
- Wasserflasche im Fahrzeug?

H-off Fahrzeug

- Datensicherung
- Messequipment-Check
- Fahrtenbuch prüfen + ggf. ergänzen (km-Zahl, Probandentest VP-Nr)
- Desinfizieren
- Tanken/Laden
- Dok 2_H-off_MARKE_Handbuch noch im Fahrzeug?
- Wasserflasche im Fahrzeug?

0_Leitfaden_B

Übergabe Fahrzeug Gruppe B

Legende:

Gerader, normaler Text: Hinweise

Gerader, fett markierter Text: Ortswechsel/Ortsangabe

Kursiver, Text: Sprechteile

Gelbe Markierung: Verweis auf andere Dokumente

Grüner Text: Verweis auf LimeSurvey Umfrage

Vorbereitung: kurz zuvor

- H-off Fahrzeug
  - In Innenhof stellen
  - Fahrtenbuch_OEM bereitlegen und vorausfüllen: Probandentest + VP-Nummer; km-Beginn; Datum
  - Dok 2_H-off_MARKE_Handbuch bereitlegen (ist im Handschuhfach)
  - Dok 10_Leitfaden_Sicherheitsfahrer bereitlegen (ist im Handschuhfach)
- Im Besprechungsraum/0330 (Niklas'Büro)
  - Allgemeines
    - Spender zur Handdesinfektion + Tücher für Flächendesinfektion
    - Selbsttests
    - Stift (2x, VL und Proband)
    - Getränke + Süßigkeiten
  - Versuchsleiter
    - Dok 0_Leitfaden_B
    - Dok 7_Aufwandsentschaedigung-Bestaetigung
    - Dok 8_Einverstaendniserklärung VP-Datenbank
    - Laptop für Lime FOT Interview (i.d.R. eigener Laptop)
    - Fahrzeugschlüssel + Schrankentransponder
    - Klemmbrett VL
      - Dok 5_H-off_InstruktionL2
      - Dok 6_VL-Protokoll
  - Proband
    - Laptop für Lime H-off Nachbefragung
- Geplante Route auf Staus/Unfälle checken

0_Leitfaden_B

Abholen und Begrüßen

Treffpunkt Probanden: Fahrstühle Hof 3, EG

Proband begrüßen und **zu Besprechungsraum/0330 (Niklas' Büro) führen (oder direkt zum Fahrzeug)**

Im Besprechungsraum/0330 (Niklas'Büro)

> *„Herzlich Willkommen zu der Studie „Automatisiertes Fahren im Straßenverkehr". Zunächst einmal vielen Dank für Ihre Teilnahme an der Studie und dass Sie sich heute Zeit genommen haben. In dem heutigen Termin werden Sie das System Teilautomation L2 in einem unserer Versuchsfahrzeuge erleben. Zunächst werde ich Sie in das* ***Fahrzeug einweisen****. In einer gemeinsamen* ***Trainingsfahrt*** *werde ich Ihnen alle wichtigen Funktionen zeigen. Danach folgt eine* ***Versuchsfahrt****, die ca. 45min – 1 Std. dauert. Im Anschluss an die Fahrt bitten wir Sie, Ihre Eindrücke in einem* ***Fragebogen und dann in einem kurzen Interview*** *zu schildern. Der Termin dauert etwa zweieinhalb Stunden.*
>
> *Für den heutigen Termin gilt 2G+. Daher würde ich Sie nun bitten, mir einen entsprechenden Impf- bzw. Testnachweis zu zeigen."*

Impf- /Testnachweis checken

Führerschein checken

Getränke & Snacks (Niklas' Büro) anbieten

Auf Toilette hinweisen

0_Leitfaden_B

Fahrzeugerprobung im Stillstand

Erklärungen im Fahrzeug

kurz halten, da für h-off eher nicht relevant,
die Fahrer sollten aber kurz Zeit haben, sich mit dem Fahrzeug und auch mit dem Kombi vertraut zu machen!

Grundbedienung

- Sitz/Lenkrad/Spiegel einstellen
- Motor anschalten
- Automatik erklären (D für Drive, R für Rückwärts, P für Handbremse, ggf. zusätzliches)
- Licht: An- und Ausschalten (falls nicht automatisiert)
- Scheibenwischer (falls nicht automatisiert)

Anzeigeelemente (Non-Assistenz)

- Kombi:
  - Kilometerstand (→ hier checken, wie viel schon gefahren + Fahrtenbuch)
  - Uhrzeit (→ hier checken wie lange schon gefahren + Fahrtenbuch)
  - Tank-/Ladeanzeige & Reichweitenanzeige
  - Weitere relevante Icons/Anzeigeelemente (Non-Assistenz)
  - Hinweis: keine dauerhaften Änderungen am Layout
- Head-up Display
  - Relevante Icons/Anzeigeelemente (Non-Assistenz)
- CID: Generelle Struktur + Bedienung (sehr kurzhalten, Hinweis auf Navi und Mediasystem inkl. Radio)
- Ggf. weitere relevante Anzeigeelemente / Bedienelemente (z.B. Touchscreen, iDrive, etc.)

Funktionen (Non-Assistenz)

- Klimaanlage: Allgemeintemperatur und Scheibenlüftung zeigen
- CID: Hinweis: Alles darf geändert werden, **außer Einstellungen Fahrerassistenz**
  - Ausnahmen: s. Dok 2_H-off_MARKE_Handbuch #2
- Hinweis: Im Zweifelsfall Rücksprache mit Studienteam

Allgemeines

- Kamerapositionen zeigen (ggf. Hinweis auf Sonnenblende)
- Trigger-Knopf → Dok 2_H-off_MARKE_Handbuch #1 → beide Varianten ausprobieren lassen

Probefahrt auf Gelände

1-2 Male im Kreis fahren, dann Stop vor Lehrstuhl

0_Leitfaden_B

Instruktion L2H-off

„Ich erkläre Ihnen nun das Automationssystem, das Sie in der Studie überwiegend nutzen werden. Bitte lesen Sie sich dafür zunächst diese allgemeine Beschreibung zu Teilautomation L2 durch."

→ Dok 5_H-off_InstruktionL2

Fragen klären

Anleitung L2

→ Dok 2_H-off_MARKE_Handbuch #3f.

„In Ihrem Fahrzeug wird das L2 System „XY" genannt. Im Folgenden zeige ich Ihnen die wichtigsten Bedienschritte für das System. Bitte beachten Sie, dass dies nur einen ersten Überblick darstellt. Konsultieren Sie die Bedienungsanleitung im Handschuhfach für wichtige weitere Hinweise zu Anzeige, Bedienung und Grenzen des Automationssystems."

Hinweise aktives L2:

- Icon/Anzeige und Farbe beschreiben
- Wenn L2 aktiv ist, übernimmt es die Quer- und die Längsführung. Es kann sein, dass aufgrund von Verkehrs- oder Umweltbedingungen Teile des Systems deaktiviert werden. Dann verändert sich das markierte Icon in der Farbe oder in der Form.
  Beachten Sie, dass Sie als Fahrer immer für die Fahraufgabe verantwortlich sind.
  → ggf. Änderungen des Icons/der Anzeige bei (Teil-)Deaktivierung kurz beschreiben

Bedienschritte (durchgehen und auf Tasten hinweisen, aber nicht ausprobieren lassen)

- Aktivieren
- Deaktivieren
- Geschwindigkeit einstellen
- Abstand einstellen
- Spurwechsel durchführen

Falls Tasten ausprobiert werden → Zündung an & aus **und** kontrollieren, ob Systemstatus unverändert (z.B. ACC und/oder Spurhalteassistent noch im Standby aktiv ist)

„Haben Sie hierzu aktuell noch Fragen? Wir werden alle Anzeige- und Bedienelement gleich in einer Trainingsfahrt gemeinsam austesten."

0_Leitfaden_B

Fahrzeugerprobung auf Straße

Vor der Fahrt: **Individuelle Kalibrierung** der (auf Augen gerichteten) Kameras

- VP sollte ab jetzt Maske absetzen
- Fahrposition einnehmen → auf Punkte blicken
- **Trigger 2x drücken** (Markierung Kalibrierung in den Daten)

1. Entspannt Geradeaus
2. Rückspiegel
3. CID
4. Mittelkonsole unten, z.B. Drehdrücksteller
5. Kombi
6. Lenkradtasten links
7. Linker Seitenspiegel
8. Rechter Seitenspiegel
9. Ggf. Head-up Display

Start Trainingsfahrt (Fahrt: A9 nach 66 Pfaffenhofen P&R)

- Beginn Trainingsfahrt: A9 Richtung Norden; insgesamt ~15-20 Min
- L2 ausprobieren
- Anweisungen Versuchsleiter jederzeit Folge leisten
- Informieren, falls Unwohlsein oder Pause benötigt
- Proband frühzeitig navigieren und auf Fußgänger hinweisen

Start Autobahn

Protokoll führen→ Dok 6_VL-Protokoll

- Proband: Beobachtungen, z.B. nimmt Hände sofort weg, Misuse, etc.)
- Auto: Beobachtungen, seltsames Verhalten, z.B. Eyes Off Warning, obwohl Blick auf Straße, ...

Zu beachten

- Sehr individuell nach Bedürfnissen gestalten, ausreichend Zeit lassen, ggf. Pausen einlegen
- Auf Fehlverhalten hinweisen
- Auf Verantwortung des Fahrers, Sicherheit als höchste Prio & Maximalgeschwindigkeit hinweisen
- Fragen klären, proaktiv nach Fragen fragen
- Ggf. Hinweis: Erkennung aktuelle Geschwindigkeit nicht immer zuverlässig

Manöver

- Manuelle Manöver: Bremsen, Beschleunigen, Spurwechsel
- L2
    - Aktivierung
    - Deaktivierung (verschiedene Wege)
    - Zielgeschwindigkeit anpassen
    - Abstand regeln (nicht bei iX, da dort über CID)
    - Spurwechsel durchführen
- Trigger-Knopf
    - Drücken & ausprobieren (sprechen lassen) → wichtig: *„auch wenn ungewohnt, bitte verwenden, da hilfreich für uns"*

0_Leitfaden_B

Ende

- **Trigger 2x drücken** (während Zündung aktiv) → dies markiert das Ende der Trainingsfahrt & bedeutet, dass als nächstes die Versuchsfahrt folgt

Versuchsfahrt

Dauer >45min

Markierung Beginn der Fahrt (auf Autobahn) durch **3x Druck des Triggers (<1s)** – ACHTUNG VL macht das – nicht der Proband

Route 1: A9: Von 66-Pfaffenhofen nach 63-Manching, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~17 min sein; falls deutlich kürzer, dann weiterfahren nach 62-Ingolstadt-Süd oder 61-Ingolstadt-Nord (statt 63-Manching; bei langem Baustellenabschnitt empfehlenswert).

Alternativen:

Route 2: A9+A92: Von 66-Pfaffenhofen nach 10-Moosburg-Süd, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~28min sein; falls deutlich kürzer, dann weiterfahren nach 12-Landshut-West (statt 10-Moosburg-Süd).

Route 3: A9+A92+A99+A8: Von 66-Pfaffenhofen nach 77-Sulzemoos, wenden und zurück nach 70-Garching-Nord → sollten ab 66-Pfaffenhofen ~33min sein; falls deutlich kürzer, dann weiterfahren nach 76-Odelzhausen (statt 77-Sulzemoos).

Hinweise:

- während der Fahrt Auffälligkeiten über Trigger kommentieren, nicht direkt mit VL kommunizieren, da lediglich Funktion Sicherheitsfahrer
- bei Baustellen: bei An- und Abfahrt in Baustelle Hände ans Lenkrad + rechte Spur

Bei Fehlverhalten des Probanden oder sicherheitskritischen Situationen
→ Dok 10_Leitfaden_Sicherheitsfahrer

Protokoll führen → Dok 6_VL-Protokoll

- Proband: Beobachtungen, z.B. nimmt Hände sofort weg, Misuse, etc.
- Auto: Beobachtungen, seltsames Verhalten, z.B. Eyes Off Warning, obwohl Blick auf Straße, ...
- Weiteres: Auffälligkeiten, Probleme, Kommentare

Schranke

Ggf. aussteigen & selbst öffnen

Fahrzeug parken

Im Hof

Fahrtenbuch ausfüllen

Toilette, Snacks & Getränke vor Befragung anbieten

0_Leitfaden_B

Nachbefragung

Im Besprechungsraum/0330 (Niklas' Büro)

Laptop Proband: Lime H-off Nachbefragung laden → ausfüllen lassen

Währenddessen

- ggf. Datensicherung starten
- mit Lime FOT Interview beginnen (Übertragung → Dok 6_VL-Protokoll)

Interview

Laptop VL: Lime FOT Interview laden → VP & Termin B angeben
→ Notizen aus Dok 6_VL-Protokoll übertragen
→ als Leitfaden nutzen, direkt ausfüllen, selbst sortieren, in welche Felder Antworten einzuordnen sind.

Abschied

Auszahlung: Dok 7_Aufwandsentschaedigung-Bestaetigung
→ i.d.R. 50,- €

Bei Interesse: Dok 8_Einverstaendniserklaerung VP-Datenbank

Nachbereitung:

Dokumente

- Dok 6_VL-Protokoll abheften
- Dok 7_Aufwandsentschaedigung-Bestaetigung abheften
- Dok 8_Einverstaendniserklaerung VP-Datenbank abheften

Desinfizieren von Tischen, Stiften und Instruktionen

H-off Fahrzeug

- Datensicherung
- Messequipment-Check
- Fahrtenbuch prüfen + ggf. ergänzen (km-Zahl, Probandentest VP-Nr)
- Desinfizieren
- Tanken/Laden
- Dok 2_H-off_MARKE_Handbuch noch im Fahrzeug?
- Wasserflasche im Fahrzeug?

5 Evaluation of Hypotheses on System Design

The centerpiece of experimentally controlled data collections on specific system design hypotheses of this project were four driving simulator studies including 60-80 participants each (see Figure 5-1). Each of the four studies focused on a specific system design hypothesis or potential challenge for L2H-off functions, as derived from the knowledge basis build by the first three subprojects (briefly described in Chapter 5.1). The four independent data collections were aligned regarding, e.g., the general system design for implementation in the driving simulator, metrics for data analysis and instructions with the goal to aggregate results across all of the four studies to answer challenges and questions as comprehensively as possible.

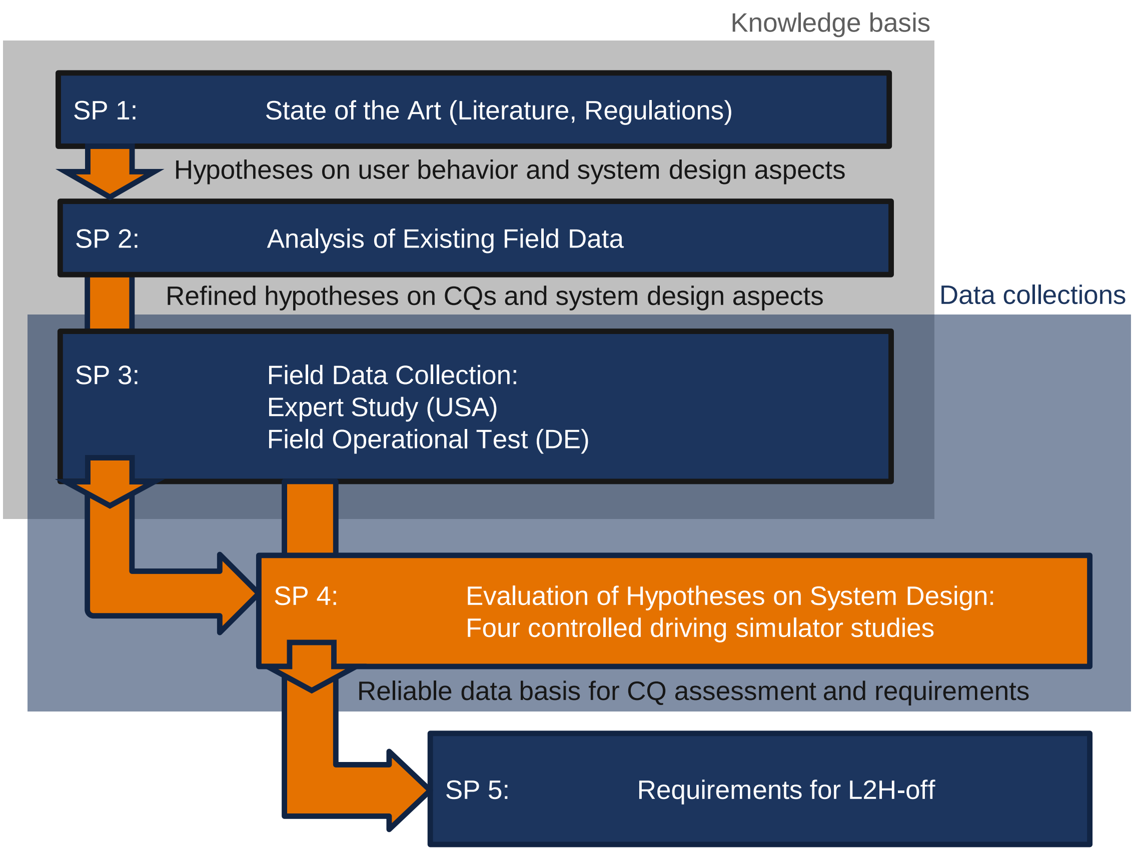


Figure 5-1: Overview on the five subprojects and the role of SP 4 within the project.

5.1 Procedure on Selecting Study Foci and Design Hypotheses for Testing

Documentation by P. Dautzenberg (Institut für Kraftfahrzeuge, RWTH Aachen University)

Based on the state of the art analyses conducted in previous work packages, four controlled experimental studies within this project (SP4) provide the opportunity to address the most relevant research questions and study foci in order to fill gaps in the current state of the art and to derive specific system design hypotheses with respect to L2H-off systems. In this respect, the five challenges and questions (CQs; 1. Hands-off = mind-off?; 2. Prolonged transition times; 3. Foreseeable misuse; 4. Mode confusion; 5. Safety level) motivating this project are considered and addressed across multiple studies. The project as well as the respective user studies aim at providing and increasing insights into challenges potentially related to L2H-off systems. Furthermore, findings and impressions gathered in previous work packages (see e.g., SP1 and SP3) establish the basis from which main experimental foci have been derived, based on the main design elements driver monitoring, HMI and function (Section 2.5).

The process for identifying relevant research questions and study foci involved two key steps:

1. First, a **review of existing regulations** for current L2H-on with a focus on UN ECE R79 (2018) has been conducted. Additionally, UN ECE R157 (2021) has been analyzed with a defined focus on DMS criteria of L3 regulations, as these already consider supplementary criteria next to hands-on detection. Thereby, it has been assessed which passages/requirements could either be adopted for L2H-off systems or would need to be adapted and/or supplemented and from which passages a delimitation should be created.

2. Subsequently, the identified adaptation of regulation needs have been considered as well as the expected driver behavior when using L2 against the background of the five **challenges and questions (CQ)** to consider the need to address further research questions apart from those targeting adaptation needs of existing functionalities. This has been done while keeping in mind potential compensation strategies from state of the art solutions and scientific literature.

The results of these considerations have been used to derive a prioritization of CQs for the project and for the experimental studies specifically.

5.1.1 Results

In accordance with the two-step approach to derive relevant research questions, the results are described below according to these two steps.

5.1.1.1 Review of existing regulations

While reviewing the UN ECE R79, one subchapter (5.6.2.2.5) has been identified that may need to be adapted with regard to L2H-off systems. Central subject of this subsection is the detection of “hands-on steering control” as a mandatory element of L2H-on systems. Since monitoring “hands-on steering control” while using a H-off systems is not feasible, an alternative driver monitoring approach other than “hands-on steering control” is needed that determines whether the driver is sufficiently attentive and available to take over the driving task. Subsequently, the question arises which criteria may be used to determine whether the driver is sufficiently attentive and available.

Regulation UN ECE R157 is considered for a clear delimitation from SAE Level 3 systems. UN ECE R157 focusses on automated lane keeping systems (ALKS) that are operational up to 130 km/h. Since such systems perform the complete driving task within the respective ODD, the driver only serves as a fallback for the system when prompted and not as a supervisor of the system with continued responsibility for the driving task. Thus, this regulation is based on different assumptions regarding necessary driver interventions. However, UN ECE R157 provides the reference to “eyes on road” as a DMS criterion that may be used with hands-free systems. In contrast to UN ECE R79 with its sole focus on hands-on detection, UN ECE R157 defines multiple characteristics to determine whether the driver is present, available and attentive (subchapter 6.3.1.1.), one being the following:

“The driver is deemed to be attentive when at least one of the following criteria is met:

a) Driver gaze direction is confirmed as primarily looking at the road ahead;
b) Driver gaze direction is being confirmed as looking at the rear-view mirrors; or,
c) Driver head movement is confirmed as primarily directed towards the driving task.”

Accordingly, one conclusion from the analysis of regulation UN ECE R157 is that a DMS that determines "eyes-on road" could replace the L2H-on specific "hands-on steering control“ in cases where hands-free monitoring is admitted.

Another passage of subsection (5.6.2.2.5) in UN ECE R79 addresses the warning cascade employed whenever hands-on steering control cannot be confirmed. Based on the assumption that an alternative DMS is monitoring the driver for "eyes on road", there could be a need for adaptation with regard to the number of stages and their timing.

5.1.1.2 Consideration of the CQs

As described in the procedure (see 5.1), the identified adaptation needs listed in the previous chapter were subsequently considered against the background of the five challenges and questions (CQ) to determine the need to address further research questions within the project and the respective user studies.

The main adaptation derived from the analysis of existing regulations concerns the design of DMS and warning cascades for L2H-off. Based on the expected user behavior, it needs to be evaluated whether the five CQs could be sufficiently addressed by adaptations solely to the DMS and warning cascades. The evaluation is based on publications collected for the state of the art analysis (SP1) and conclusions from the expert study in the USA (SP3; comparative assessment of series-production L2H-on and L2H-off systems).

Based on both literature and expert assessment, a feasible DMS is considered as a potential advance to address most of the CQs, further stressing this aspect as a relevant study focus. However, potentially arising mode confusion (CQ4) might not be sufficiently addressed by DMS design alone. In fact, there is limited research on the occurrence of mode confusion when switching between different L2 function designs (H-on and H-off). Therefore, more input is needed to assess whether hands-free L2 systems potentially increase mode confusion and if so which countermeasures can be implemented on the system design side to counteract this. Findings in related fields, such as the automation expectation mismatch, indicate that directing users' attention to the right places in the right time might not always be sufficient to guarantee a safe interaction with L2 automation (e.g., Tivesten, Broo, & Ljung Aust, 2022; experience from SP3.4). An adapted DMS or warning cascade seems therefore not sufficient to address the potential problem of mode confusion in all usage situations. These findings show that in addition to the previously identified research foci (DMS and attentiveness alert design), mode confusion should also be addressed to fill existing gaps and to derive respective system design hypotheses.

5.1.2 Selected Research Questions and Design Hypotheses

Based on the two previous analysis steps, it becomes apparent that especially with regard to the design of the DMS, the attention alert/warning cascade and mode awareness (research) questions are pending, the answers to which could be relevant for the conclusion on potential challenges related to hands-free driving and for requirements of L2H-off systems.

Figure 5-1 gives an overview of the foci to be addressed within the experimental studies based on the discussion of findings from the state of the art and the input from the Scientific Advisory Board. Furthermore, in this figure the four studies planned within this project are assigned to the respective foci.

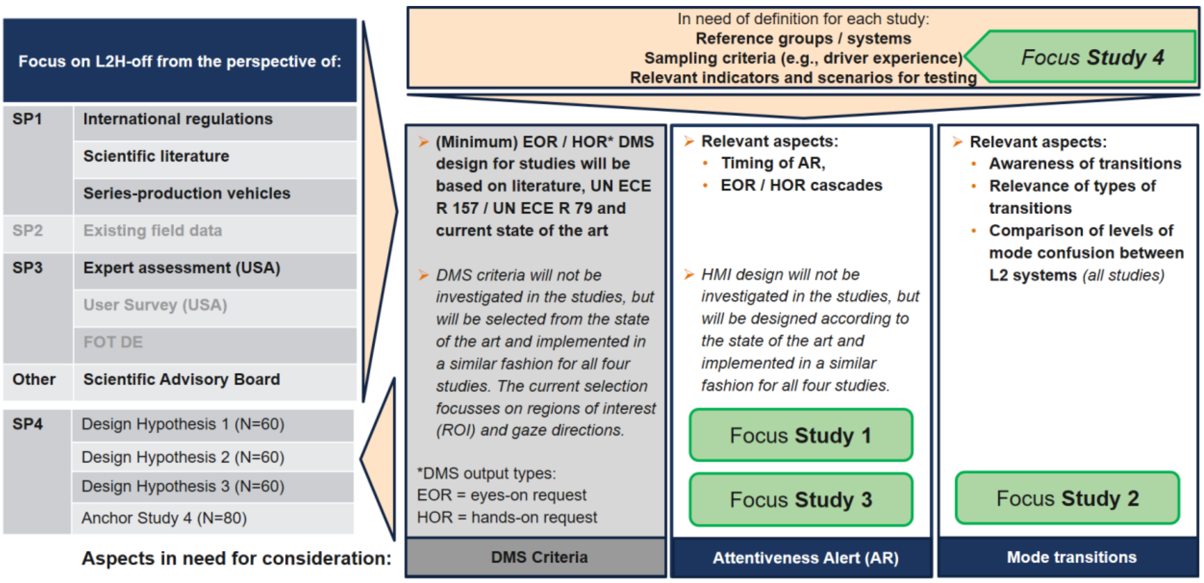


Figure 5-1: Research foci to be addressed in the four controlled studies. SP-Aspects printed bold were used primarily for deriving the research foci.

Table 5-1 provides an overview of the research questions addressed in the respective studies. The study-specific assumptions and derivation regarding potential design hypotheses are presented in the respective subsections of the studies.

Table 5-1: Overview of the four user studies including their *main* research questions. The crosses indicate which challenges/questions (CQs) are in the focus of each study. Whenever possible, all CQs are taken into account to some degree by each of the studies.

Study	Focused CQ(s)					Research Question(s)
	1	2	3	4	5	
Study 1	x					1. Are there differences between L2H-on systems with hands-on detection (HOD) and L2H-off systems with eyes-on detection (EOD) with regard to attention and user behavior? 2. Are there differences between manual driving and L2 automated driving with regard to attention and user behavior?
Study 2				x		1. Are there differences between L2H-on and L2H-off systems regarding mode confusion? 2. Are there differences between clear-cut transitions (L0 - L2, and vice versa) and more complex multi-step transitions (L0 – L1 – L2 and vice versa) regarding mode confusion?
Study 3	x					1. Are there differences between L2H-off systems with eyes-on requests (EOR) and L2H-on systems with Hands-on request (HOR) with regards to attention and user behavior? 2. Are there differences between L2H-off systems with differently timed eyes-on requests (EOR) with regards to attention and user behavior?
Study 4		x			x	1. Are L2H-off and L2H-on systems comparable with regard to… ... driver-detected intervention needs (controllability of system failures)? ... system-indicated transitions (TOR timing)? 2. Can findings be generalized over different samples? ... regarding prior expertise with L2 systems (none vs. EOR vs. HOR)? ... regarding cultural differences (US vs. DE sample)?

5.1.3 Discussion on studies‘ comparability

The following aspects have been discussed as part of the experimental design of the studies within this project, either to keep them comparable across studies or to vary them intentionally between studies:

- HMI design (for displaying warnings, requests or other information),
- Prototypical DMS (underlying criteria/parameters and areas of interest) for experimental studies,
- Warning cascades (timing and modalities of DMS alerts and take-over requests),
- Number of assistance modes and logic of transitions between them (i.e., L0, L1, L2),
- Scenarios to be experienced within the simulation (transition types, parametrization, anticipation possibilities, system behavior),
- Data (subjective and objective) to be collected and analyzed (vehicle-related metrics, gaze data, subjective (standardized) questionnaires and/or single items),
- Reference group(s) (L2H-off vs. L2H-on vs. L0) and
- Sample characteristics (e.g. ADAS experience, age, gender).

A detailed description of the operationalization of each of these aspects can be found in the respective study chapters.

5.1.4 References

United Nations Economic Commission for Europe (2021). *Automated Lane Keeping Systems (ALKS).* UN Regulation No. 157, pp. 75-137. Retrieved from https://unece.org/transport/documents/2021/03/standards/un-regulation-no-157-automated-lane-keeping-systems-alks

United Nations Economic Commission for Europe (2018). *Uniform provisions concerning the approval of vehicles with regard to steering equipment.* UN Regulation No. 79 Retrieved from https://unece.org/transport/documents/2021/02/standards/un-regulation-no-79-rev4-amend3

Tivesten, E., Broo, V., & Ljung Aust, M. (2022). *The Influence of Alcohol and Automation on Drivers' Visual Behavior During Test Track Driving*. Available at SSRN 4165931.

5.2 Experimental Study 1

Documentation by A. Feierle, D. Albers, B. Biebl, N. Grabbe, M. Hübner, T. Hecht, K. Bengler (Lehrstuhl für Ergonomie, TU München)

Experimental Study 1 focused on CQ1 “Hands-off = Mind off?”.

5.2.1 Research Questions

While L2H-on functions are equipped with a hands-on detection (HOD) and a corresponding hands-on request (HOR), L2H-off functions use an eyes-on detection (EOD) and a corresponding eyes-on requests (EOR) to avoid inattention and potential mind off. In order to investigate whether an L2H-off function shows different effects on mind off and the driver’s visual attention, an L2H-off function should be compared with an L2H-on function. Additionally, a comparison of both L2-functions with manual driving (L0) was aimed at. This aimed to investigate the following research questions in Experimental Study 1:

- RQ1: Are there differences between L2H-on with HOD and L2H-off with EOD with regard to visual attention and user behavior?
- RQ2: Are there differences between manual driving and L2 automated driving with regard to visual attention and user behavior?

5.2.2 Method

5.2.2.1 Sample

Participants were recruited using social media advertising, notices on the campus, and the participant database of the Chair of Ergonomics at the Technical University of Munich. Requirement for participation was that participants have held a driving license for at least 5 years, which was intended to prevent the results from being influenced by novice drivers. Participants received 50€ as a compensation.

A total of 60 participants took part in the experiment. Twenty participants experienced a manual drive, 20 the L2H-off function and 20 the L2H-on function. Mean age for the L2H-on group was M = 33.60 years (SD = 13.04), for the L2H-on group M = 32.60 years (SD = 16.81) and for the L0 group M = 33.60 years (SD = 13.62). Each group consisted of 7 women and 13 men. Participants of the L2H-on group had a driver’s license for M = 15.75 years (SD = 12.05), participants of the L2H-of group for M = 14.95 years (SD = 16.39) and participants of the L0 group for M = 16.10 years (SD = 13.23). The driving frequency of the participants showed similar mean values across the three groups (Figure 5-1). Participants usually drove several times a week. The highway trips were less frequent, while the L0 group and L2H-on group showed a slightly higher variance than the L2H-off group. Mileage in the last 12 months showed that participants across all three groups drove predominantly between 1 and 20,000 km; there were no drivers with over 50,000 km per year (Figure 5-2).

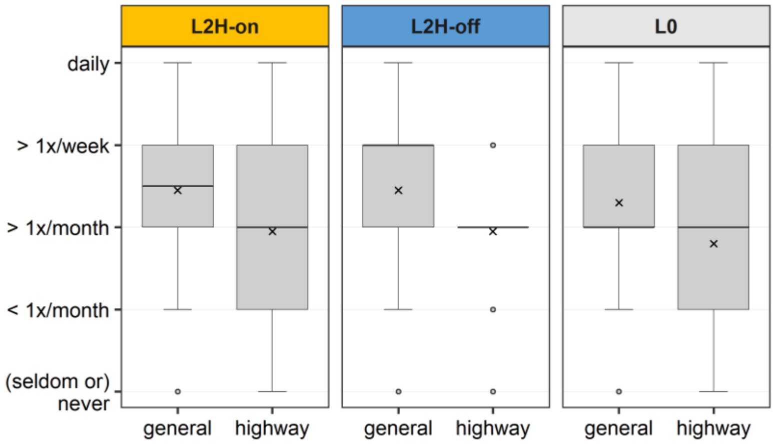


Figure 5-1: Boxplots showing the driving experience in general and on highways of the participant groups. The sample size was n = 20 for all groups.

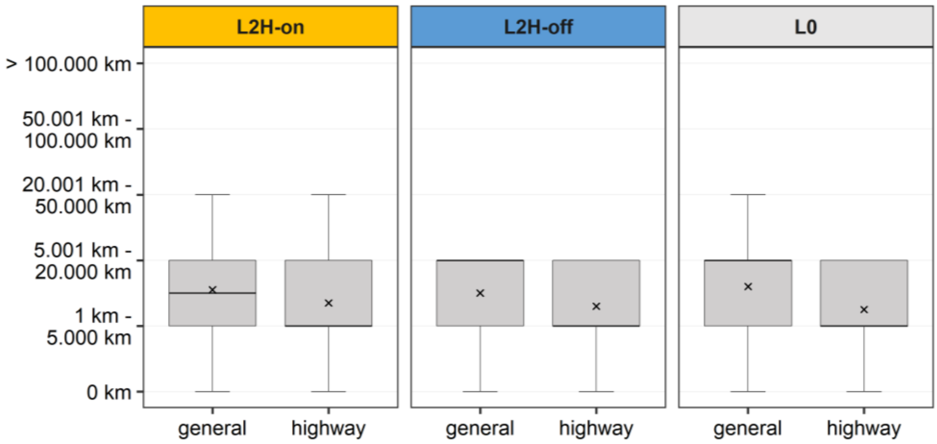


Figure 5-2: Boxplots showing the mileage during the last 12 months of the participant groups. The sample size was n = 20 for all groups.

Figure 5-3 presents the familiarity of participants with six different driver assistance systems: cruise control (CC), adaptive cruise control (ACC), lane keeping assistant (LKA), traffic jam assistant (TJA), parking assistant (PA), and L2 function (L2). On a descriptive level, the results showed similar familiarity between the different groups. CC was used regularly by many participants. In contrast, ACC, LKA, TJA, PA and L2 were used regularly by only a few participants. TJA, PA and L2 was even unknown to many. Overall, there was a high degree of variance within individual participants, but also across participants.

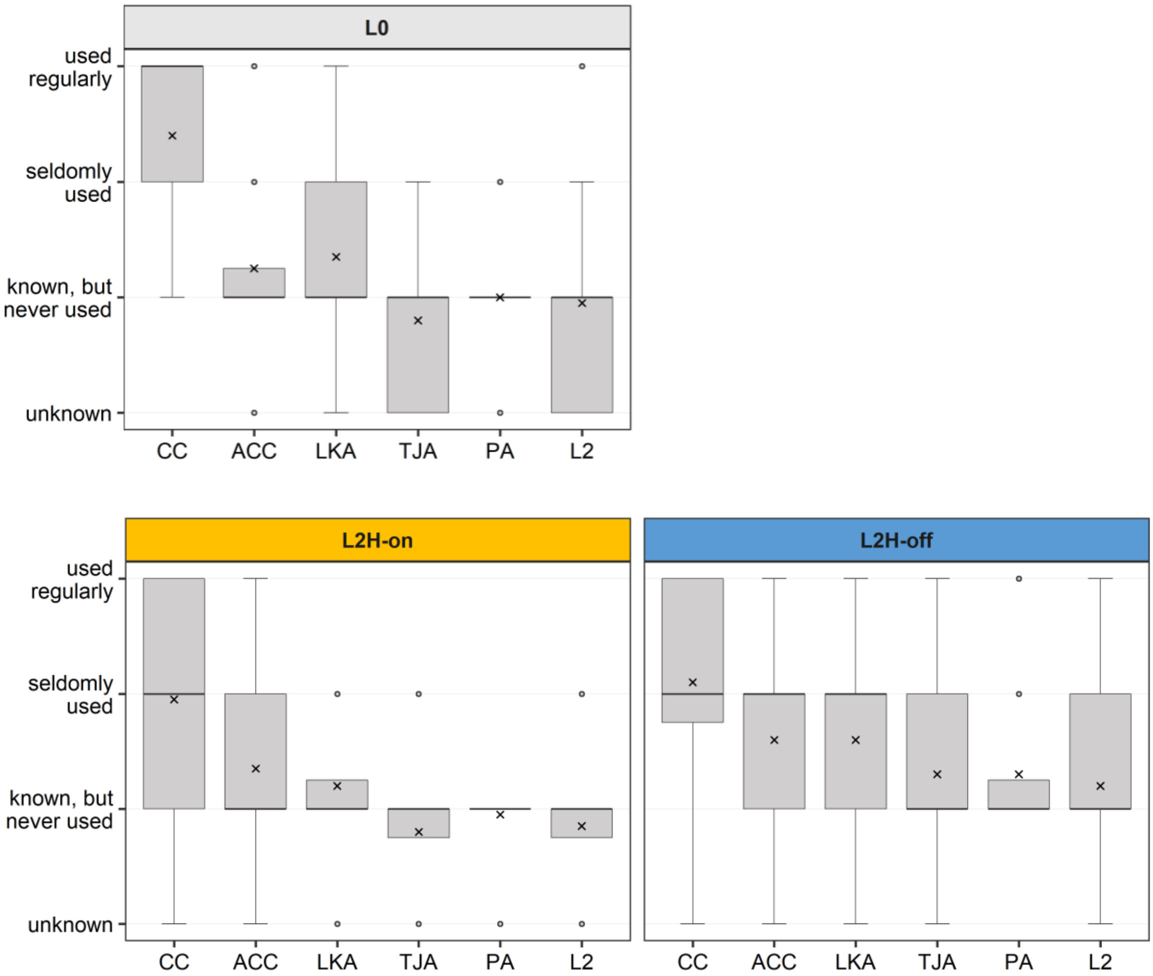


Figure 5-3: Boxplots showing the familiarity with driver assistance systems (CC: Cruise Control; ACC: Adaptive Cruise Control; LKA: Lane Keeping Assistant; TJA: Traffic Jam Assistant; PA: Parking Assistant; L2: L2 function). The sample size was n = 20 for all groups.

The affinity for technology interaction scale (Franke, Attig, & Wessel, 2019) showed similar mean values across all groups. The participants could be described as slightly technology-affine, whereby the variance at L0 was smaller than in the other two groups (see Figure 5-4). Overall, there were no outliers.

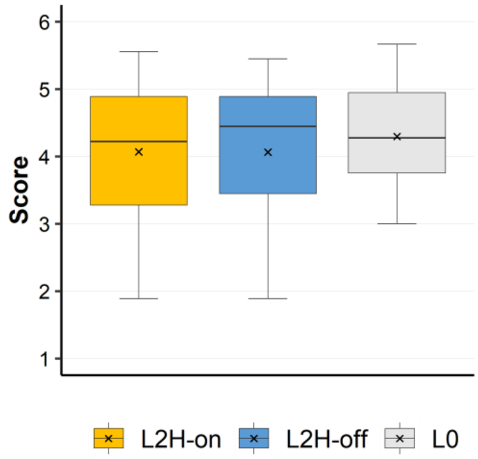


Figure 5-4: Boxplots showing results for the affinity for technology interaction scale. The sample size was *n* = 20 for all groups. Higher scores indicate a higher affinity for technology.

5.2.2.2 Experimental Design

A between-subject design was applied for Study 1 (see Figure 5-5). The function design consisting of L2H-on, L2H-off and L0 was the between-subject factor leading to three experimental groups. In the L2H-on group, participants experienced a drive in a vehicle that had an L2H-on function including hands-on detection and a HOR. Participants in the L2H-off group experienced a drive with an L2H-off function equipped with a DMS based primarily on visual attention for EOR as well as hands-on monitoring for HOR in subsequent DMS warning stages. The L0 group made the drive manually without the support of assistance systems for longitudinal and lateral guidance. The experiment was approved by the Ethics Committee of the Technical University of Munich.

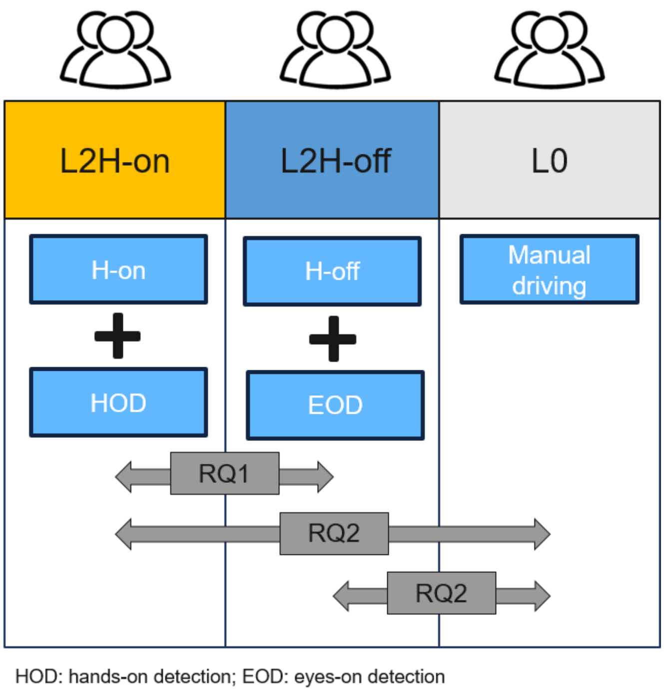


Figure 5-5: Experimental design for answering the research questions.

5.2.2.3 Apparatus

Study 1 was conducted at the static driving simulator of the Chair of Ergonomics at the Technical University of Munich (Figure 5-6). The driving simulator was based on a BMW E64 mock-up and a 6-channel projection system. Three projectors provided a front view of 180° and three additional projectors visualized the view of the rear and side mirrors. A sound system provided engine and environmental sounds. SILAB 6.5 from the Würzburg Institute for Traffic Sciences has been used as driving simulation software. The driving simulation, data recording and the visualization of the projectors run with 60 Hz. The remote eye-tracking system Smart Eye was used to record gaze behavior. Both L2 functions were based on a clear-cut principle (see Study 2). Participants could activate the L2 function via a button on the steering. Deactivation of the L2 function was possible using the same button, by braking, by steering or by accelerating. The set speed of the L2 function was 120 km/h, while participants were not able to adjust this set speed. L1 (ACC only) was not available during the driving simulation session. All three groups had in common that the vehicle was equipped with an additional emergency brake assistant.

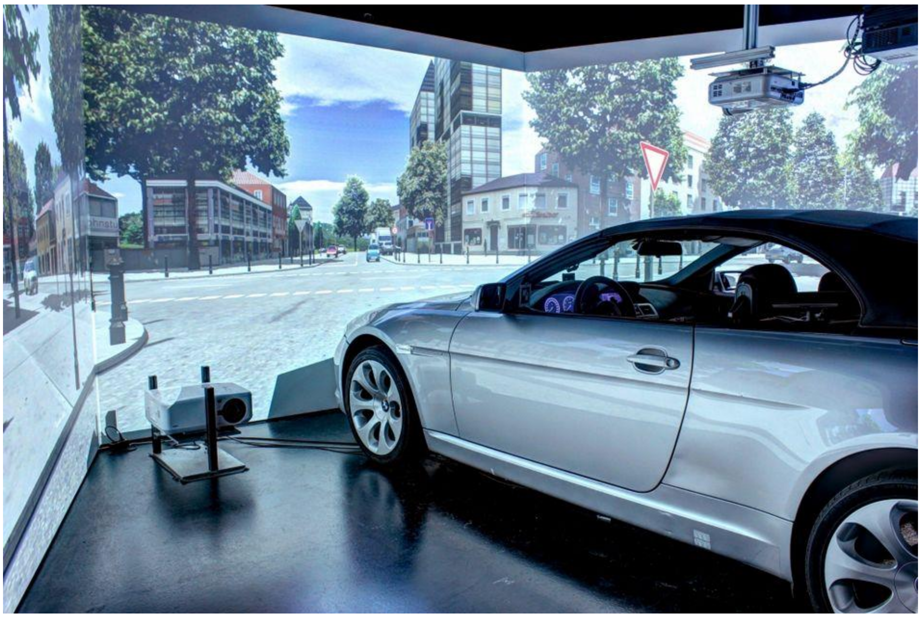


Figure 5-6: Static driving simulator of the Chair of Ergonomics at the Technical University of Munich.

DMS and HMI

A three-stage warning cascade for the attentiveness alert of the L2H-off and L2H-on function has been chosen (e.g., Kurpiers, Lechner, and Raisch (2019), Blanco et al. (2015)) which was presented to the participants in German. An overview of the warning cascades is presented in Figure 5-7.

L2H-off	**Warning Cascade Design**					
	1. Warning Stage		**2. Warning Stage**		**3. Warning Stage**	
HMI	**5 s**	Eyes-on request (visual + acoustic alert) **“Aufmerksam bleiben!“**	**8 s**	Hands-on request (visual + acoustic alert) **“Aufmerksam bleiben, Hände ans Lenkrad!“**	**13 s**	DMS direct control request (visual + acoustic alert; braking to standstill) **“Fahrzeug bremst. Übernehmen!"**
Driver’s task		Eyes on road		Eyes on road + hands on wheel		Direct control of driving task

L2H-on	**Warning Cascade Design**					
	1. Warning Stage		**2. Warning Stage**		**3. Warning Stage**	
HMI	**15 s**	Hands-on request (visual alert + acoustic alert) **„Hände ans Lenkrad!“**	**20 s**	Hands-on request (visual + acoustic alert) **„Hände ans Lenkrad!”**	**25 s**	DMS direct control request (visual + acoustic alert; braking to standstill) **“Fahrzeug bremst. Übernehmen!"**
Driver’s task		Hands on wheel		Hands on wheel		Direct control of driving task

Figure 5-7: Attentiveness alert and the corresponding warning stages for L2H-off and L2H-on.

L2H-off:

For the L2H-off function, an EOR was issued if the driver did not look to the road for 5 s continuously. This time period was consistent with vehicles in the FOT and the US study. Also, a similar time period has already been used in Kurpiers et al. (2019). During the first warning stage, a "Stay Attentive" text box appeared on the instrument cluster, the LEDs on the steering wheel flashed yellow, and a warning tone appeared. The warning cascade was terminated when the driver looked back to the road for at least 0.2 s. Looking to the road was defined as a glance to the windshield. The rearview mirror was not defined as part of the windshield. The second warning stage after additional 3 s resulted in an attention request and hands-on request in the instrument cluster, a red flashing of the LEDs on the steering wheel, and an acoustic warning. To end the warning, the driver had to put at least one hand back on the steering wheel and watch the road. During the third warning stage after additional 5 s, compared to warning stage 2, there was an additional DMS direct control request in the instrument cluster and the vehicle started to brake to a full stop. After taking over direct control, the driver drove again manually and had to activate the L2 function again.

L2H-on:

The first warning stage of the L2H-on function occurred after 15 s when no hands were detected on the steering wheel according to UN ECE R79 (2018). Here, a text box with "Hands on the steering wheel!" and an additional symbol appeared in the instrument cluster, as well as a yellow flashing of the LEDs on the steering wheel and an auditory warning. The warning was terminated when the driver had one of his two hands on the steering wheel again. Hands-on detection was realized via a steering wheel with capacitive sensors. The second warning stage after additional 5 s resulted in a second hands-on request in the instrument cluster, a red flashing of the LEDs on the steering wheel, and an auditory warning. Again, the warning was terminated when the driver had one of his hands back on the steering wheel. The third warning stage after additional 5 s was similar to the L2H-off function. There was an additional DMS direct control request in the instrument cluster and the vehicle started to brake to a full stop. After taking over direct control, the driver drove again manually and had to activate the L2 function again.

The HMI in the instrument cluster was designed simple to minimize visual distraction to the instrument cluster due to the HMI design (Kraft, Naujoks, Wörle, & Neukum, 2018). The HMI was divided into three areas. The left area displayed the current speed, the speed limit, and the set speed of the active L2 function. Below this were the two icons which collectively represented the active L2 function. In the right area, the current gear was displayed. In the middle area, warnings and requests were presented according to the warning levels of the warning cascade. The comparison between the display of L2H-on and L2H-off is shown in Figure 5-8. The HMI of the L0 group was consistently designed, also showed the current speed limit, but no icons for the L2 function.

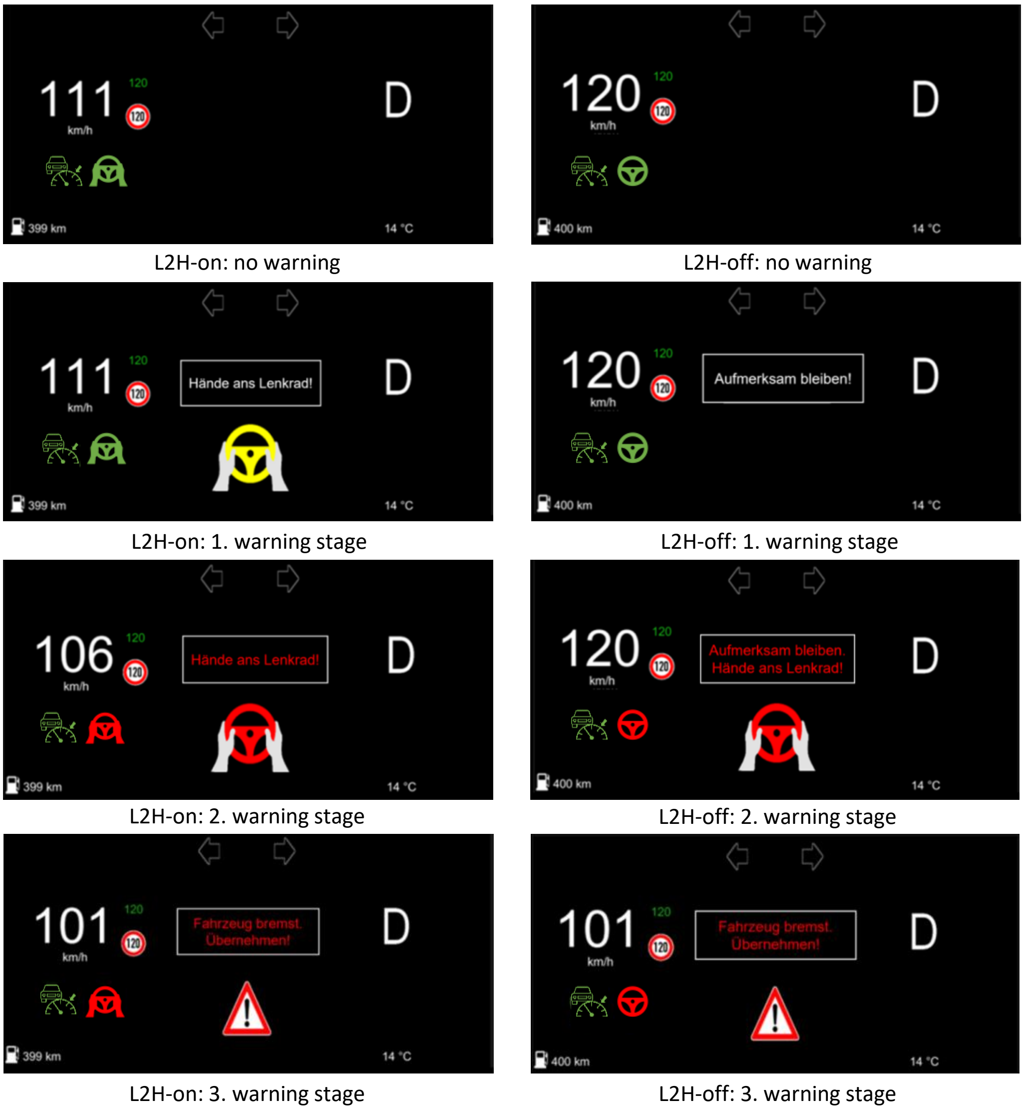


Figure 5-8: HMI design of the L2H-on function (left) and the L2H-off function (right) for L2-driving without any warnings and the three warning stages

Non-Driving Related Task

To investigate whether different DMS are equally suited to enable the driver to anticipate and handle possible system limits, visual distraction of the driver needs to be influenced in a way that is reasonable and admissible for the driving context. A well interruptible visual-motoric NDRT was designed to be in line with the requirements for vehicle guidance while increasing variance in visual attention of the driver. Therefore, a 9.6-inch tablet was mounted in front of the central information display. During the NDRT, participants had to read a text with 60 words on average which was displayed for 30 s on the tablet. After 30 s or by confirming a button on the tablet, a question regarding the text was shown on the tablet. The answer consisted of a single word and had to be entered by the participants using the touch display keyboard of the tablet. Only one hand was needed. Again, 30 s were allocated for this before the next text was displayed. By instruction, it was up to the participants for how long they wanted to engage in the NDRT, depending on whether they felt safe. They were offered an incentive of 30 € if they performed particularly well on the NDRT. However, they were informed that if safety-critical situations occurred or their behavior was classified as safety-critical by the L2 function, they would not receive an incentive.

5.2.2.4 Experimental Track and Scenario Design

The driving duration was approx. 45 minutes. The highway consisted of two lanes in each direction with a Level of Service A (TRB, 2000). The participants were instructed to obey the traffic rules, such as keeping on the right lane when there is no reason to pass. During the test drive, four system limit/failure scenarios occurred that were relatively equally distributed over the experimental track (Figure 5-9). The speed limit of the experimental track was 100 or 120 km/h except for specific scenarios. The set speed of the L2 function was 100 or 120 km/h accordingly.

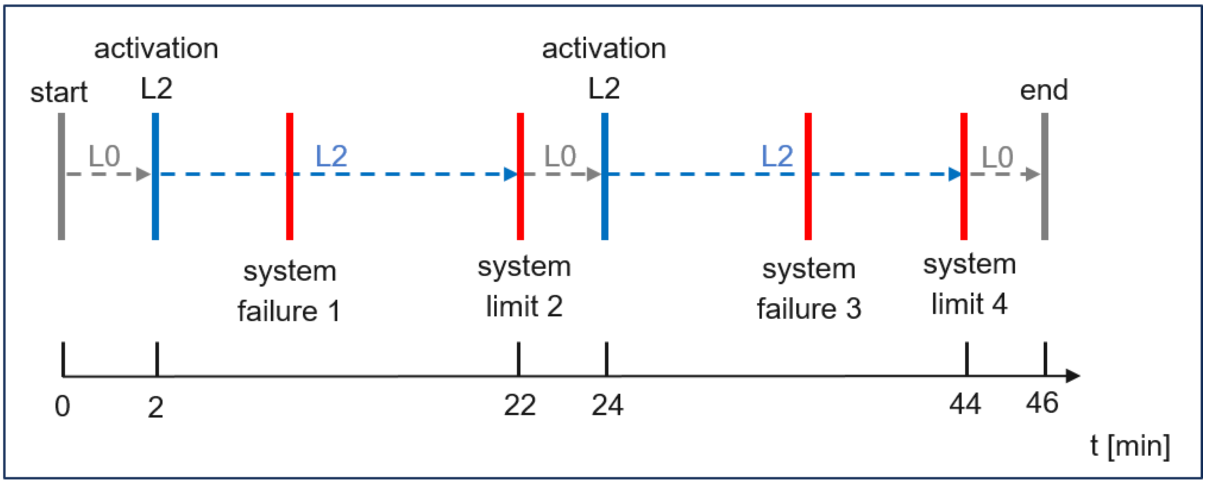


Figure 5-9: Setup of the experimental track over the driving time.

The system failure Scenarios 1 and 3 were designed identically. Participants were driving on the highway with a speed limit of 120 km/h. During the scenario, a new traffic sign with a speed limit of 100 km/h occurred which was not detected by the traffic sign recognition of the vehicle and thus not shown in the instrument cluster. Accordingly, the vehicles with L2 function also continued to drive at 120 km/h when there was no deactivation of the L2 function by the participants. The participants had to recognize on their own that a deactivation was necessary in these scenarios. After one kilometer, the speed limit was set to 120 km/h again.

Scenario 2 was a system limit caused by roadworks where the right lane was blocked (Figure 5-10). This scenario was classified as well anticipatable. Seven hundred meters before the construction site, there was a speed reduction from 120 km/h to 100 km/h, followed by a traffic sign indicating the construction site 600 m before the construction site. Three hundred meters before the construction site, the speed was reduced to 80 km/h, followed by another sign indicating the construction site. Another traffic sign 200 m before the construction site indicated that a lane change to the left would be necessary before the road was blocked by traffic bollards on the right lane. During the scenario, the vehicle in front drove on the right lane at an equivalent speed to the ego vehicle and at a distance of 147 m. Immediately before the construction site, the front vehicle changed to the left lane and cleared the view onto the blocked lane. The drivers then had a time budget of 5 s (at 80 km/h) until the emergency brake assistant intervened 36 m before the roadworks, if the lane had not been changed by then. During the lane change, there were no other vehicles in the left lane in relevant distance behind or next to the ego vehicle.

Figure 5-10: Visualization of Scenario 2 (system limit caused by roadworks blocking the right lane)

Scenario 4 was caused by a broken-down vehicle blocking the right lane (Figure 5-11). The scenario was classified as not being able to be anticipated at an early stage. In this scenario, a vehicle was also driving on the right lane in front of the ego vehicle. Both vehicles were driving 120 km/h according to the speed limit while the vehicle in front had a distance of 252 m to the ego vehicle. The vehicle ahead performed a lane change to the left immediately in front of the broken-down vehicle, leaving the drivers a time budget of approximately 5 s before the vehicle initiated an emergency braking maneuver beginning 95 m in front of the broken-down vehicle, if the driver had not changed lanes by then. During the scenario no vehicles were behind or next to the ego vehicle in a relevant distance.

Figure 5-11: Visualization of Scenario 4 (broken-down vehicle blocking the right lane)

5.2.2.5 Dependent Variables

In Study 1, both objective and subjective metrics were collected, which are presented in Table 5-1 and Table 5-2. Although the focus of this study was CQ1, further metrics were used to provide a basis for addressing the other CQs in the project. For objective data, a distinction was made as to which phase of the trip it referred to (analysis of system limits only or the entire 45-minute trip). For L2 groups, the 45-minute evaluations referred to the time when the L2 function was active.

The data collection was done with the eye tracking system Smart Eye, the driving simulation software SILAB and questionnaires via LimeSurvey. The questionnaires are listed in the Appendix.

Table 5-1: List of objective metrics assigned to CQs

<table>
<tr><th>Construct</th><th>Metric</th><th>Unit</th><th>Time/Event</th><th>Database</th><th>CQ</th></tr>
<tr><td rowspan="2">Visual attention; perceptual readiness</td><td>Eyes-off road glances above 2 s</td><td>Number</td><td rowspan="2">45 min interval (L2 groups during active L2 function)</td><td rowspan="2">Smart Eye</td><td rowspan="3">1</td></tr>
<tr><td>Attention ratios (eyes-on road, instrument cluster/steering wheel, other)</td><td rowspan="3">Percentage</td></tr>
<tr><td>Recognition and reaction to speed limits</td><td>Reaction rate</td><td>Scenario 1/3</td><td rowspan="10">SILAB</td></tr>
<tr><td>The motoric ability for safe vehicle guidance</td><td>Hands-off proportion</td><td>45 min interval in L2 mode</td><td>1, 5</td></tr>
<tr><td>Monitoring</td><td>Hands-off/eyes-off warnings</td><td>Number</td><td>45 min interval in L2 mode</td><td>1</td></tr>
<tr><td rowspan="3">Timing of driver actions</td><td>Mean reaction time to Hands-off-/Eyes-off warnings</td><td rowspan="3">Time in s</td><td>DMS warnings</td><td rowspan="3">2</td></tr>
<tr><td>H-on time</td><td rowspan="6">Scenario 2/4</td></tr>
<tr><td>Time to direct control</td></tr>
<tr><td rowspan="4">Controllability of driver- & system initiated deactivations</td><td>Emergency braking</td><td>Number</td><td rowspan="4">2, 5</td></tr>
<tr><td>Driving trajectory</td><td>Visualization in m</td></tr>
<tr><td>Minimum TTC</td><td>Time in s</td></tr>
<tr><td>Maximum lateral acceleration</td><td>m/s²</td></tr>
<tr><td>Distraction</td><td>NDRT engagement (solved tasks)</td><td>Number/Percentage</td><td>45 min interval</td><td>3</td></tr>
<tr><td>Behavior-based confusion</td><td>Attempted activations of L2 although not available</td><td>Number</td><td></td><td>4</td></tr>
<tr><td>Safety</td><td>Accidents, incidents</td><td>Number</td><td>Scenario 2/4</td><td>SILAB/Observation</td><td>5</td></tr>
</table>

Table 5-2: List of subjective metrics assigned to CQs

Construct	Metric	Description	CQ
Disuse and misuse; Perceived safety	Trust	Trust in automation (Körber, 2019)	3,5
	Acceptance	Car Technology Acceptance Research Model:(Osswald, Wurhofer, Trösterer, Beck, & Tscheligi, 2012) CTAM . Included subscales: Performance expectancy (item PE2 excluded); Effort expectancy; Attitude towards using technology; Facilitating conditions (item FC4 excluded); Behavioral intention to use the system; Perceived safety Excluded subscales: Anxiety; Self-Efficacy; Social Influence Selected and translated by fka, ika, & LfE	
	L2 intention to use	1 item Created by LfE	
Knowledge-based confusion	System understanding	11 statements Created by fka, ika, & LfE	4
	Role understanding	8 statements Created by fka, ika, & LfE	
Distraction	Subjective NDRT engagement	List of 8 activities + free text field for further activities Created by fka, ika, & LfE	3
Monitoring	Estimated monitoring performance	1 item Created by LfE	1
Other	Estimated influence of test setting	1 item + free text field for elaborations Created by LfE	3
Other	Rating of DMS and reported reactions	8 items, created by LfE	3,4,5

5.2.2.6 Procedure

The duration of the experiment was planned with 90 minutes. With invitation to the experiment, participants were asked to complete an online socio-demographic questionnaire prior to the driving simulation session. After the welcome, participants had to read and sign the safety instruction and the declaration of consent. This was followed by a manual in written form, in which the participants were instructed about the driving simulator, the L2 functionalities, the responsibilities and driving task, and the NDRT.

Next, the participants sat down in the vehicle and adjusted the seat and mirrors. The Smart Eye eye-tracking system was then calibrated and a 12-minute familiarization drive took place. In this familiarization drive, participants in the L2 function groups drove both manually and with the group-specific L2 function while the manual group experienced only manual driving. This was followed by an introduction to the use of the tablet for the NDRT. After that, the 45-minute experimental drive was conducted. After the experimental drive, the participants had to fill out a follow-up questionnaire, followed by an interview with the investigator. The experiment ended with the payment of the compensation to the participant.

5.2.2.7 Data Analysis

Data preparation was done with Matlab and MS Excel. JASP 0.16.3.0 was used for the statistical analysis. Normality distribution and equality of variances were analyzed using the Shapiro-Wilk test and the Levene's test. If these assumptions for parametric tests were not met, we used the non-parametric alternative. We performed ANOVA regardless of the violation of the normal distribution, as it appears to be robust to the violation of this condition (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017). In case of violation of equality of variances, Welch ANOVA was performed. The significance level was set to α = .05. In case of multiple comparison, a Bonferroni-Holm correction was applied.

5.2.3 Results

5.2.3.1 Subjective Data

In the following subchapters the results of the follow-up questionnaire are presented. The questionnaire was adjusted for the three groups. Participants of the subgroup L0 did not receive questions that refer to the L2 function. This excluded questions on trust, acceptance, L2 intention to use, system and role understanding as well as the rating of the DMS warnings and the participants' reactions to it.

Trust

The subgroups L2H-on und L2H-off reported their trust in the experienced L2 function via the questionnaire Trust in Automation (Körber, 2019). Figure 5-12 depicts the results of the questionnaire. It comprises the overall score and the following six subscales: *Reliability/Competence; Understanding/Predictability; Familiarity; Intention of Developers; Propensity to Trust; Trust in Automation*.

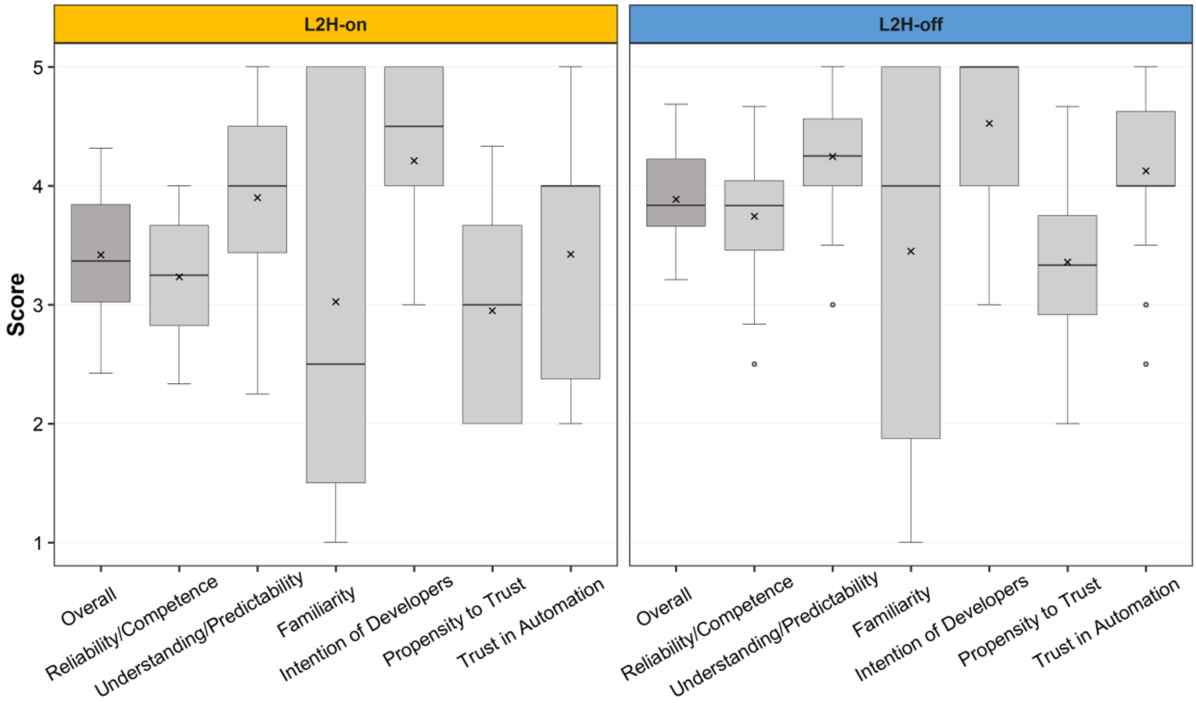


Figure 5-12: Results of the questionnaire Trust in Automation (Körber, 2019) provided by the L2H-on (n = 20) and L2H-off (n = 20) subgroups. The overall score (first scale) bases on the six subscales (second to last scale). Higher scores indicate a higher level of trust.

For all subscales, L2H-off reported higher median and mean values and smaller standard deviations (Table 5-3), indicating a higher trust. In both groups, the subscale *Familiarity* showed the highest variance. For the overall trust score a t-test was conducted that produced significant results with $t(38) = 2.93$, $p = .006$ (L2H-on: $M = 3.42$, $SD = 0.57$; L2H-off: $M = 3.89$, $SD = 0.42$) and a large effect ($d = .93$).

Table 5-3: Descriptive results of the questionnaire Trust in Automation (Körber, 2019)

Scale	**L2H-on (*n* = 20)**			**L2H-off (*n* = 20)**		
	Med	*M*	*SD*	*Med*	*M*	*SD*
Overall	3.37	3.42	0.57	3.84	3.89	0.42
Reliability/Competence	3.25	3.23	0.52	3.83	3.74	0.54
Understanding/Predictability	4.00	3.90	0.73	4.25	4.25	0.53
Familiarity	2.50	3.02	1.66	4.00	3.45	1.56
Intention of Developers	4.50	4.21	0.75	5.00	4.53	0.68
Propensity to Trust	3.00	2.95	0.76	3.33	3.36	0.70
Trust in Automation	4.00	3.42	1.02	4.00	4.12	0.69

Both subgroups showed a level of trust that is medium to high. Participants of the subgroup L2H-off had a significantly higher trust than participants of the subgroup L2H-on. Increased interaction with the DMS in the subgroup L2H-off (see 5.2.3.2) may have led to higher trust scores.

Acceptance

The CTAM (Osswald et al., 2012) was used to measure the acceptance of the L2 function. Figure 5-13 depicts the results of the six subscales *Performance Expectancy; Effort Expectancy; Attitude towards using Technology; Facilitating Conditions; Behavioral Intention to use the System; Perceived Safety*. The model does not include an overall score; thus, t-tests were calculated for all subscales.

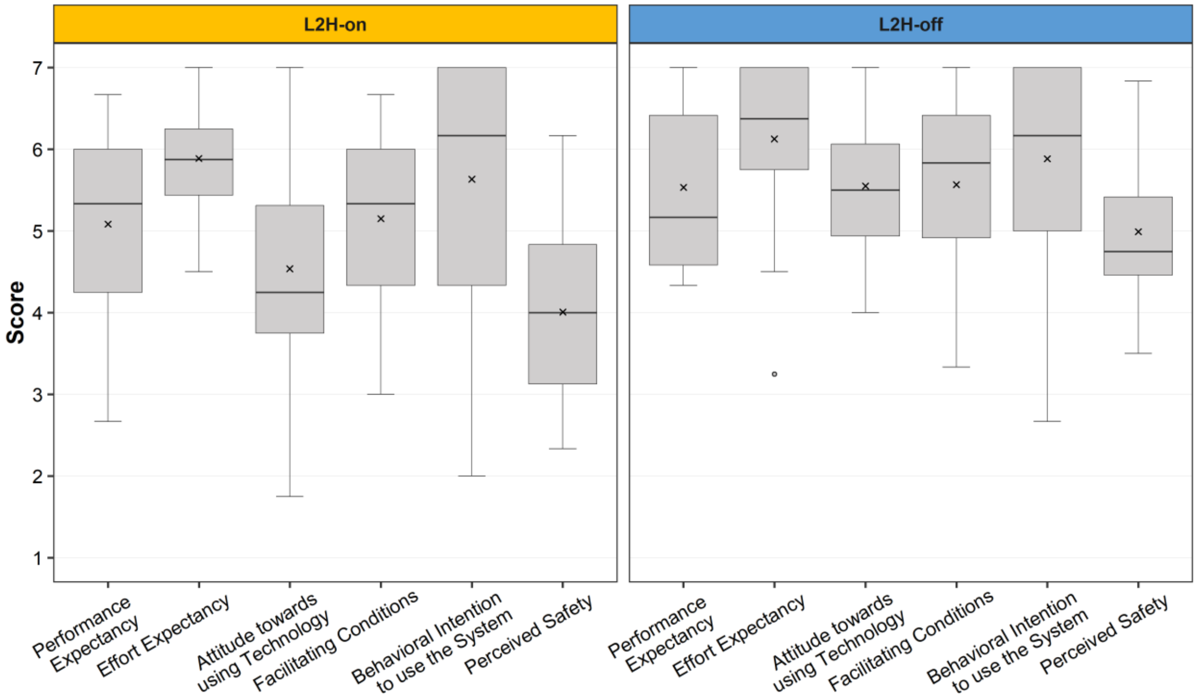


Figure 5-13: Results of the questionnaire CTAM (Osswald et al., 2012) provided by the L2H-on (*n* = 20) and L2H-off (*n* = 20) subgroups. Higher scores indicate a higher level of acceptance.

For most subscales, the L2H-off group reported higher median and mean values and smaller standard deviations (Table 5-4), indicating a higher acceptance. Two subscales produced significant results with a large effect size each. The subscale *Attitude towards using Technology* resulted in $t(38) = 2.93$, $p = .006$ (L2H-on: $M = 4.54$, $SD = 1.22$; L2H-off: $M = 5.55$, $SD = 0.95$) and a large effect ($d = .93$). The subscale *Perceived Safety* resulted in $t(38) = 2.95$, $p = .005$ (L2H-on: $M = 4.01$, $SD = 1.13$; L2H-off: $M = 4.99$, $SD = 0.97$) and a large effect ($d = .93$). The other four subscales showed no significant results.

Both subgroups showed a level of acceptance that is medium to high. Participants of the subgroup L2H-off had a significantly higher acceptance than participants of the subgroup L2H-on for two of the six subscales. The descriptive results strengthen the tendency of higher acceptance ratings of participants with the L2H-off function. Increased interaction with the DMS in the subgroup L2H-off (see 5.2.3.2) may have led to higher acceptance scores in the subscale *Perceived* Safety. The significantly higher ratings of the subgroup L2H-off in the subscale *Perceived Safety* may have induced the significantly higher ratings in the subscale *Attitude towards using Technology*.

Table 5-4: Descriptive results and test statistics of the questionnaire CTAM (Osswald et al., 2012)

Scale	**L2H-on (*n* = 20)**			**L2H-off (*n* = 20)**			**t-Test *df* = 38**
	Med	*M*	*SD*	*Med*	*M*	*SD*	
Performance Expectancy	5.33	5.08	1.18	5.17	5.53	1.03	*t* = 1.28, *p* = .208
Effort Expectancy	5.88	5.89	0.73	6.38	6.12	0.99	*t* = 0.87, *p* = .392
Attitude towards using Technology	4.25	4.54	1.22	5.50	5.55	0.95	*t* = 2.93, *p* = .006, *d* = .93
Facilitating Conditions	5.33	5.15	1.01	5.83	5.57	1.17	*t* = 1.21, *p* = .234
Behavioral Intention to use the System	6.17	5.63	1.53	6.17	5.88	1.27	*t* = 0.56, *p* = .577
Perceived Safety	4.00	4.01	1.13	4.75	4.99	0.97	*t* = 2.95, *p* = .005, *d* = .93

L2 Intention to Use

The following self-created items inquired on the intention to use the overall L2 function and its components. Only L2H-off participants rated the h-off component, i.e., the possibility to drive hands-free. Figure 5-14 depicts the results.

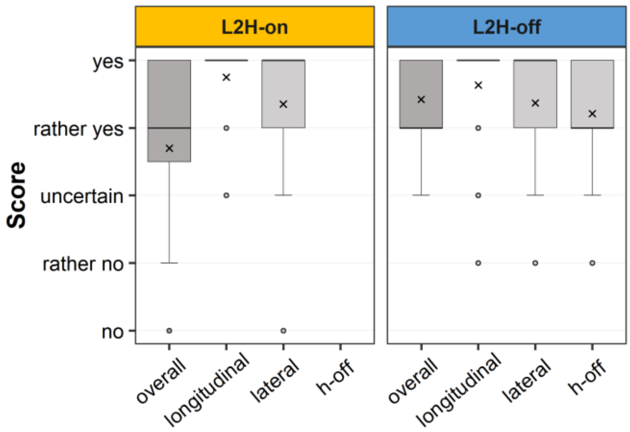


Figure 5-14: Results of the intention to use the L2 function and its components provided by the L2H-on (*n* = 20) and L2H-off (*n* = 20) subgroups. Only L2H-off participants rated the h-off component.

Table 5-5 shows the results of the descriptive and inferential analysis. There were no significant differences between the subgroups.

Both subgroups showed a high intention to use longitudinal assistance and a medium to high intention to use lateral assistance and the overall L2 function. In line with the significantly higher acceptance scores in the subscale *Attitude towards using Technology* (see 0), participants of the subgroup L2H-off showed a slightly higher – non-significant – intention to use the overall system than participants of the subgroup L2H-on.

Table 5-5: Descriptive results and test statistics of the intention to use the L2 function and its components. Only L2H-off participants rated the h-off component.

Component of L2 function	**L2H-on (n = 20)**			**L2H-off (n = = 19)**			**Wilcoxon rank sum test**
	Med	*M*	*SD*	*Med*	*M*	*SD*	
Overall	3.00	2.70	1.34	3.00	3.42	0.61	W = 243.5, p = .105
Longitudinal	4.00	3.75	0.55	4.00	3.63	0.83	W = 185.5, p = .873
Lateral	4.00	3.35	1.04	4.00	3.37	0.83	W = 181.5, p = .801
H-off				3.00	3.21	0.92	

System & Role Understanding

A set of self-developed items inquired on the participants' understanding of the system and the resulting role of the driver. The mean values were high indicating a high system and role understanding among the subgroups. However, there were two outliers in each group with scores below 70% correct answers. The descriptive results depicted in Figure 5-15 indicate that there is no difference between the subgroups. The inferential tests support this observation: The scale *System* resulted in $t(38) < 0.01$, $p = 1.000$ (L2H-on: M = 87.3%, SD = 12.7%; L2H-off: M = 87.3%, SD = 11.6%). The scale *Role* resulted in $t(38) = 0.11$, $p = .910$ (L2H-on: M = 87.5%, SD = 20.3%; L2H-off: M = 88.1%, SD = 13.7%).

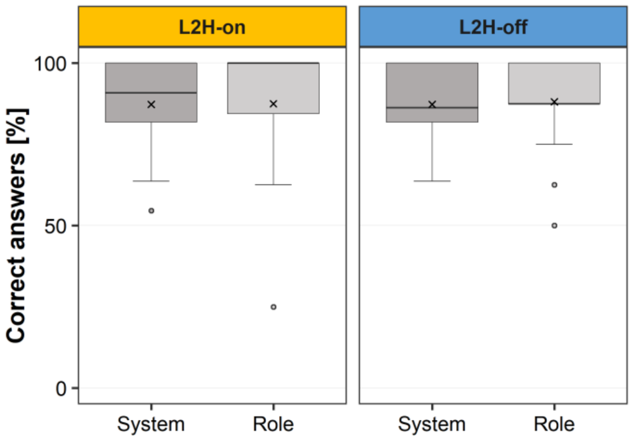


Figure 5-15: Results of the system and role understanding provided by the L2H-on (n = 20) and L2H-off (n = 20) subgroups.

Rating of DMS Warnings and Reactions to the Warnings

Participants of the subgroups L2H-on and L2H-off were requested to evaluate the warnings of the DMS. Figure 5-16 visualizes the differences between the two subgroups. Median values, means and standard deviations as well as the significance tests are listed in Table 5-6.

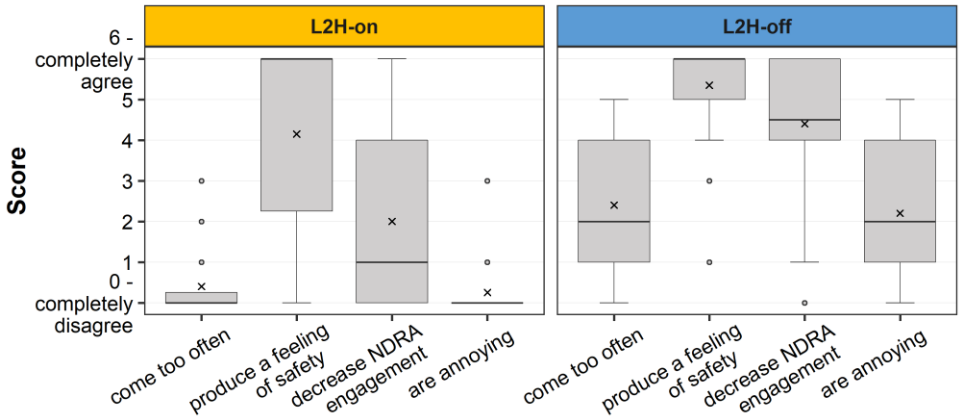


Figure 5-16: Visualization of the evaluation of DMS warnings. The group size is *n* = 20 each.

While most participants of L2H-on completely disagreed that warnings come too often, participants of the subgroup L2H-off showed a higher variance and higher mean and median scores. The difference was significant with a large effect size (W = 343.0, p < .001, d = 1.69). Most participants of both subgroups indicated that the warnings produced a feeling of safety. The standard deviation was lower for the subgroup L2H-off. The Wilcoxon rank sum test was not significant. The answers vary widely in both subgroups regarding the effect of the warnings on the NDRT engagement. While participants of L2H-off rather agreed that the warnings decrease the NDRT engagement, L2H-on participants tended to disagree with this statement. The difference was significant with a large effect size (W = 315.5, p = .002, d = 1.17). Most of the L2H-on participants completely disagreed that the warnings are annoying. Most of the L2H-off participants indicated responses that neither completely agree nor completely disagree. The difference is significant with a large effect size (W = 351.0, p < .001, d = 1.91).

Table 5-6: Descriptive results and test statistics of the evaluation of DMS warnings.

Statement: Warnings …	**L2H-on (*n* = 20)**			**L2H-off (*n* = 20)**			**Wilcoxon rank sum test**
	Med	*M*	*SD*	*Med*	*M*	*SD*	
… come too often	0	0.40	0.82	2.00	2.40	1.57	W = 343.0, p < .001, d = 1.69
… produce a feeling of safety	6.00	4.15	2.58	6.00	5.35	1.31	W = 244.0, p = .175
… decrease NDRT engagement	1.00	2.00	2.27	4.5	4.40	1.90	W = 315.5, p = .002, d = 1.17
… are annoying	0	0.25	0.72	2.00	2.20	1.54	W = 351.0, p < .001, d = 1.91

Participants of the subgroups L2H-on and L2H-off were requested to evaluate their reactions to the warnings of the DMS. Figure 5-17 visualizes the differences between the two subgroups. Median values, means and standard deviations are listed in Table 5-7.

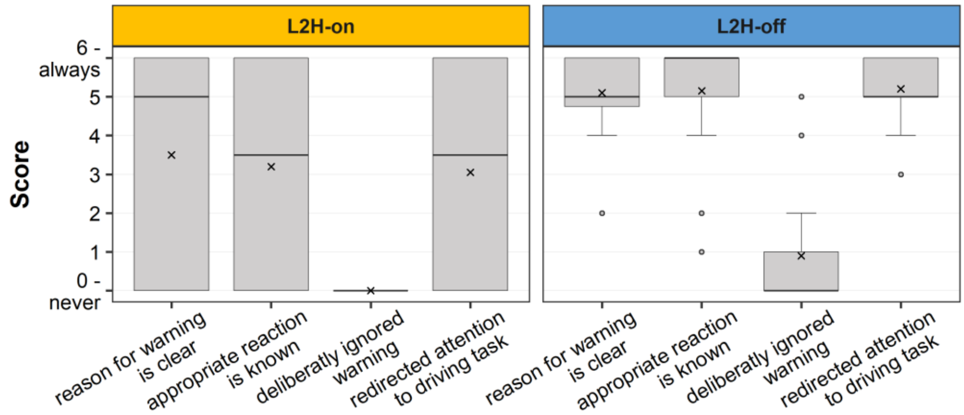


Figure 5-17: Visualization of the evaluation of the reaction to the DMS warnings. The group size is n = 20 each.

For three of the four statements the variance was very high for the subgroup L2H-on and very small for the subgroup L2H-off. The statements refer to the reason for the warning, the appropriate reaction to it and whether the warning had redirected the attention to the driving task. While the median and mean values of the L2H-on subgroup were rather in the center of the scale for these four statements, participants of L2H-off produced high median and mean values. None of the three Wilcoxon rank sum tests was significant.

The fourth statement refers to whether participants have deliberately ignored the warning. All participants of L2H-on reported to have never done this. Some of the L2H-off participants indicated a tendency to deliberately ignore warnings. The Wilcoxon rank sum test produced significant results with a large effect size (W = 290.0, p = .001, d = 1.23).

Table 5-7: Descriptive results and test statistics of the evaluation of the reaction to the DMS warnings.

Statement	**L2H-on (*n* = 20)**			**L2H-off (*n* = 20)**			**Wilcoxon rank sum test**
	Med	*M*	*SD*	*Med*	*M*	*SD*	
The reason for the warning is clear.	5.00	3.50	2.86	5.00	5.10	1.07	W = 236.5, p = .302
The appropriate reaction to the warning is known.	3.50	3.20	2.75	6.00	5.15	1.39	W = 268.0, p = .052
I deliberately ignored the warning.	0	0	0	0	0.90	1.41	W = 290.0, p = .001, d = 1.23
The warning redirected the attention to the driving task.	3.50	3.05	2.93	5.00	5.20	0.89	W = 257.5, p = .103

Increased interaction with the DMS in the subgroup L2H-off (see 5.2.3.2) may explain the differences between the two subgroups. Most of the participants in subgroup L2H-on have not experienced warnings at all or only few warnings. Therefore, participants in this subgroup could only provide hypothetical replies when evaluating the warnings and anticipate their behavior regarding their reactions. In comparison, participants of the subgroups L2H-off frequently experienced warnings. The average replies in Figure 5-17 indicate that the warnings fulfill their purpose. High standard deviations in some statements, e.g., *Warnings are annoying* or *Warnings come too often*, indicate that the acceptance of the L2 function could decrease in prolonged usage periods. The design of the warnings should be the subject of future research in naturalistic settings.

Subjective NDRT Engagement

Figure 5-20 visualizes the self-reported NDRT engagement, referring to an anticipated engagement in NDRT. The tasks were assigned to three categories: visual and motoric (yellow), primary motoric (green) and primary auditory (red). Based on the descriptive analysis, no systematic differences were identified for the different categories. The median of the subgroup L2H-off was higher for most of the different NDRTs (Figure 5-18).

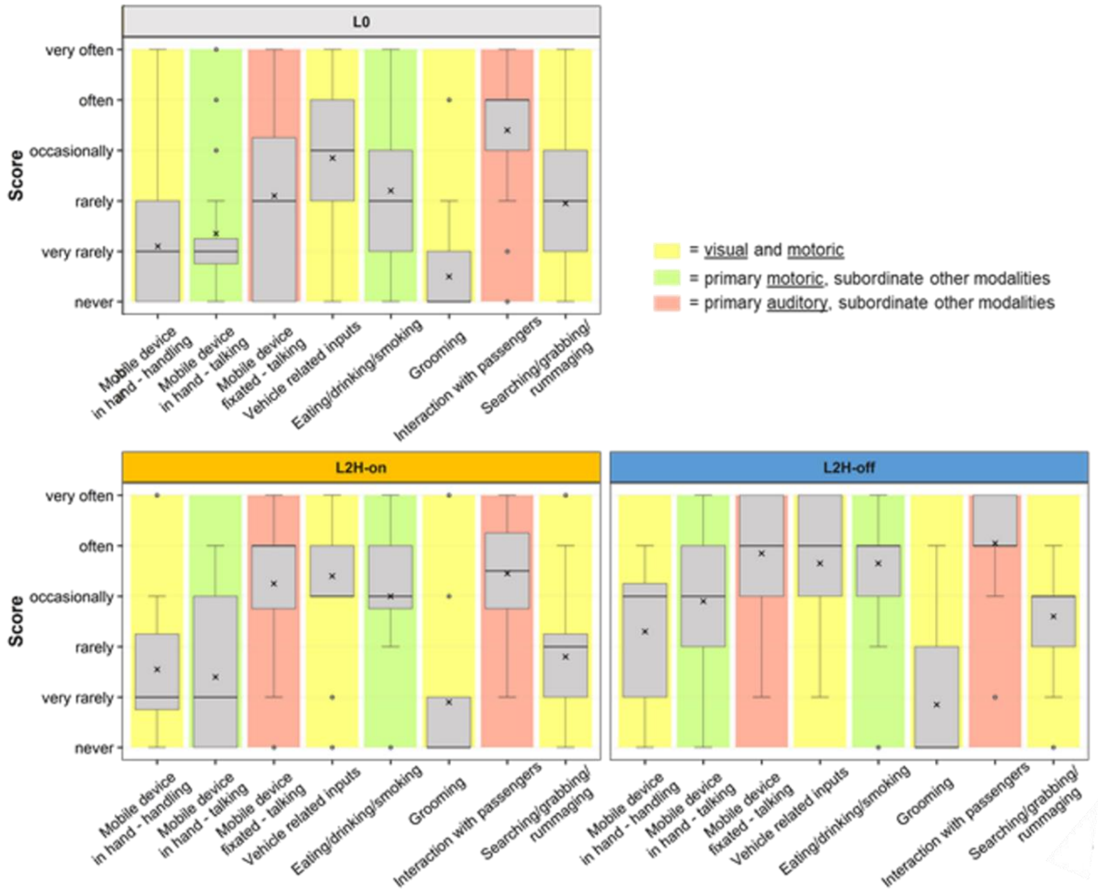


Figure 5-18: Visualization of the self-reported engagement with NDRTs. Participants were instructed to refer to the anticipated engagement with NDRTs when using the system experienced in the experiment. The group size is *n* = 20 each.

Table 5-8: Descriptive results of the self-reported engagement with NDRTs.

NDRT	L0 (*n* = 20)			L2H-on (*n* = 20)			L2H-off (*n* = 20)		
	Med	*M*	*SD*	*Med*	*M*	*SD*	*Med*	*M*	*SD*
Mobile device in hand – handling	1.00	1.10	1.21	1.00	1.55	1.36	3.00	2.30	1.49
Mobile device in hand – talking	1.00	1.35	1.39	1.00	1.40	1.35	3.00	2.90	1.55
Mobile device fixated - talking	2.00	2.10	1.89	4.00	3.25	1.71	4.00	3.85	1.18
Vehicle related inputs	3.00	2.85	1.42	3.00	3.40	1.27	4.00	3.65	1.23
Eating/drinking/smoking	2.00	2.20	1.36	3.00	3.00	1.75	4.00	3.65	1.14
Grooming	0	0.50	1.00	0	0.90	1.59	0	0.85	1.23
Interaction with passengers	4.00	3.40	1.43	3.50	3.45	1.23	4.00	4.05	0.94
Searching/grabbing/rummaging	2.00	1.95	1.15	2.00	1.80	1.44	3.00	2.60	1.23

The descriptive analysis indicates that the anticipated engagement with NDRTs may be increased for the L2H-off function compared to L0 and L2H-on driving. However, the differences were not large and were accompanied with high standard deviations in all subgroups.

Self-reported Monitoring Performance

Overall, participants reported to have been neither inattentive nor always attentive while monitoring the driving task. The variance was large among all three subgroups. The median was lowest in the subgroup L0 (*Med* = 2.0) and highest in subgroup L2H-off (*Med* = 4.0; L2H-on: *Med* = 2.5). This indicates that the self-reported monitoring performance was not decreased by the use of a L2 function but may even be increased. The distribution is visualized in Figure 5-19.

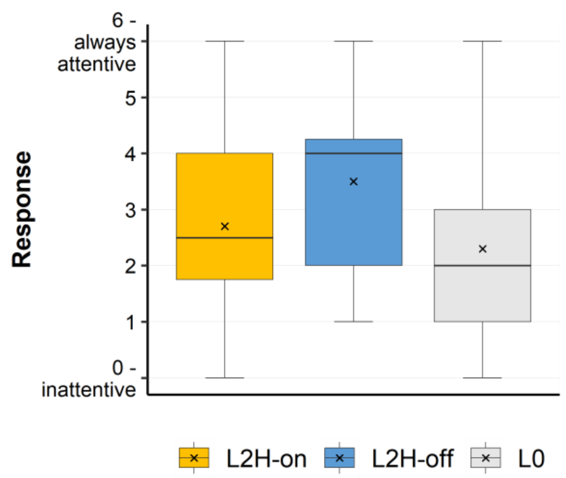


Figure 5-19: Visualization of the self-reported monitoring performance during the experiment. The group size is $n = 20$ each.

Self-Reported Influence of Test Setting

About three quarters of participants of all three subgroups reported to be affected by the test setting. The distribution among the subgroups is similar (see Figure 5-20).

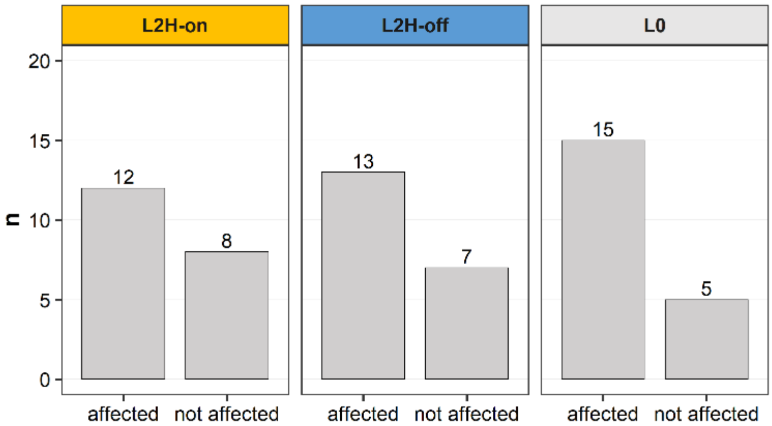


Figure 5-20: Distribution of participants reporting whether they were affected by the test setting. All three subgroups are displayed (each group: n = 20).

The majority of participants reported that they would have engaged less in the NDRT if they weren't in the test setting. This implies that the objective results regarding NDRT engagement should be treated with caution when transferring them to naturalistic driving setting. The proportion of participants indicating an influence of the test setting and their explanations are similar among all three subgroups. Therefore, the relative comparability among the three groups remains unscathed.

5.2.3.2 Objective Data

Due to faulty data recording objective data files of one participant of the L2H-off group and one participant of the L2H-on group had to be excluded.

Overall Drive

One participant of the L2H-on group did not show a data availability of at least 70%, which led to a further exclusion from data analysis according to ISO/TS 15007-2 (2013). This resulted in 57 valid eye-tracking data sets, which showed a good data availability of L2H-on with M = 93.25% (SD = 7.61%), L2H-off with M = 94.34% (SD = 2.07%) and L0 with M = 93.26% (SD = 4.83%).

Eyes-off Road Glances >2s

Eyes-off road glances were calculated only for the driving phase with active L2 function for the L2H-on and L2H-off groups and for the phase when L2 function was available in principle in the other groups for the L0 group. Welch ANOVA showed a significant effect of the eyes-off road glances above 2 s between the groups ($F(2, 30.346) = 9.19$, $p < .001$, $\eta_p^2 = .22$). The results are presented in Figure 5-21 and Table 5-9. According to the post-hoc tests, there was only a significant difference between L2H-off and L2H-on ($p_{Holm} < .001$; Table 5-10), resulting in L2H-off participants looking away more often for at least 2 s.

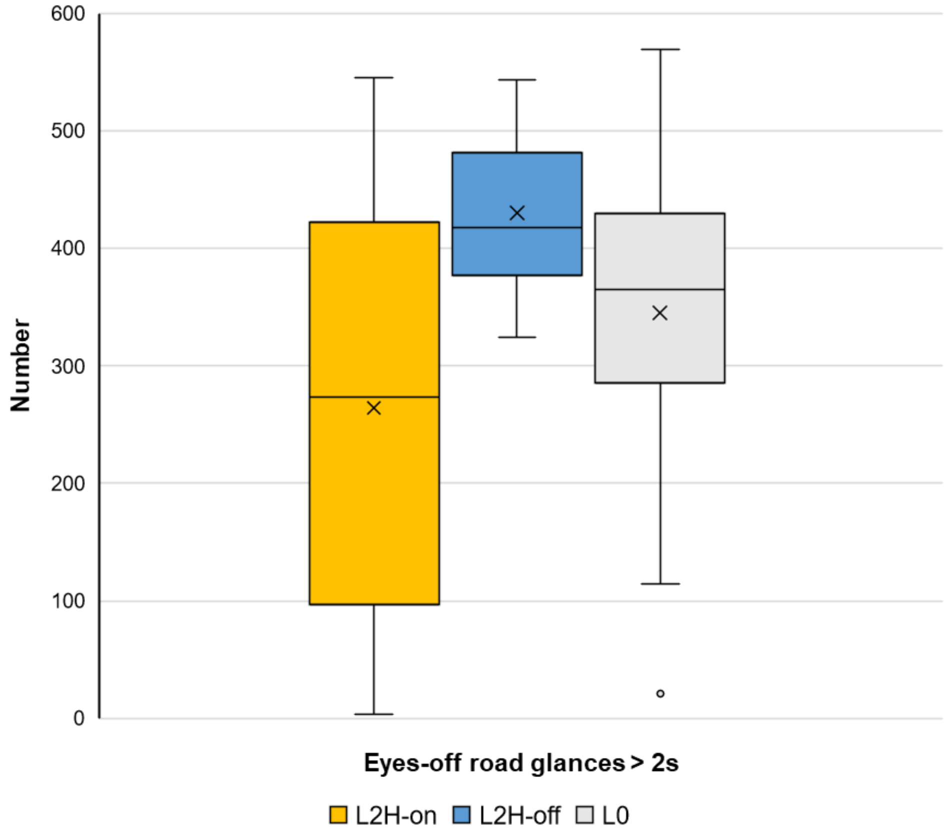


Figure 5-21: Boxplots showing number of eyes-of road glances above 2 s for L2H-on (*n* = 18), L2H-off (*n* = 19) and L0 (*n* = 20).

Table 5-9: Descriptive results for number of eyes-of road glances above 2 s.

Driving Mode	***M***	***SD***
L2H-on (*n* = 18)	264.06	176.70
L2H-off (*n* = 19)	430.21	62.96
L0 (*n* = 20)	345.20	126.39

Table 5-10: Post-hoc comparisons for number of eyes-of road glances above 2 s between L2H-on (*n* = 18), L2H-off (*n* = 19) and L0 (*n* = 20).

Driving Mode		p_{Holm}
L2H-off	L2H-on	< .001
L2H-off	L0	.091
L2H-on	L0	.091

Visual Attention Ratio

Visual attention ratio was analyzed for eyes on road, eyes on instrument cluster and eyes on other. Eyes on road corresponded to the AOI windshield without the rearview mirror. Eyes on instrument cluster was identical to the AOI instrument cluster and eye on other included all other detected glances including the AOI for the NDRT. We conducted a Welch ANOVA for the visual attention ratio on the AOI eyes on road, which showed a significant difference ($F(2, 21.552) = 28.74$, $p < .001$, $\eta_p^2 = .48$). Post-hoc tests revealed that all three groups differed significantly from each other ($p < .001$, Table 5-12) while the L0 group showed the highest and the L2H-on group showed the lowest proportion for eyes on road.

According to the results of visual attention ratio and eyes-off road glances above 2 s, L2H-off led to more frequent eyes off road glances above 2 s, but overall lower overall eyes off road time compared to L2H-on. This may reduce the risk of overlooking something by the L2H-off group compared to the L2H-on group. The results of the visual attention ratio are presented in Figure 5-22 and Table 5-11.

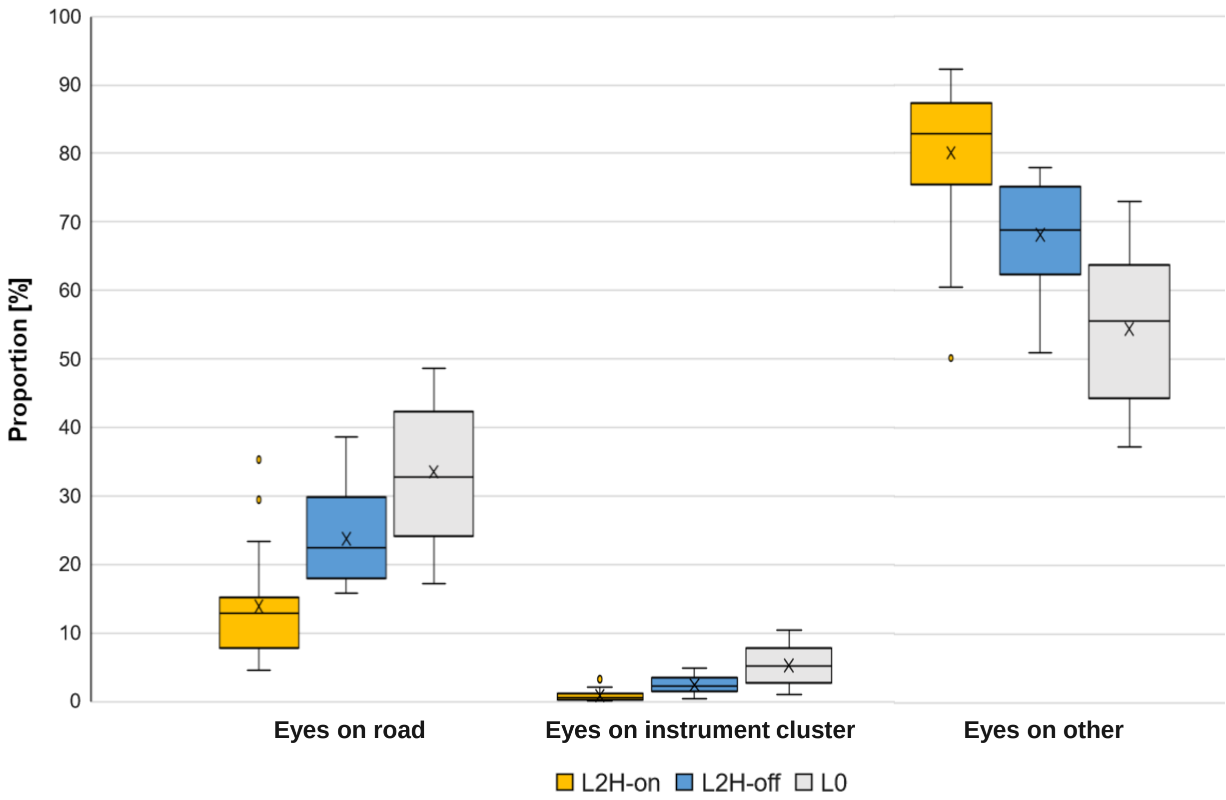


Figure 5-22: Boxplots showing the visual attention ratio for L2H-on (n = 18), L2H-off (n = 19) and L0 (n = 20).

On a descriptive level, the AOIs instrument cluster and other were also considered. L0 resulted in the largest and L2H-on in the lowest eyes on instrument cluster ratio. This could be because participants in the L0 group had to continuously check the speed in the instrument cluster and maintain it manually. The L2H-off group experienced more DMS warnings than the L2H-on group, which were presented in the instrument cluster (see hands-off/eyes-off warnings) and thus might cause higher visual attention. Eyes on other resulted accordingly from the other AOIs and thus showed the highest ratio for L2H-on and the lowest ratio for L0.

Table 5-11: Descriptive results for visual attention ratio.

Driving Mode	Eyes on road [%] *M (SD)*	Eyes on instrument cluster [%] *M (SD)*	Eyes on other [%] *M (SD)*
L2H-on (n = 18)	13.90 (8.22)	0.85 (0.82)	80.10 (10.90)
L2H-off (n = 19)	23.76 (6.89)	2.45 (1.31)	68.14 (7.74)
L0 (n = 20)	33.57 (10.15)	5.26 (2.66)	54.43 (11.36)

Table 5-12: Post-hoc comparisons for visual attention ratio regarding AOI eyes on road between L2H-on (n = 18), L2H-off (n = 19) and L0 (n = 20).

Driving Mode		p_{Holm}
L2H-off	L2H-on	< .001
L2H-off	L0	< .001
L2H-on	L0	< .001

Hands-off Proportion

For the hands-off proportion, Welch ANOVA revealed a significant difference (F(2, 25.542) = 89.68, p < .001, η_p^2 = .85). Post-hoc tests resulted in significant differences of the L2H-off group compared to the L2H-on and the L0 group. Thus, L2H-off showed the largest proportion of hands off, but still showed a high variance between its participants, which is consistent with the design differences of the two L2 functions. Additionally, L2H-on and L0 participants showed hands-off proportions greater than zero but at a similar level. Therefore, it can be assumed that the participants of the L2H-on had an appropriate hands-on proportion.

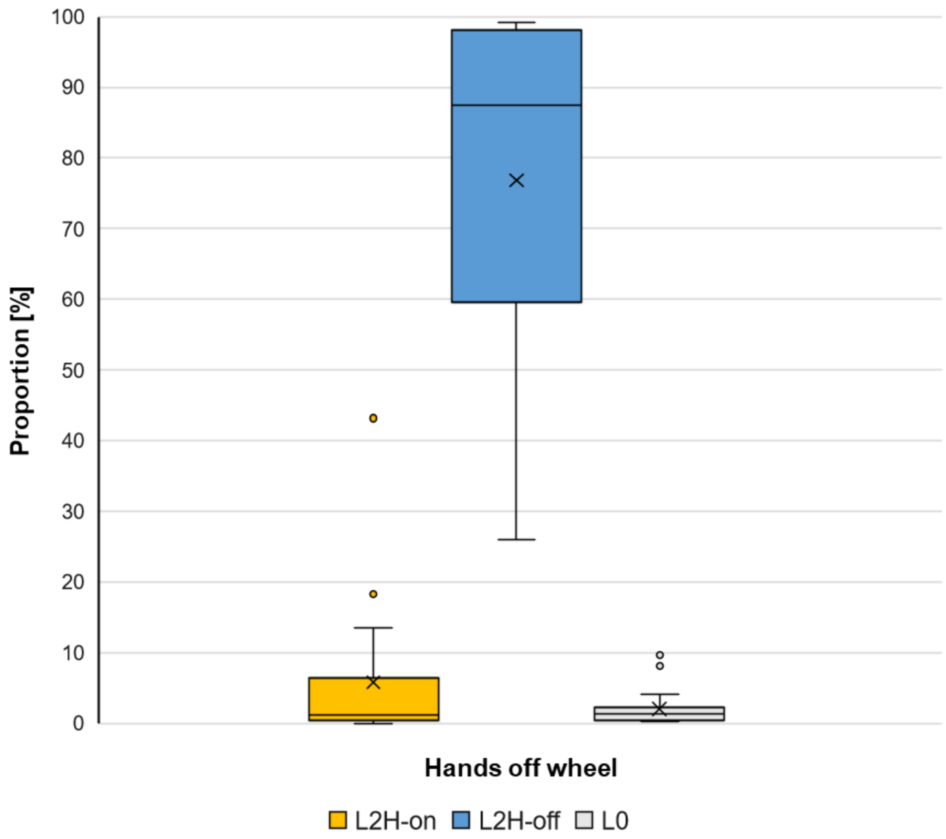


Figure 5-23: Boxplots showing hands-off wheel proportion for L2H-on (*n* = 19), L2H-off (*n* = 19) and L0 (*n* = 20).

Table 5-13: Descriptive results for hands-off wheel proportion.

Driving Mode	***M***	***SD***
L2H-on (*n* = 18)	5.79%	10.43%
L2H-off (*n* = 19)	76.83%	24.00%
L0 (*n* = 20)	2.04%	2.56%

Table 5-14 Post-hoc comparisons for hands-off wheel proportion between L2H-on (*n* = 18), L2H-off (*n* = 19) and L0 (*n* = 20).

Driving Mode		p_{Holm}
L2H-off	L2H-on	< .001
L2H-off	L0	< .001
L2H-on	L0	.440

Hands-off/Eyes-off Warnings

L2H-off showed significantly higher number of eyes-off warnings (warning stage 1) compared to the number of hands-off warnings in L2H-on (U = 361.00, p < .001, r = 1.00; Figure 5-24).

The first warning stage started after 5 s for the L2H-off group and after 15 s for the L2H-on group. Warning stages 2 and 3 were not triggered by any participant in the L2H-on group. In contrast, the L2H-off group also showed warnings of warning stage 2 and 3 (see Table 5-15).

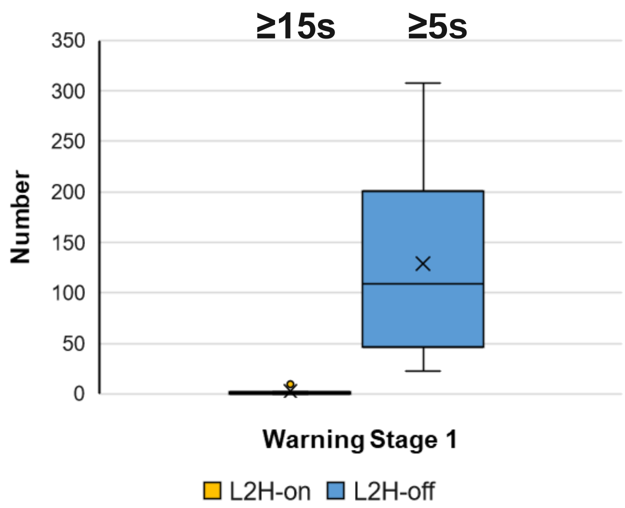


Figure 5-24: Boxplots showing the number of hands-off warnings (L2H-on, *n* = 19) and number of eyes-off warnings (L2H-off, *n* = 19).

Table 5-15: Descriptive results for number of number of hands-off warnings (L2H-on) and number of eyes-off warnings (L2H-off).

Driving Mode	***n* with # > 0**	***M***	***SD***
	Warning Stage 1		
L2H-on (*n* = 19)	10	2.53	4.65
L2H-off (*n* = 19)	19	128.68	80.97
	Warning Stage 2		
L2H-on (*n* = 19)	0	-	-
L2H-off (*n* = 19)	13	10.79	16.35
	Warning Stage 3		
L2H-on (*n* = 19)	0	-	-
L2H-off (*n* = 19)	6	2.74	4.42

Reaction time to Hands-off-/Eyes-off warnings

A t-test was used to analyze the mean reaction time to hands-off warnings in L2H-on and eyes-off warnings in L2H-off. No significant difference was found (t = 1.66, p = .108, d = .65). Participants reacted after M = 1.79 s (SD = 0.74 s) to hands-off warnings in L2H-on and M = 1.39 s (SD = 0.55 s) to eyes-off warnings in L2H-off.

Figure 5-25: Boxplots showing mean reaction time to DMS warnings for L2H-on (n = = 19) and L2H-off (n = 19).

NDRT engagement

The number of solved tasks of the NDRT showed a significant difference ($F(2, 56) = 5.04$, $p = .010$, $\eta_p^2 = .15$; Figure 5-26).

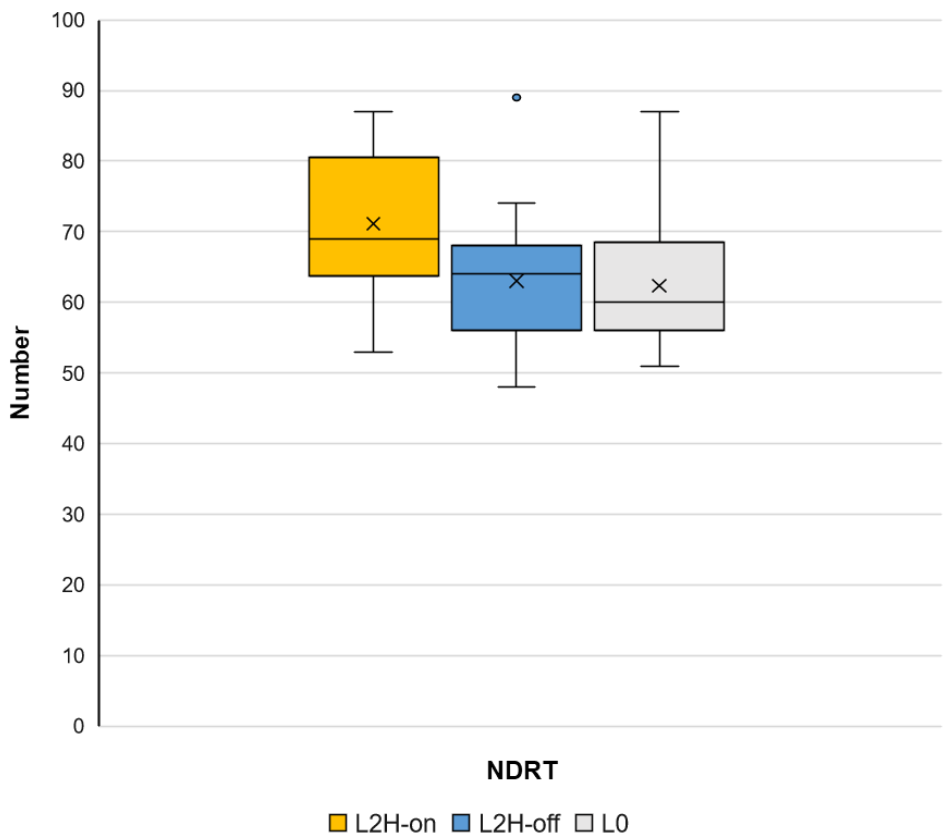


Figure 5-26: Boxplots showing number of completed tasks for L2H-on (n = 20), L2H-off (n = 19) and L0 (n = 20). The hypothetical maximum number of tasks presented to each participant was 100.

We conducted post-hoc tests that revealed a significant difference between L2H-on and L2H-off and between L2H-on and L0 (Table 5-16). In both comparisons, more tasks were solved in L2H-on. No significant differences were found during the post-hoc comparison between L2H-off and L0. Here, L2H-off and L0 showed a similar level of the average number of solved tasks.

Table 5-16: Descriptive results for number of completed tasks. The hypothetical maximum number of tasks presented to each participant was 100.

Driving Mode	***M***	***SD***
L2H-on (n = 20)	71.10	9.80
L2H-off (n = 19)	63.05	9.76
L0 (n = 20)	62.35	9.41

Table 5-17: Post-hoc comparisons for number of completed tasks between L2H-on (n = 20), L2H-off (n = 19) and L0 (n = 20).

Driving Mode		p_{Holm}
L2H-off	L2H-on	.024
L2H-off	L0	.821
L2H-on	L0	.018

As a second metric for evaluating NDRT engagement, the proportion of correctly solved tasks was used. A Welch ANOVA showed no significant differences between the different driving modes (F(2, 35.52) = 2.55, p = .092). Here, L0 showed a higher variance than both L2 groups resulting in a lower average proportion of correctly solved tasks (Figure 5-27, Table 5-18). The results showed that the L2H-on group not only completed more tasks, but also solved more tasks correctly in absolute terms compared to L2H-off and L0.

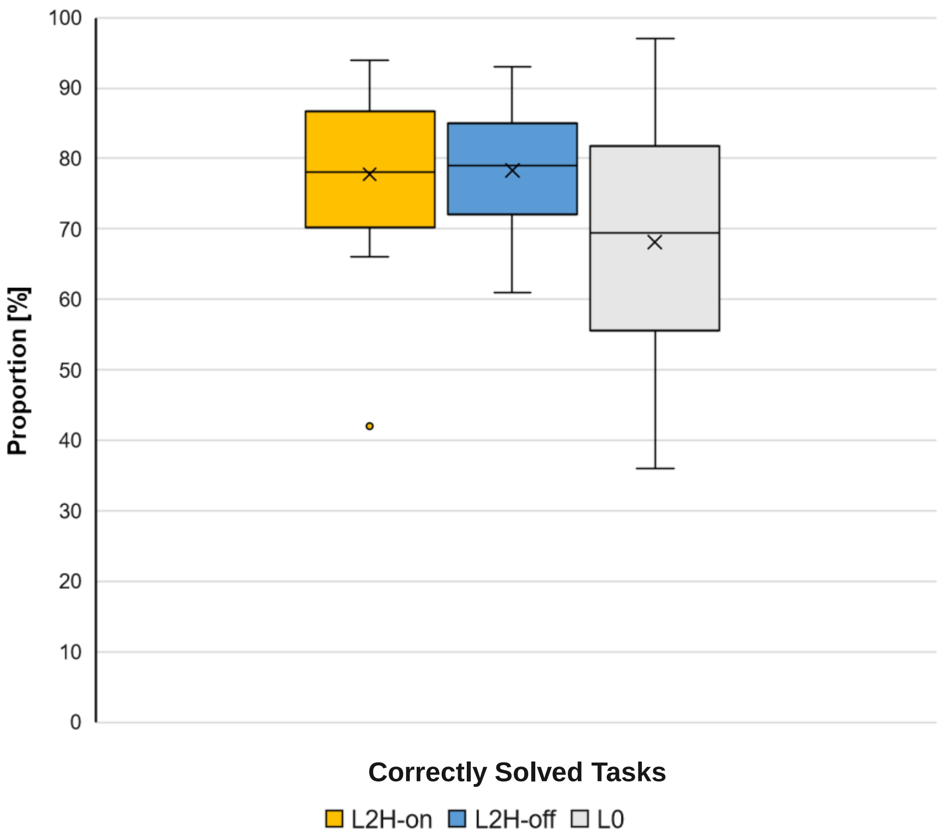


Figure 5-27: Boxplots showing the proportion of correctly solved tasks for L2H-on (*n* = 20), L2H-off (*n* = 19) and L0 (*n* = 20).

Table 5-18: Descriptive results for the proportion of correctly solved tasks.

Driving Mode	*M*	*SD*
L2H-on (*n* = 20)	77.75%	11.92%
L2H-off (*n* = 19)	78.26%	8.93%
L0 (*n* = 20)	68.15%	18.10%

Attempted activations of L2 although not available

The number of attempted activations of L2, although L2 was not available, can generally be described as low for L2H-on and for L2H-off (Figure 5-28).

One outlier existed in L2H-off, while there was no significant difference between L2H-on and L2H-off according to the conducted Wilcoxon rank sum test (W = 189, p = .805). However, the L2 function was only unavailable at the beginning, while passing the construction site (1.80 km) and at the end of the experimental drive, which would generally not be expected to produce high numbers of attempted activations of L2, if L2 was not available.

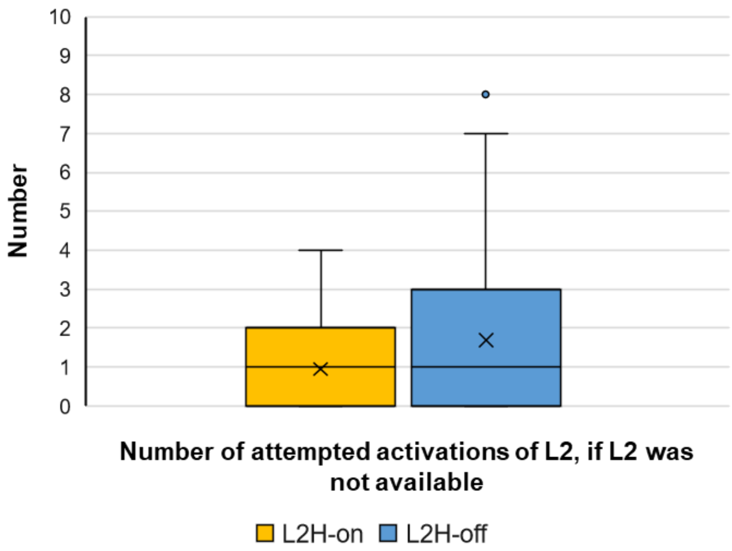


Figure 5-28: Boxplots showing the number of attempted activations of L2, if L2 was not available for L2H-on (*n* = 19) and L2H-off (*n* = 19).

Table 5-19: Descriptive results for number of attempted activations of L2, if L2 was not available.

Driving Mode	*M*	*SD*
L2H-on (*n* = 19)	0.95	1.08
L2H-off (*n* = 19)	1.68	2.47

Scenario 1/3

Reaction Rate

In Scenario 1 and 3, where a failure of the speed limit detection occurred, the proportion of reactions by regaining direct control that required a recognition of the failure was analyzed. This proportion was examined only for L2H-on and L2H-off because reliable verification of whether participants in the L0 group reacted to the speed limit, as they were constantly in direct control, was not possible. Deceleration via adaption of set-speed was not possible in this study. During the first speed limit failure (Scenario 1), 32% of the participants in the L2H-on group (*n* = 19) and 26% of the L2H-off users (*n* = 19) reacted to the speed limit failure by regaining direct control and slowing the vehicle down. In Scenario 3, 32% of the L2H-on group (*n* = 19) and 32% of the L2H-off (*n* = 19) reacted to the speed limit. Overall, there was no difference between the two L2 systems on a descriptive level. Nevertheless, the speed limits did not seem to be recognizable for a majority of the participants. However, it cannot be ruled out that individual participants intentionally did not react in order to drive faster.

Scenario 2/4

H-on time

Because participants had at least one hand on the steering wheel most of the time during L2H-on and L0, H-on time was examined only for using L2H-off (see Figure 5-29). During the anticipatory scenario roadworks, the participants of the L2H-off group took their hands back on the steering wheel at *M* = -5.73s (*SD* = 3.85s). Accordingly, with the exception of two outliers, many participants had their hands back on the steering wheel before the traffic bollards and the blocked lane were directly visible in their own lane. During the non-anticipatory cut-out scenario, participants of L2H-off had their hands back on the steering wheel after *M* = 5.1s (*SD* = 1.99s). One outlier had at least one hand back on the steering wheel before the broken-down vehicle was visible, which probably happened at random.

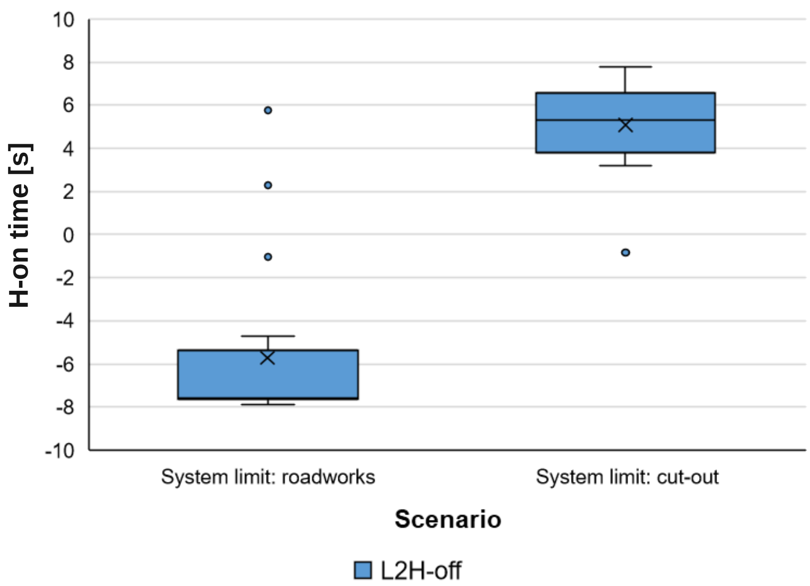


Figure 5-29: Boxplots showing hands-on time for scenarios roadworks (Scenario 2) and cut-out (Scenario 4) for L2H-off (*n* = 19).

Time to direct control

During Scenario 2, four participants of L2H-on and two participants of L2H-off did not deactivate the automation by steering, braking or by pressing a button before coming to a standstill due to the emergency braking function. These participants are not included in Figure 5-30 presenting the time to direct control and the statistical tests. The roadworks scenario showed no significant difference between L2H-on (*n* = 15) and L2H-off (*n* = 17) according to the Mann-Whitney-U-test (*U* = 95.00, *p* = .230). In contrast to Scenario 2, no participants were additionally excluded in the cut-out scenario (Scenario 4). In this scenario, no significant difference was found between L2H-on (n = 15) and L2H-off (*n* = 17) based on the t-test (*t*(36) = -1.26, *p* = .216), showing an average time to direct control greater than the time budget of 5 s for both groups.

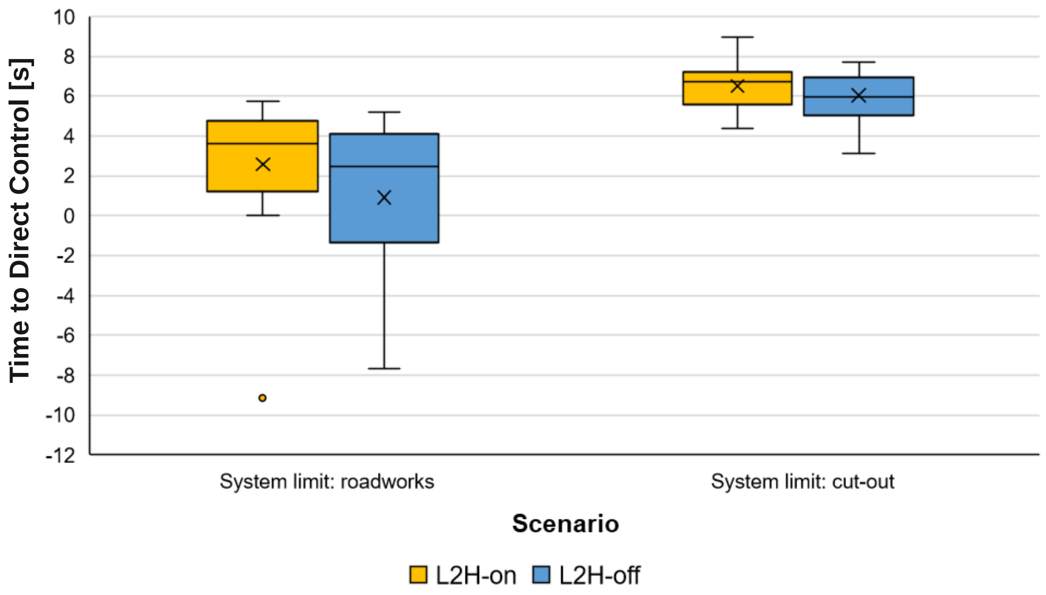


Figure 5-30: Boxplots showing time to direct control during the roadworks (L2H-on: *n* = 15, L2H-off: *n* = 17) and cut-out scenarios (L2H-on: *n* = 19, L2H-off: *n* = 19). The zero point corresponds to the beginning of the scenario when the blocked lane became visible.

Table 5-20: Descriptive results for time to direct control in the roadworks and cut-out scenarios. The zero point corresponds to the beginning of the scenario when the blocked lane became visible.

Driving Mode	***M***	***SD***
	Roadworks	
L2H-on (*n* = 15)	2.58 s	3.64 s
L2H-off (*n* = 17)	0.91 s	4.29 s
	Cut-Out	
L2H-on (*n* = 19)	6.51 s	1.13 s
L2H-off (*n* = 19)	6.04 s	1.18 s

Emergency Braking Maneuvers

In both scenarios, the emergency brake assistant was activated after the time budget of 5 s was reached, if the ego-vehicle was still in the right lane. L0 participants showed the lowest number of emergency braking maneuvers and L2H-on the highest number of emergency braking maneuvers in the roadworks scenario (see Table 5-21). Anyway, no clear trend was identified. In the cut-out scenario, 17 out of 19 participants in each group experienced the emergency braking.

Table 5-21: Descriptive results for the number of emergency braking maneuvers during the roadworks and cut-out scenarios.

Driving Mode	***Emergency Braking***	***No Emergency Braking***
	Roadworks	
L2H-on (*n* = 19)	13	6
L2H-off (*n* = 19)	11	9
L0 (*n* = 20)	9	11
	Cut-Out	
L2H-on (*n* = 19)	17	2
L2H-off (*n* = 19)	17	2
L0 (*n* = 19)	17	2

Minimum TTC

The minimum TTC was analyzed separately for both scenarios. In the roadworks scenarios, Welch ANOVA revealed a significant effect ($F(2, 27.58) = 6.52$, $p = .005$, $\eta_p^2 = .15$). These results are presented in Table 5-22 and Figure 5-31. Post-hoc tests showed a significant difference between L2H-off and L2H-on (see Table 5-23). L2H-on resulted in a lower and thus more critical TTC. No significant differences were found for the comparison of the L2 groups with L0. For the cut-out scenario, no significant effect was found regarding the minimum TTC ($F(2, 54) = 0.65$, $p = .529$). The cut-out scenario resulted in lower minimum TTC than the roadworks scenarios, regardless of the driving mode. Since most participants in all groups experienced an emergency braking maneuver in the cut-out scenario, the results of the minimum TTC in this scenario are more due to the emergency braking maneuvers than to the human drivers.

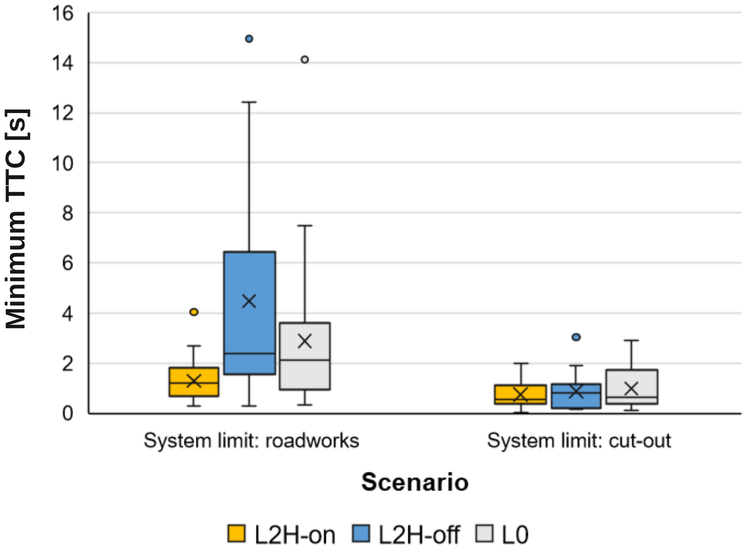


Figure 5-31: Boxplots showing the minimum time-to-collision (TTC) in the roadworks (L2H-on: *n* = 19, L2H-off: *n* = 19, L0: *n* = 20) and cut-out scenarios (L2H-on: *n* = 19, L2H-off: *n* = 19, L0: *n* = 19).

Table 5-22: Descriptive results for the minimum time-to-collision (TTC) during the roadworks and cut-out scenarios.

Driving Mode	*M*	*SD*
	Roadworks	
L2H-on (*n* = 19)	1.30 s	0.93 s
L2H-off (*n* = 18)	4.48 s	4.40 s
L0 (*n* = 20)	2.89 s	3.91 s
	Cut-Out	
L2H-on (*n* = 19)	0.72 s	0.55 s
L2H-off (*n* = 18)	0.88 s	0.72 s
L0 (*n* = 19)	0.97 s	0.77 s

Table 5-23: Post-hoc comparison for minimum time-to-collision for the scenario roadworks between L2H-on (*n* = 19), L2H-off (*n* = 19) and L0 (*n* = 20).

Driving Mode		p_{Holm}
L2H-off	L2H-on	.010
L2H-off	L0	.247
L2H-on	L0	.247

Driving Trajectory

The visualization of the ego-vehicle's driving trajectories during the roadworks scenario are presented in Figure 5-32 to Figure 5-34. It is noticeable that individual participants in L2H-off took direct control earlier and changed lanes to the left compared to L2H-on and thus appeared to anticipate the blocked lane due to the road works earlier.

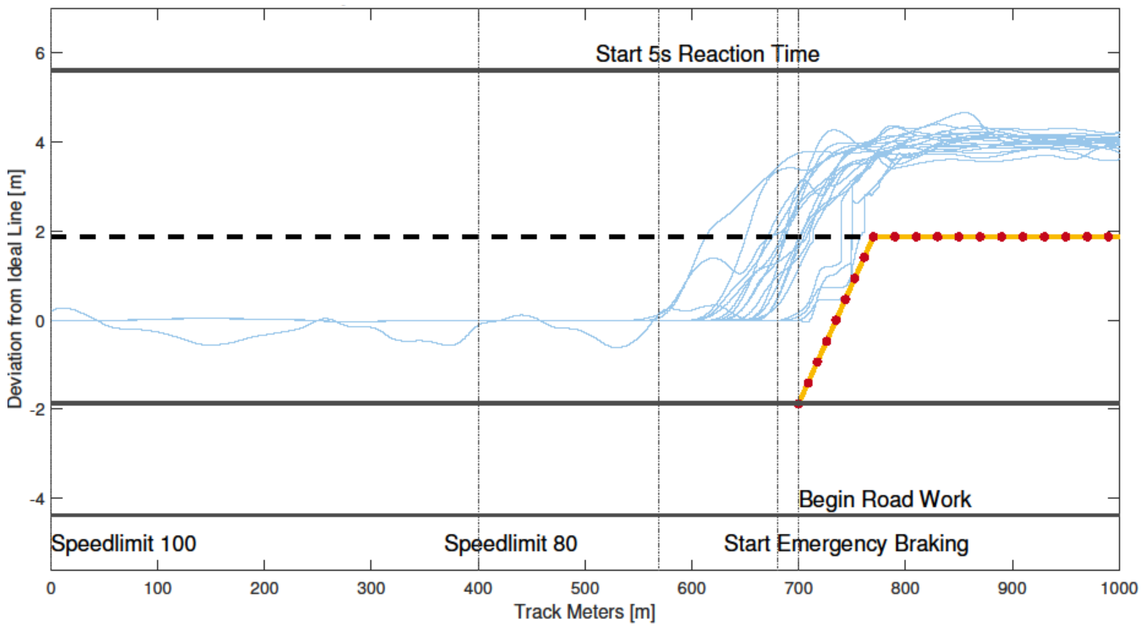


Figure 5-32: Ego-vehicle's driving trajectory of the L2H-on group during the roadworks scenario. Each participant is plotted as a single line (*n* = 19).

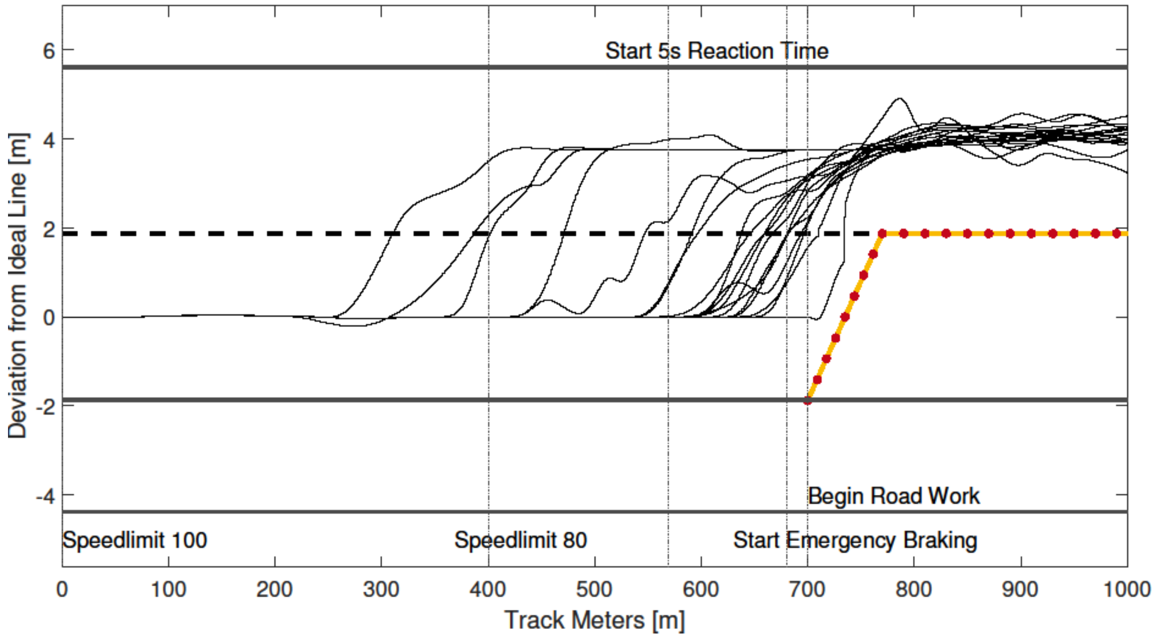


Figure 5-33: Ego-vehicle's driving trajectory of the L2H-off group during the roadworks scenario. Each participant is plotted as a single line (n = 19).

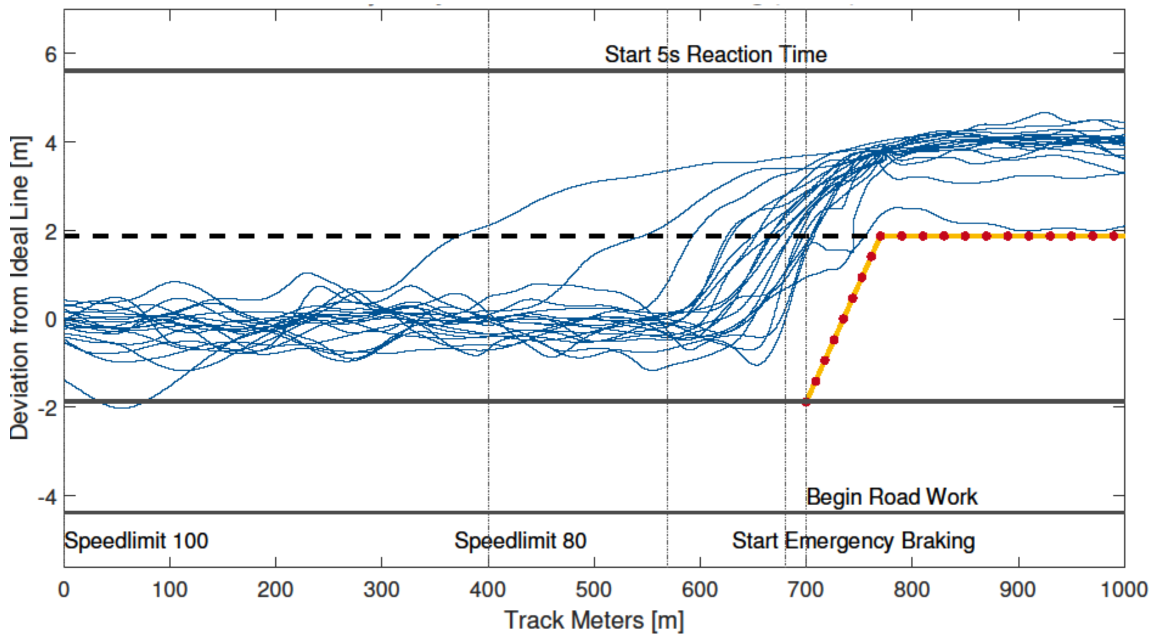


Figure 5-34: Ego-vehicle's driving trajectory of the L0 group during the roadworks scenario. Each participant is plotted as a single line (n = 19).

The visualization of the ego-vehicle's driving trajectories during the cut-out scenario are presented in Figure 5-35 to Figure 5-37. No clear differences can be observed between the different driving modes.

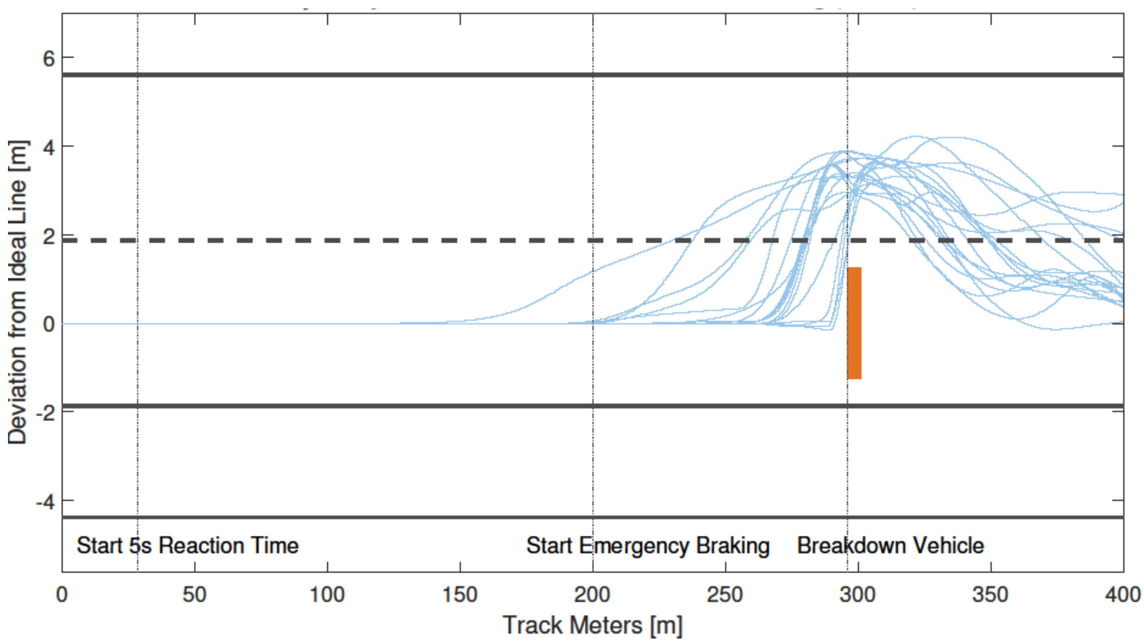


Figure 5-35: Ego-vehicle's driving trajectory of the L2H-on group during the cut-out scenario. Each participant is plotted as a single line (n = 19).

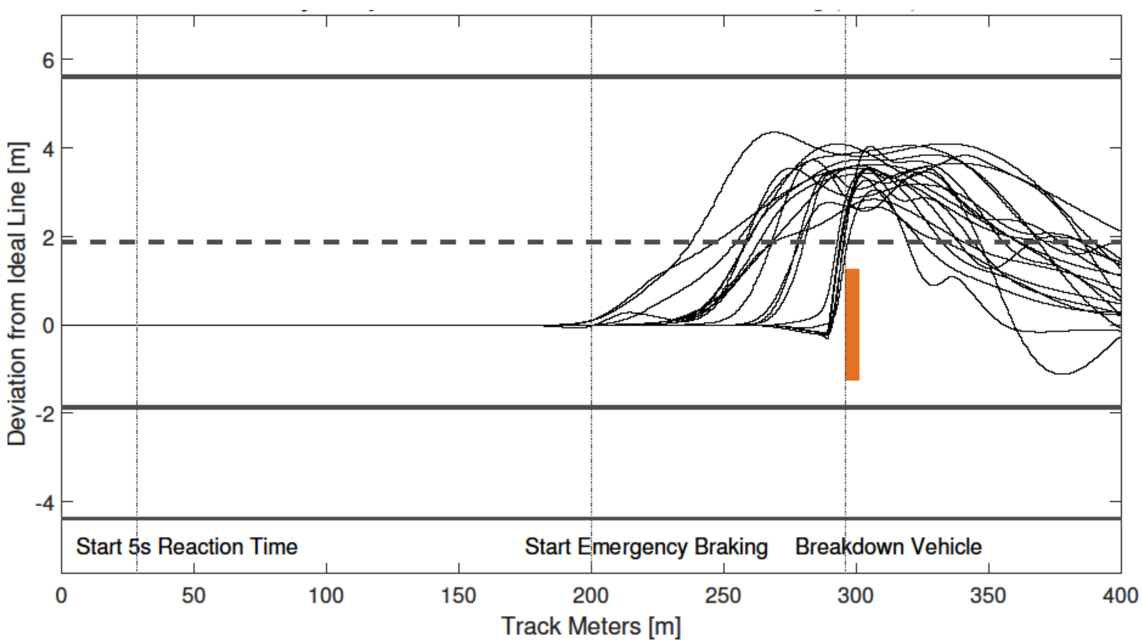


Figure 5-36: Ego-vehicle's driving trajectory of the L2H-off group during the cut-out scenario. Each participant is plotted as a single line (n = 19).

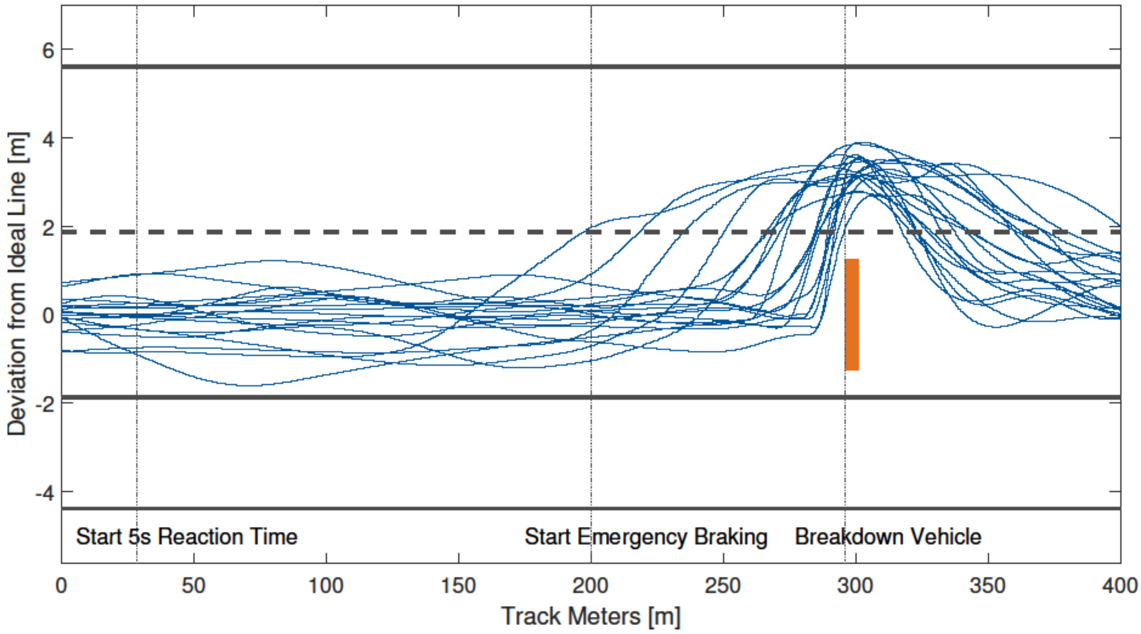


Figure 5-37: Ego-vehicle's driving trajectory of the L0 group during the cut-out scenario. Each participant is plotted as a single line (n = 19).

Maximum lateral acceleration

The maximum lateral acceleration showed a significant difference in the roadworks scenario based on the conducted Welch ANOVA ($F(2, 54) = 5.37$, $p = .007$, $\eta_p^2 = .17$). Results of the maximum lateral acceleration (calculated from the start to the end of the scenario) are presented in Figure 5-38 and Table 5-24.

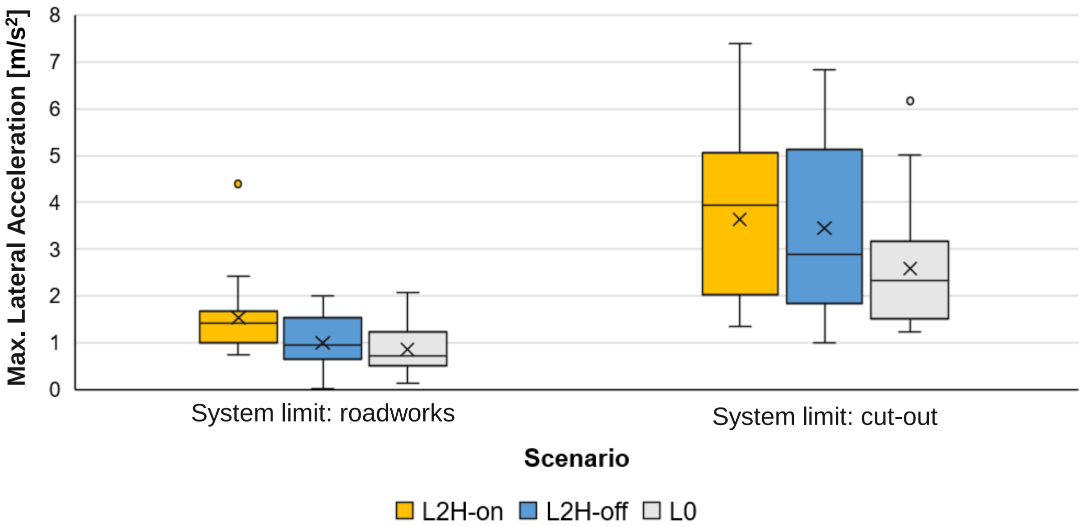


Figure 5-38: Boxplots showing maximum lateral acceleration during the roadworks (L2H-on: n = 19, L2H-off: n = 18, L0: n = 20) and cut-out scenarios (L2H-on: n = 19, L2H-off: n = 18, L0: n = 19).

Post-hoc test revealed a significant difference between L2H-on and L2H-off and a significant difference between L2H-on and L0 (Table 5-25). L2H-on led to higher maximum lateral acceleration. No significant difference was found between L2H-off and L0. The lateral maximum acceleration in the cut-out scenario showed no significant differences depending on the driving mode ($F(2, 33.80) = 2.76$, $p = .077$). In general, the cut-out scenario resulted in overall higher maximum lateral acceleration values compared to the road work scenario. The results should only be interpreted relatively, since no lateral accelerations are perceptible in a static driving simulator.

Table 5-24: Descriptive results for maximum lateral acceleration during the roadworks and cut-out scenarios.

Driving Mode	***M***	***SD***
	Roadworks	
L2H-on (n = 19)	1.53 m/s^2	0.83 m/s^2
L2H-off (n = 18)	1.00 m/s^2	0.63 m/s^2
L0 (n = 20)	0.86 m/s^2	0.51 m/s^2
	Cut-Out	
L2H-on (n = 19)	3.64 m/s^2	1.69 m/s^2
L2H-off (n = 18)	3.45 m/s^2	1.92 m/s^2
L0 (n = 19)	2.59 m/s^2	1.26 m/s^2

Table 5-25: Post-hoc comparison for maximum lateral acceleration for scenario roadworks between L2H-on (n = 19), L2H-off (n = 18) and L0 (n = 19).

Driving Mode		p_{Holm}
L2H-off	L2H-on	.037
L2H-off	L0	.536
L2H-on	L0	.009

Incidents/Accidents

There were no incidents, such as accidents or other highly critical situations. In this context, it should be mentioned that the emergency braking assistant might have influenced this and thus prevented potential longitudinal collisions.

5.2.4 Conclusions

In the following, the former mentioned specific results are discussed on a more abstract level to provide insights regarding the research questions of this study (CQ1) and the project's other CQs.

CQ1: Hands-off = Mind off?

The implemented L2H-off function in Study 1 does not lead to mind off. Although a higher number of eyes-off road glances above 2 s was found in L2H-off compared to L2H-on, this was nevertheless associated with a higher visual attention ratio to the road in L2H-off. According to this, participants in L2H-off looked away more often, but for lesser time at a single long glance, which reduced the risk of overlooking potential hazards compared to L2H-on. The high number of eyes-off road glances with L2H-off also led to consequently high numbers of DMS warnings at L2H-off compared with L2H-on, but with the effect of returning attention to the road earlier at L2H-off than at L2H-on. The self-reported monitoring performance is not decreased by the use of a L2 function compared to L0 but may even be increased.

CQ2: Prolonged transition times

The results of Study 1 give no evidence that L2H-off leads to prolonged transition times compared to L2H-on in transition relevant scenarios. Reaction time to warnings and the time to direct control in the system limit scenarios showed no differences between L2H-off and L2H-on. Accordingly, this was also reflected in the number of emergency braking maneuvers during this scenario. Moreover, the comparison of lane changes show that participants might be able to anticipate and react to certain scenarios faster in L2H-off than in L2H-on leading to differences in the driving trajectories, minimum TTCs and maximum lateral accelerations in favor of L2H-off.

CQ3: Foreseeable misuse

According to the results of Study 1, L2H-off does not produce a higher level of foreseeable misuse. Participants did not engage more with the NDRT than at L0, while compared to L2H-on an even lower engagement was observed. Thus, participants of L2H-off rather agreed that the warnings decrease the NDRT engagement, while L2H-on participants tended to disagree with this statement.

CQ4: Mode confusion

No evidence was found that L2H-off led to a higher level of mode confusion as there were no differences between L2H-off and L2H-on regarding the attempted activations although L2 was not available, and no differences regarding the subjective system and role understanding. In general, the level of mode confusion can be described as low in Study 1.

CQ5: Safety

We found no accidents or highly safety-critical incidents in Study 1. The participants regardless of L2H-off, L2H-on and L0 benefited from the existing emergency braking assistant. No effects were recognizable in this study according to which L2H-off would lead to a greater safety risk than in L2H-on or L0. In particular, scenarios that cannot be anticipated and become critical appear to constitute a challenge for all drivers independent of the driving mode.

5.2.5 References

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552–557. doi:10.7334/psicothema2016.383

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J. …& Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (No. DOT HS 812 182). Washington, DC.

Franke, T., Attig, C., & Wessel, D. (2019). A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human–Computer Interaction*, 35(6), 456–467. doi:10.1080/10447318.2018.1456150

International Organization for Standardization (2013). *Road vehicles —Measurement of driver visual behaviour with respect to transport information and control systems — Part 2: Equipment and procedures.* (ISO/TS Standard, 15007-2). Retrieved from https://www.iso.org/standard/56622.html

Körber, M. (2019). Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (Vol. 823, pp. 13–30). Cham: Springer International Publishing. doi:10.1007/978-3-319-96074-6_2

Kraft, A.-K., Naujoks, F., Wörle, J., & Neukum, A. (2018). The impact of an in-vehicle display on glance distribution in partially automated driving in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, 52, 40–50. doi:10.1016/j.trf.2017.11.012

Kurpiers, C., Lechner, D., & Raisch, F. (2019). *The Influence of a Gaze Direction Based Attention Request to Maintain Mode Awareness* (26th International Technical Conference on the Enhanced Safety of Vehicles (ESV): Technology: Enabling a Safer Tomorrow). Eindhoven Netherlands. Retrieved from https://www-esv.nhtsa.dot.gov/Proceedings/26/26ESV-000245.pdf

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car. In A. L. Kun (Ed.), *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12* (51–58). New York, New York, USA: ACM Press. doi:10.1145/2390256.2390264

Transportation Research Board (2000). *Highway capacity manual.* Washington, DC, 2(1).

United Nations Economic Commission for Europe (2018). *Uniform provisions concerning the approval of vehicles with regard to steering equipment.* UN Regulation No. 79 Retrieved from https://unece.org/transport/documents/2021/02/standards/un-regulation-no-79-rev4-amend3

5.2.6 Appendix

5.2.6.1 Pre-Questionnaire

Beschreibung	Frage	Antwortformat
Datenzuordnung		
Probanden-code	Bitte generieren Sie Ihren persönlichen Versuchsperso-nen-Code für die Studie. Dieser Code besitzt den Vor-teil, dass Sie den Code mittels der Fragen jederzeit neu generieren können, außenstehende Dritte jedoch kaum. Wir benötigen diesen Code, um Ihre Daten der Vorbefra-gung mit den Daten der Versuchsfahrt zu verknüpfen.	[…]
Name	Bitte geben Sie Ihre Kontaktdaten an. Diese Daten dienen ausschließlich der Kontaktauf-nahme nach Zuordnung zu einer Versuchsgruppe (A o-der B). Die Daten werden getrennt von den weiteren im Fragebogen erhobenen Daten aufbewahrt und mit Ab-schluss der Datenerhebung gelöscht.	[…]
Soziodemographisch		
Alter	Bitte geben Sie Ihr Alter in Jahren an.	[…]
Geschlecht	Bitte geben Sie Ihr Geschlecht an.	• Männlich • Weiblich • Divers • k.A.
Haendigkeit	Welche Hand bevorzugen Sie bei alltäglichen Verrich-tungen (z.B. eine Schere benutzen)?	• Rechts • Links • Kein Unterschied
Sehschwaech e	Benutzen Sie beim Autofahren eine Sehhilfe?	• Ja, ich benutze Sie auch während des Versuchs (Brille/Kontaktlinsen). • Ja, ich benutze Sie je-doch nicht während des Versuchs. • Nein
Far-bfehlsichtigkeit	Liegt bei Ihnen eine Farbfehlsichtigkeit vor?	• Ja, rot-grün Sehschwä-che • Ja, blau-gelb Sehschwä-che • Nein
Ho-erschwaeche	Liegt bei Ihnen eine Hörschwäche vor? Wenn ja, ist diese korrigiert?	• Ja, sie wird auch wäh-rend des Versuchs korri-giert.

		• Ja, sie wird während des Versuchs jedoch nicht korrigiert. • Nein
Kontext Fahren		
Fuehrerschein	In welchem Jahr haben Sie Ihren Pkw-Führerschein gemacht?	[…]
Fahrtfrequenz	Wie oft sind Sie in den letzten zwölf Monaten im Durchsch	• Täglich • Mehrmals pro Woche • Mehrmals pro Monat • Weniger als einmal pro Monat • (Selten bis) Nie
Fahrtstrecke	Wie viele Kilometer sind Sie in den letzten zwölf Monaten circa mit dem Auto gefahren?	• 0 km (keine Fahrt) • 1 km - 5.000 km • 5.001 km - 20.000 km • 20.001 km - 50.000 km • 50.000 km - 100.000 km • Mehr als 100.000 km
FrequenzAutobahn	Wie oft sind Sie in den letzten zwölf Monaten im Durchschnitt auf Autobahnen Auto gefahren?	s. Fahrtfrequenz
StreckeAutobahn	Wie viele Kilometer sind Sie in den letzten zwölf Monaten circa mit dem Auto auf Autobahnen gefahren?	s. Fahrtstrecke
KenntnisAS	Welche Erfahrungen haben Sie persönlich gesammelt mit: 1. Tempomat (CC) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit.] 2. Abstandsregeltempomat (ACC) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Abstand zum Vorderfahrzeug ein.] 3. Aktiver Spurhalteassistent [Dieses System erkennt die Fahrstreifenbegrenzungen und hält das Fahrzeug in den Begrenzungen.] 4. Stauassistent [Dieses System regelt die Geschwindigkeit und den Abstand zum Vorderfahrzeug im Stau und hält dabei das Fahrzeug auf dem Fahrstreifen] 5. Park Assist [Dieses System übernimmt während des Einparkvorgangs nur die Lenkbewegung.] 6. Teilautomation (L2) [Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Abstand zum	• Unbekannt • bekannt, aber nie benutzt • selten genutzt • regelmäßig genutzt
FiltHerstellerCC	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Tempomat (CC)	• BMW • VW • Mercedes • Audi • Tesla • Weitere Marken • […]

FiltHerstellerACC	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Abstandsregeltempomat (ACC)	s. FiltHerstellerCC
FiltHerstellerSpurha	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Aktiver Spurhalteassistent	s. FiltHerstellerCC
FiltHerstellerStauAs	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Stauassistent	s. FiltHerstellerCC
FiltHersteller-ParkAs	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Park Assist	s. FiltHerstellerCC
FiltHerstellerTeilau	Geben Sie bitte für das jeweilige System an, von welcher/n Automarke(n) Sie das System kennen: Teilautomation (L2)	s. FiltHerstellerCC
Fahrstil	DSQ [15 Items; Uebersetzung durch fka+ika+LfE]	French et al., 1993
Technikaffinitaet	ATI-S [9 Items]	Franke et al., 2019

5.2.6.2 Follow-up-Questionnaire

Beschreibung	Frage	Antwortformat
Metadaten		
VP	VP-Nummer	[…]
Fahrt	Gib an, zu welcher Fahrt diese Befragung geführt wird.	• L0 • L2H-on • L2H-off
Probanden-code	Bitte generieren Sie Ihren persönlichen Versuchspersonen-Code für die Studie. Dieser Code besitzt den Vorteil, dass Sie den Code mittels der Fragen jederzeit neu generieren können, außenstehende Dritte jedoch kaum. Wir benötigen diesen Code, um Ihre Daten der Vorbefragung mit den Daten der Versuchsfahrt zu verknüpfen.	[…]
Subjektive Metriken --- nur L2H-on und L2H-off		
Vertrauen	TiA Körber [19 Items]	Körber, 2019
Akzeptanz	CTAM [Subskalen: Performance expectancy (-PE2); Effort expectancy; Attitude towards using technology; Facilitating conditions (-FC4); Behavioral intention to use the system; Perceived safety ausgeschlossen: Subskalen: Anxiety; Self-Efficacy; Social Influence; Auswahl und Uebersetzung durch fka+ika+LfE]	Osswald et al., 2012
Systemverstaendnis	19 Items [Zusammenstellung von LfE] • Das System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer. • Das System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden. • Ich muss das system stets überwachen, wenn das System aktiviert ist. • Wenn das System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit. • Ich darf mich mit fahrfremden Tätigkeiten wie z.B. E-Mails schreiben beschäftigen, wenn das System aktiviert ist. • Der Fahrer muss das System bewusst aktivieren. • Das System passt die Geschwindigkeit an die des vorausfahrenden Fahrzeuges an. • Das System kann Fahruntauglichkeit durch Müdigkeit des Fahrers ausgleichen. • Der Fahrer darf sich von der Überwachung des Verkehrsraums abwenden, wenn das System aktiviert ist und eine andere Person im Fahrzeug diese Aufgabe für mich übernimmt.	• Nicht zutreffend • Zutreffend • Unsicher

	• Der Fahrer darf sich von der Überwachung des Verkehrsraums abwenden, wenn das System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind. • Das System erkennt immer, wenn es eine Situation nicht meistern kann. • Das System lenkt automatisch. • Der Fahrer muss innerhalb von Sekunden die Fahraufgabe übernehmen können. • Ich muss auch bei aktivierter Automation zu jedem Zeitpunkt wach bleiben. • Ich sollte niemals unaufgefordert in die Automation eingreifen. • In der folgenden Situation kann es passieren, dass das System die Situation nicht richtig einschätzen kann und der Fahrer eingreifen muss: Es sind Schlaglöcher auf der Straße, die das Erkennen der Fahrbahnmarkierung erschweren. • Wenn das Level 2 System aktiv ist, hält es einen voreingestellten Mindestabstand zum vorausfahrenden Fahrzeug ein. • Wenn das Level 2 System aktiv ist, steuert es die Geschwindigkeit des Fahrzeugs entsprechend der eingestellten Geschwindigkeit. • Wenn das Level 2 System aktiv ist, führt es Fahrstreifenwechsel durch, ohne dass der Fahrende dabei selbst lenken muss.	
Bewertung Warnungen	4 Items [Zusammenstellung von fka+ika+LfE] Bitte geben Sie an, inwiefern Sie den Aussagen zustimmen. • Die Warnungen kamen zu häufig. • Bei einer L2 automatisierten Fahrt würde ich mich mit dem Warnsystem sicherer fühlen als ohne das Warnsystem. • Ohne die Warnungen würde ich mich mehr mit fahrfremden Tätigkeiten (z.B. am Tablet spielen, Handy bedienen, trinken, etc.) beschäftigen. • Ich empfand die Warnungen als lästig.	• 1 Stimme überhaupt nicht zu • 2 • 3 • 4 • 5 • 6 • 7 Stimme voll zu
Reaktion auf Warnungen	4 Items [Zusammenstellung von fka+ika+LfE] Bitte geben Sie an, wie oft Sie auf folgende Aussagen reagiert haben. • Ich konnte nachvollziehen, warum eine Warnung ertönt ist. • Die richtige Reaktion nach Auftreten einer Warnung war mir klar. • Ich habe die Warnungen bewusst ignoriert. • Die Warnungen lenkten meine Aufmerksamkeit wieder auf das Fahrgeschehen.	• 1 Nie • 2 • 3 • 4 • 5 • 6 • 7 Immer
Subjektive Metriken --- all Gruppen		
NDRTs	8 Items [Zusammenstellung von fka+ika+LfE]	• Nie • Sehr selten • Selten

	Stellen Sie sich vor, Sie wären in Ihrem privaten Fahrzeug. Bitte geben Sie an, wie oft Sie sich mit folgenden Tätigkeiten beschäftigen würden, wenn Sie auf der Autobahn fahren *(und das System aktiviert ist)*. • NDRT1 Handy oder ähnliches Gerät (Laptop, externes Navi, Tablet, …) in der Hand – Bedienung SMS/WhatsApp Nachrichten verfassen oder lesen; Browsing; … • NDRT2 Handy oder ähnliches Gerät (Tablet, …) in der Hand - Sprechen Telefonieren ohne Freisprechanlage; Aufnehmen von Sprachnachrichten; ... • NDRT3 Handy oder ähnliches Gerät (Tablet, …) fest installiert bzw. mit Freisprechanlage verbunden - Sprechen Telefonieren mit Freisprechanlage; aufnehmen von Sprachnachrichten über Sprachbefehle; ... nicht: Unterhaltung mit Beifahrern, also Personen im Fahrzeug befinden • NDRT4 Bedienung von Systemen im Fahrzeug (nicht direkt relevant für die Fahraufgabe) Bedienen der integrierten Navigation; Einstellungen im Infotainmentsystem vornehmen; Verstellen des Sitzes; Einstellen der Klimaanlage; … • NDRT5 Essen/Trinken/Rauchen Öffnen einer Dose; Essen eines Apfels; Anzünden einer Zigarette; … • NDRT6 Körperpflege/ Make-Up/... Frisieren; Make-Up; Nagelpflege; … Nicht: kurze, unbewusste Handlungen (z.B. kratzen) • NDRT7 Interaktionen mit Beifahrern Unterhalten mit Beifahrer; Gestikulieren vor Beifahrer; Blicke zum Beifahrer; … • NDRT8 Suchen; Greifen; Kramen; … Suchen nach Objekt(en) und hingreifen, z.B. in einer Tasche	• Oft • Sehr oft
MatrixNDRTs-Freitext	Freitextfeld	[…]
SubjUeberwachungsguete	Wie aufmerksam waren Sie bei der Fahrt? / Wie aufmerksam haben Sie die Systemleistung überwacht, wenn Sie L2 aktiviert hatten?	• 0 Unaufmerksam • 1 • 2 • 3 • 4 • 5 • 6 Stets aufmerksam
SubjEinflussSetting	Hätten Sie sich anders verhalten, wenn Sie die letzte Fahrt nicht im Rahmen einer Studie durchgeführt hätten? Z.B. Beschäftigung mit fahrfremden Tätigkeiten oder Ähnliches?	• Ja • Nein

SubjEinflussSetting-Komm	Sie haben die vorige Frage mit "ja" beantwortet. Bitte beschreiben Sie kurz, inwiefern Sie sich anders verhalten hätten.	[…]
Kommentare	Haben Sie Kommentare zu der heutigen Fahrt bzw. dem erlebten System?	[…]
SonstKommentare	Haben Sie sonstige Kommentare oder Anmerkungen zur Studie?	[…]

5.2.6.3 Interview

Beschreibung	Frage	Antwortformat
Metadaten		
VP	VP-Nummer	[…]
Fahrt	Gib an, zu welcher Fahrt diese Befragung geführt wird.	• L0 • L2H-on • L2H-off
VL-Protokoll		
L0VP	L0 Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt	[…]
L0Fzg	L0 Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, ...	[…]
HonVP	L2H-on Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt	[…]
HonFzg	L2H-on Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, ...	[…]
HoffVP	L2H-off Anmerkungen zum Probanden, z.B. während Einweisung oder Fahrt	[…]
HoffFzg	L2H-off Anmerkungen zum Fahrzeug, z.B. Abwurf ohne Grund, Fehlermeldung, …	[…]
VLProt-Sonst	Sonstige Anmerkungen	[…]
Interview-Leitfaden		
FahrerTrans	Fahrerinitierte Transitionen Während der Autobahnfahrten sind Sie einen Großteil der Zeit mit L2 gefahren. In verschiedenen Situationen - zum Beispiel bei einer Baustellensituation - haben Sie das System aktiviert oder deaktiviert, um damit in einen anderen Fahrmodus wechseln. Gab es hierbei Situationen oder Aspekte, über die Sie uns gerne mehr erzählen würden? Für manuelle Fahrt: Gab es kritische Situationen oder Momente, über die Sie uns gerne mehr erzählen würden?	[…]
SystemTrans	Nur HandsOn und HandsOff Systeminitiierte Transitionen Während der Autobahnfahrten sind Sie einen Großteil der Zeit mit L2 gefahren. In verschiedenen Situationen hat das System von sich L2 teilweise oder komplett deaktiviert oder Sie zu einer Übernahme aufgefordert.	[…]

	Gab es hierbei Situationen, über die Sie uns gerne mehr erzählen würden?	
DMS	Nur HandsOn und HandsOff DMS Wenn L2 aktiv war, gab es ein Fahrerbeobachtungssystem, das gewarnt hat, wenn Sie ON: die Hände zu lange vom Lenkrad genommen haben oder OFF: zu lange von der Straße weggesehen haben. Möchten Sie dieses Fahrerbeobachtungssystem noch einmal genauer kommentieren oder bewerten oder Ihre Erlebnisse mit dem Fahrerbeobachtungssystem beschreiben?	[…]
HMI	HMI Für die Bedienung von L2 haben Sie verschiedene Tasten am Lenkrad verwendet. Je nach Systemzustand wurden Ihnen im Anschluss unterschiedliche Anzeigen im Kombi-Display angezeigt. Möchten Sie allgemein die Bedienung des L2 Systems oder die verwendeten Anzeigen noch einmal genauer kommentieren bzw. bewerten? Für Manuelle Fahrt: Möchten Sie allgemein die verwendeten Anzeigen im Kombi-Display noch einmal genauer kommentieren bzw. bewerten?	[…]
Geschwin digkeit	Während der Fahrt gab es verschiedene Geschwindigkeitsbeschränkungen. Möchten Sie das Fahrverhalten des Systems diesbezüglich noch einmal genauer kommentieren bzw. bewerten oder Ihre Erlebnisse beschreiben?	[…]
Systemverhal ten	Nur HandsOn und HandsOff Systemverhalten Während der Nutzung von L2 haben Sie die aktive Spurführung, die Geschwindigkeitskontrolle, das Halten des Abstandes und die Interaktion mit anderen Verkehrsteilnehmern in verschiedenen Situationen erlebt. Möchten Sie dieses Fahrverhalten des Systems noch einmal genauer kommentieren bzw. bewerten oder Ihre Erlebnisse beschreiben?	[…]
L2PrivNut zung	Nur HandsOn und HandsOff Bewertung L2 In diesem Versuch konnten Sie Erfahrungen mit L2 sammeln. Würden Sie dies auch gern in Ihrem privaten Fahrzeug nutzen?	• Nein • Eher nein • Unsicher • Eher ja • Ja • keine Angabe • nicht anwendbar (L0)

L2Kom-ponenten	Nur HandsOn und HandsOff Bewertung L2 Welche Komponente von L2 würden Sie auch gern in Ihrem privaten Fahrzeug nutzen? [3 Items] • Längsführung/ACC • Querführung/Spurhaltung/Lenken • *Nur für Hands-off:* H-off/Hände frei nutzen können	• Nein • Eher nein • Unsicher • Eher ja • Ja • Nicht anwendbar [Spalte nur für A-2 H-on Interview & Frage 3]
L2Bewer-tungKomm	Nur HandsOn und HandsOff Bewertung L2 Anmerkungen zu den vorigen Fragen/Antworten vom Probanden zur L2-Bewertung.	[…]
Sonstiges	Nur HandsOn und HandsOff Haben Sie noch weitere Kommentare oder Anmerkungen zu der erlebten Fahrt mit L2?	[…]
Studie	Haben Sie noch weitere Kommentare oder Anmerkungen zu der Studie?	[…]

5.3 Experimental Study 2

Documentation by P. Dautzenberg (Institut für Kraftfahrzeuge, RWTH Aachen University)

Study 2 focuses on the research subject “Mode Confusion” (CQ4). Mode confusion means that the driver has lost track of the currently active automation mode. The likelihood for mode confusion thereby further increases if the system or alternating system modes appear similar to the user.

Mode confusion is one possible reason for deficient mode awareness. Mode awareness combines two major aspects (Boos, Feldhütter, Schwiebach, & Bengler, 2020; Kurpiers, Biebl, Mejia Hernandez, & Raisch, 2020): Knowledge-based and behavior-based confusion. The first aspect implies the knowledge about which mode is currently active and the knowledge about the function’s abilities and limits, as well as the tasks and roles as driver (knowledge-based confusion). Understanding the L2 function and its limitations as well as understanding one's own tasks when interacting with the function appear to be essential prerequisites for mode awareness. The second aspect of mode awareness implies mode compliant behavior (behavior-based confusion). One important aspect of mode confusion is that the driver assumes that the vehicle performs a task (longitudinal and/or lateral guidance) that it does not actually perform. In this case the driver mistakenly relies on the vehicle being able to solve certain tasks or challenges without active driver input. The corresponding actions, which are appropriate for the erroneously assumed automation but inappropriate for the actual, currently active mode are unintentional, so called, mode errors.

As described in Section 2.4, there is only little to no research on the occurrence of mode confusion in direct comparison of or when switching between different L2 function designs (L2H-on and L2H-off). One possible concern is that individuals may have a reduced awareness of their responsibilities when using an L2H-off function, as they have less contact with the steering wheel and may thus feel less directly involved in the driving task. Furthermore, we assumed that mode confusion might not be addressed sufficiently by an adapted DMS design alone. Therefore, more input is needed to assess whether hands-free use of L2 functions increases mode confusion and if so, which countermeasures can be implemented on the function design side, to counteract this.

In fact, some literature provides guidance on what design principles might be helpful in reducing the risk of mode confusion. For example, functions with clear-cut modes, i.e. providing either both lateral and longitudinal assistance (*on*) or neither lateral nor longitudinal assistance (*off*), should increase mode awareness/decrease mode confusion as there are less transitions, which the driver will go through (Consumer Reports, 2020). Furthermore, there are indications that systems providing gaze-based attentiveness requests should increase mode awareness or decrease mode confusion (Kurpiers et al., 2020). These assumptions were examined by means of the studies with a focus on mode confusion in Study 2.

5.3.1 Research Questions & Hypotheses

Study 2 aimed to answer three main research questions (RQ):

- RQ1: Are there differences between L2H-on and L2H-off functions with regard to mode confusion?
- RQ2: Do clear-cut transitions (L0 - L2 and vice versa) increase mode awareness/decrease mode confusion compared to successive/multi-step transitions (L0 – L1 – L2 and vice versa)?
- RQ3: Which transitions (L0 - L1; L0 - L2; L1 – L2 and vice versa) are especially prone to mode confusion?

Regarding RQ1, we assessed whether people using L2H-off functions are more prone to mode confusion, based on the concerns mentioned beforehand. Regarding RQ2, it was assumed that clear-cut modes of either on or off should increase mode awareness/decrease mode confusion, as there are less transitions the driver may go through and thus less variability in levels of direct control. When using a clear-cut system, drivers need to either execute longitudinal and lateral control themselves or to supervise the execution of both by the L2 functionality. The relevant task is therefore either executing or supervising. When using a multi-step system, the driver also has to likewise execute manual driving or supervise L2 automation. However, when driving in the additional L1 mode, drivers partly execute and supervise control, e.g. they have to supervise lateral guidance him/herself and supervise longitudinal guidance as performed by ACC. The third research question was assessed mostly exploratory. Insights from the Expert Study (see Section 4.2) suggest, that mode confusion may arise when switching between L1 and L2 modes due to the unclear level in lateral assistance being provided while longitudinal assistance levels remained the same.

5.3.2 Method

5.3.2.1 Participants

In total, *N* = 60 persons participated in the study. Data from two participants had to be excluded as they had to abort the experiment due to simulator sickness. Therefore, data from *N* = 58 participants (69% male) with an average age of *M* = 29 years (SD = 11, Range: 21–74 years) were considered in the further data analysis. Participants were randomly assigned to one of the three experimental groups. A description of the three groups can be found in Table 5-1.

Table 5-1: Description of the three experimental groups.

Group	Function	Sample size	Gender	Age	ACC experience	L2 experience
A	L2H-on multi	20	7 female 13 male	*M* = 25 *SD* = 5	60%	84%
B	L2H-off multi	19	5 female 14 male	*M* = 30 *SD* = 13	53%	32%
C	L2H-off clear	19	6 female 13 male	*M* = 32 *SD* = 14	32%	21%

To ensure that the survey results were not biased by a lack of affinity for technology, the Affinity for Technology Interaction Scale (ATI Scale; Franke, Attig, & Wessel, 2019) was applied. The results showed that there were no significant differences between the experimental groups and an overall high tendency regarding technology affinity.

5.3.2.2 Simulator and Mock-Up

Study 2 was conducted by ika in a static driving simulator. Simulation and mock-up were developed and supervised by fka. The projection system of the simulator consists of a circular screen surface with a diameter of five meters, covering an angular range of 220° around the driver. Onto this surface, a coherent image of the environment is projected by three Full HD projectors. The image for the outer rear view mirrors is created by two 55-inch monitors positioned behind the vehicle. The central rear view mirror is not used in this setup and has therefore been covered. The mock-up used for the study is a BMW i3 that had been modified for use in the driving simulator. A ten inch ultra-wide monitor serves as the driver display. The basic software running the simulator is Virtual Test Drive (VTD), whereas the vehicle dynamics component is an in-house development. The automated driving function has also been developed by fka and was specifically adapted for the studies within this project, as detailed below. This includes the incorporation of the DMS and its warning stages with visual and acoustic signals (see 5.3.2.4).

5.3.2.3 Level 2 Function Design

For the current study, a L2H-on and two L2H-off functions were implemented. To assess the assumptions of the study, we additionally varied whether the function operates according to a clear-cut or a multi-step principle (i.e., with or without an explicit ACC mode). When interacting with the multi-step function, the driver first has to activate L1 before L2 can be activated. Icons in the HMI indicate when the respective mode is available. If the driver brakes (threshold: brake pedal travel ≥ 10%), the L2 function is deactivated immediately and the function falls back to L0. From here on it takes two steps (i.e., activating L1 and subsequently activating L2) to get full assistance. If the driver steers (threshold: ≥ 5.7° difference the current steering wheel angle) the system falls back to L1. From here on it takes one step to get back to L2 automation. If however, a function direct control request (FDCR) is given or a DMS direct control request (DDCR), the system falls back to L1. From here on it takes only one step (activating L2) to get back to L2 automation. The clear-cut function comprises L0 and L2 only. There is no L1 functionality implemented or available. As with the multi-step function, brake interventions lead to an immediate deactivation of the L2 function and a fallback to L0. With both (clear-cut and multi-step) functions, steering input below the threshold for overruling the function or acceleration can be used to override the L2 function for a short duration. For all functions, thresholds for the deactivation of L2 automation were 5.7° steering input and 10% brake pedal travel. Set-speed was pre-selected and could not be adapted by the driver.

Within the current study, the L2H-on group and one of the L2H-off groups experienced the multi-step function. The second L2H-off group experienced the clear-cut function. This allows the comparison between L2H-on and L2H-off functions and the comparison between multi-step and clear-cut functions according to the research questions and respective assumptions (see 5.3.1).

5.3.2.4 DMS and HMI

The Driver Monitoring System (DMS) is realized using a head-mounted eyetracking device (Tobii Glasses Pro 3) in combination with a camera facing the driver. The gaze direction provided by the former is combined with the absolute head position detected through the latter, resulting in an absolute gaze vector. The warning system is calibrated to signal "eyes off road" whenever the gaze vector is pointing outside the windshield area. For the hands-on detection, a BMW steering wheel equipped with capacitive sensors is used in conjunction with a custom evaluation system adapted to the simulator setup. The timing and warning stages of the DMS were adapted from Study 1 (cf. Section 5.2.2.3.1). Figure 5-1 summarizes the warning cascades.

L2H-off	**Request Design**					
HMI	1. Warning Stage		2. Warning Stage		3. Warning stage	
	5 s	Eyes-on Request (visual + acoustic alert) „Aufmerksam bleiben“	+3 s	Hands-on Request (visual + acoustic alert) „Aufmerksam bleiben. Hände ans Lenkrad!“	+5 s	DMS direct controrl Request (visual + acoustic alert) „Fahrzeug bremst. Übernehmen!“
Driver‘s task		Eyes on road		Eyes on road + hands on wheel		Take-over driving task

L2H-on	**Request Alert Design**					
HMI	1. Warning Stage		2. Warning Stage		3. Warning stage	
	15 s	Hands-on Request (visual + acoustic alert) „Hände ans Lenkrad.“	20 s	Hands-on Request (visual + acoustic alert) „Hände ans Lenkrad.“	25 s	DMS direct controrl Request (visual + acoustic alert) „Fahrzeug bremst. Übernehmen!“
Driver‘s task		Hands on wheel		Hands on wheel		Take-over driving task

Figure 5-1: Warning cascades implemented and assessed in the current study.

An overview of the different HMI displays for each function is given in Appendix I for the L2H-off clear-cut function and in Appendix II for the L2H-on multi-step function.

5.3.2.5 Driving Scenarios

The driving duration was approximately 45 minutes. The highway consisted of two lanes each in both directions with a Level of Service B. To enable participants to experience the complexity of the multi-step function and thus to investigate the effect of functional complexity on mode confusion, it must be ensured that the participants fall back onto different levels (at least once on L1 and at least once on L0) throughout the study. Therefore, it is necessary to implement different scenarios that increase the likelihood for different intervention behaviors/deactivation paths. Participants experiencing the multi-step function should also experience at least one situation in which they fall back to L1 by L2 deactivation, but the ACC serving as L1 fallback reacts inappropriately to the situation at hand, in order to make the potential challenges of a multi-step function tangible. Based on these assumptions, scenarios of the following types were implemented:

- **Scenario Type A:** Scenarios in which participants must provide lateral guidance. Acceleration or braking is not necessary to cope with the situation (Scenario 1 and 2).
- **Scenario Type B:** Scenarios in which participants initially only need to perform lateral guidance, but in which the current ACC setting does not provide assistance appropriate to the situation, so that braking (longitudinal guidance) becomes necessary to successfully handle the situation (Scenario 3 and 4).
- **Scenario Type C:** Scenarios in which participants must perform braking/longitudinal guidance to successfully handle the situation (Scenario 5).

By implementing two scenarios for Scenario Type A and B each, the probability that the participants of the multi-step groups show the desired behavior (no longitudinal intervention or successive adoption of lateral and longitudinal guidance) should be increased. Figure 5-2 illustrates the sequence and timing of scenarios including the potential fallback levels after each system limit. Figure 5-3 gives a more detailed visualization of the five scenarios.

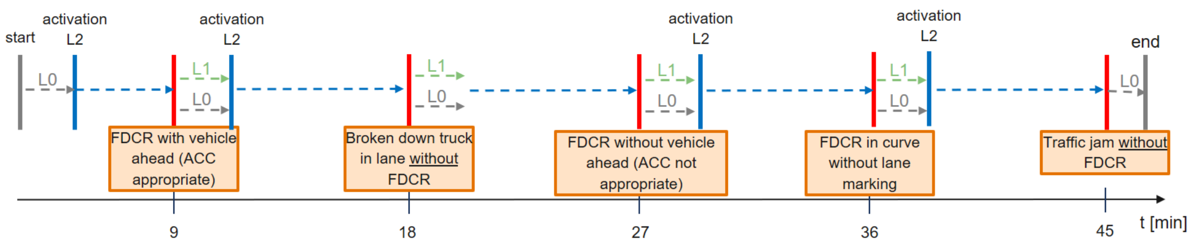


Figure 5-2: Experimental track incl. timing of scenarios, fallback possibilities and L2 availability.

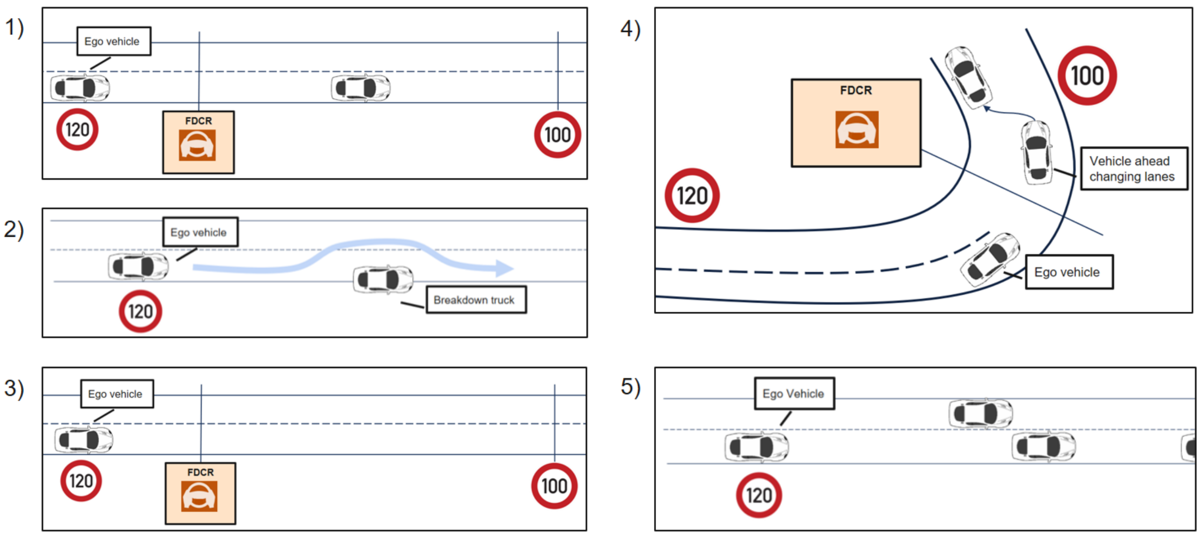


Figure 5-3: Overview on the scenarios of Study 2. The numbering corresponds to the presentation order in the study.

5.3.2.6 Study Design

A between-subject design with the factor *Level 2 Function Design* (L2H-on multi-step vs. L2H-off multi-step vs. L2H-off clear-cut) was applied. The variation of L2 functions resulted in three experimental groups. As dependent variables, both subjective questionnaire ratings as well as behavior data gathered via eyetracking and simulator were collected and assessed. Table 5-2 lists all subjective data and Table 5-3 the objective data collected and analyzed within Study 2.

Table 5-2: Subjective questionnaire data.

Questionnaires and single items		CQ

Mode confusion	• System and Driver role understanding (see
	• Appendix III)
	• Single item (Perceived System Complexity): How would you rate the overall complexity of the L2 function?
	• System Usability Scale (SUS; Brooke, 1996)
	• (Video-based) Interview questions:
	  o Do you remember this situation?
	  o Did the vehicle react as expected after you took over control in the corresponding situation?
	• After you took over control, was it clear to you which tasks you would have to perform and which tasks the vehicle would continue to perform?
		4

Trust in Automation	TiA (Körber, 2019) - subscales: • Understanding/Predictability • Trust in Automation	3, 5
Acceptance	CTAM (Osswald et al., 2012) – subscales: • Behavioral intention to use • Perceived safety	3, 5
DMS evaluation	8 items (see also Annex Study 1; *Bewertung Warnungen / Reaktion auf Warnungen*)	3, 4, 5

Table 5-3: Objective data assessed

Objective data		Unit	CQ
Eyetrack-ing	• Number of eyes-off road glances > 2s • Attention ratio: eyes-on road, eyes-on instrument cluster/steering wheel, eyes-on NDRT, eyes on other	• Number • Percentage	1
Simulator	• Number of HOR and EOR • ACC usage	• Number • Percentage	1
	• Hands-free driving while driving L2 • Number of attempted activations of L2, although not available	• Percentage • Number	4
	• Reaction time to HOR and EOR • Time to first input after FDCR • Hands-on time after FDCR • Number and description of incidents • Minimum frontal distance to object • Minimum lateral distance to object	• Time in s • Time in s • Time in s • Number • Distance in m • Distance in m	2, 5

5.3.2.7 Procedure

Participants were welcomed, asked to read and sign all legal documents, and informed about the procedure. Participants were then asked to answer demographic and mobility behavior related questions. After that participants were instructed to read a manual, explaining the L2 function. Subsequently, participants answered the items regarding system understanding. After that participants were brought to the driving simulator and started a familiarization drive (about 10 to 15 minutes) to get used to the simulator and the L2 function. The experimenter motivated participants to try different ways to override or deactivate the L2 function. Within the familiarization drive, participants also experienced an FDCR. After the familiarization drive the experiment drive (about 45 minutes) started. During the experimental drive, the five scenarios were presented as described beforehand. After the experimental drive, participants were shown videos of the five scenarios they experienced within the study on a tablet. They were interviewed regarding, if they could remember the respective situation and if the L2 function reacted according to their expectations. This interview was done to get more qualitative insight into potential mode confusion. Subsequently they were asked to answer the post-drive questionnaires (regarding system and driver role understanding, DMS evaluation, trust, and acceptance). Finally, participants were informed about the study goals, incentivized and seen off. The total duration of the experiment was about 1:30 h per participant.

5.3.2.8 Data Analysis

The data was analyzed using SPSS 27 (IBM Statistics). Normality distribution and equality of variances were analyzed using the Shapiro-Wilk test and the Levene's test. If these assumptions for parametric tests were not met, we used the non-parametric alternative. We performed ANOVA regardless of the violation of the normal distribution, as it appears to be robust to the violation of this condition (Blanca, Alarcón, Arnau, Bono & Bendayan, 2017). In case of violation of equality of variances, the Welch ANOVA was performed. If not reported otherwise, a one-way ANOVA was conducted for all subjective and objective measures. The significance level was set to $\alpha = .05$.

5.3.3 Results

5.3.3.1 Subjective Data

Mode confusion

<u>System & Driver Role Understanding</u>

Participants were asked to answer questions that would assess their understanding of the L2 function, its limits and functionality as well as their understanding of their role and responsibilities as the driver. This was done via the items already used in the FOT, Study 1 and items derived from the US Survey, extended by new study-specific items. The items were statements about the L2 function or driver tasks that participants could agree or disagree with (see Appendix I). Participants were asked to answer the items regarding system understanding before and after the test drive, to assess if the experience of the L2 function changed the understanding. The items regarding driver role understanding were assessed after the test drive only. These items were not answered before the test drive to avoid biasing participants.

For system understanding, two one-way ANOVAs were performed, one for each assessment point. Since additional, very specific items were queried after the experimental drive, the assessment points were not directly compared. For both assessment points, no main effect of function could be found (before experimental drive: $F(2, 55) = 1.15$, $p = .325$; after experimental drive: $F(2, 55) = 1.58$, $p = .215$), indicating that all three groups had a similar level of correct answers. Overall, the results indicate a good to very good system understanding (see Figure 5-4).

However, when looking at the results in detail, participants of both multi-step groups were seemingly uncertain about the item "If, while the Level 2 function is activated, the signal "Take over control" appears in the display, all assistance functions (including ACC) are cancelled and I have to drive myself as well as regulate speed and distance." (L2H-on multi: 35% correct answers; L2H-off multi: 52.63% correct answers). This result indicates that participants of both multi-step groups were not sure about what assistance level they would fall back onto after a FDCR, even after experiencing at least four FDRCs in the familiarization and experimental drive. When looking into the single item results of the L2H-off clear-cut group, the majority of participants of this group were rather uncertain about the item "There is an assistance mode in which the vehicle controls acceleration but the driver steers him/herself." (36.84% correct answers). This result indicates that participants of this group were seemingly unsure if there was a function similar to an ACC (L1) implemented in the vehicle.

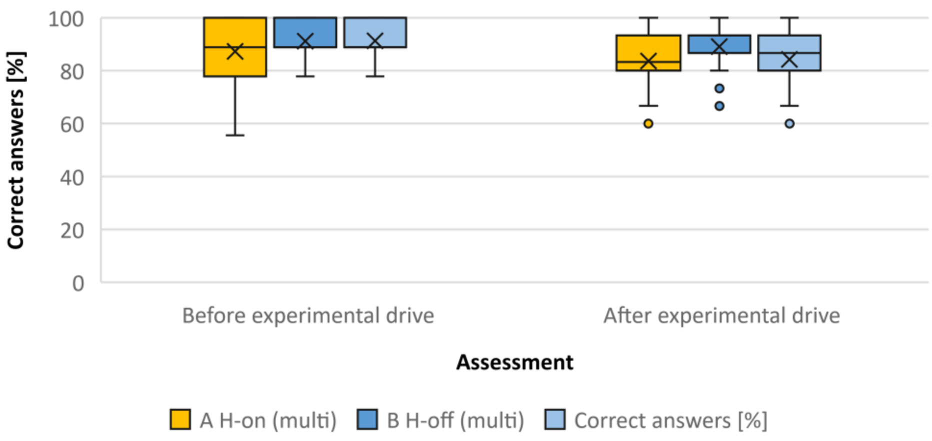


Figure 5-4: Boxplots for the proportion of correct answers for system understanding before and after the experimental drive.

For driver role understanding, a one-way ANOVA was conducted which revealed no significant differences between the three systems ($F(2, 55) = 0.36$, $p = .699$). Apart from very few outliers, the mean values indicate a flawless driver role understanding (see Figure 5-5).

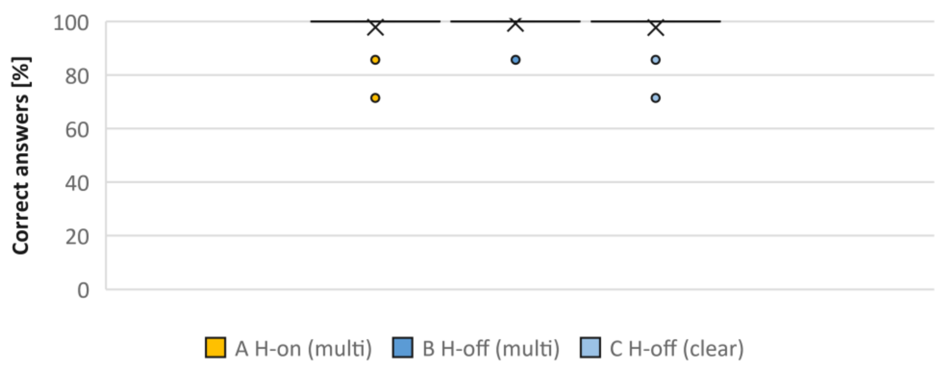


Figure 5-5: Boxplots for the proportion of correct answers for driver role understanding after the experimental drive.

Perceived Complexity

Participants were asked to rate the perceived complexity of the L2 function using a 7-point Likert scale single item. The one-way ANOVA revealed no significant differences between the three groups ($F(2, 55) = 2.68$, $p = .077$), indicating that all three L2 functions were assessed equally regarding complexity. Overall, mean values show that complexity was rated low to medium in all three groups (L2H-on multi-step: $M = 2.36$, $SD = 1.21$; L2H-off multi-step: $M = 3.53$, $SD = 1.70$; L2H-off clear-cut: $M = 3.67$, $SD = 1.88$)

System Usability Scale

To assess whether there are differences regarding usability between the three L2 function variants, the System Usability Scale (SUS, Brooke, 1996) was used. The conducted one-way ANOVA revealed no significant differences between the three groups ($F(2, 55) = 0.24$, $p = .791$), indicating that all three L2 functions are rated equally well regarding usability. The mean values furthermore indicate that usability ratings for all three function variants can be interpreted as excellent (for all M > 80.3; L2H-on multi-step: $M = 85.75$, $SD = 31.13$; L2H-off multi-step: $M = 86.58$, $SD = 7.32$; L2H-off clear-cut: $M = 84.08$, $SD = 12.86$).

Video-based interview

After the experimental drive, video clips of the five scenarios were presented to participants, to get deeper insights into potential scenario-specific mode confusion. First, participants were asked if they remembered the respective scenario. Scenarios 1, 2, 4 and 5 were remembered by all participants (100%). Regarding Scenario 3, the majority, but not all participants could remember the situation (L2H-on multi-step: 85%, L2H-off multi-step: 90%; L2H-off clear-cut: 84%). The participants who could not remember the scenario were excluded from further questions regarding this respective scenario. The remaining participants were asked whether the vehicle reacted as expected after participants took over direct control in the respective scenario. Across all scenarios and functions, the majority of participants indicated that the vehicle reacted according to their expectations (see Table 5-4), indicating that most participants didn't experience mode confusion during or after transitions in the respective situations. Descriptive data show that in Scenario 3, the vehicle seems to have most frequently failed to respond according to expectations. Participants of all three groups, who indicated that their expectations weren't matched most frequently explained that they expected the vehicle to automatically respond to the new speed sign (100 km/h). Furthermore, it was not clear to participants why the FDCR occurred within this scenario. The same explanation was given by some participants in Scenario 1. Participants who indicated that the vehicle didn't react as expected in Scenario 2 and 5 most frequently explained that they expected a FDCR in these situations or a function-triggered vehicle behavior (steering or braking) that would solve the situation. These comments might indicate an automation expectation mismatch in these situations.

Table 5-4: Proportion of participants who indicated that the vehicle responded according to expectations.

Group	*Scenario 1*	*Scenario 2*	*Scenario 3*	*Scenario 4*	*Scenario 5*
L2H-on (multi-step)	90 %	85%	75%	100%	80%
L2H-off (multi-step)	100%	95%	74%	95%	95%
L2H-off (clear-cut)	90%	79%	79%	95%	79%

Additionally, participants were asked if it was clear to them which tasks they had to perform after taking over direct control in the respective situation and which tasks the vehicle would continue to perform. Again, across all scenarios and functions, the majority of participants indicated that the distribution of tasks between them and the vehicle was clear to them (see Table 5-5), further indicating that most participants didn't experience mode confusion during or after transitions. Descriptive data show that participants seemed to be most uncertain in scenario 3. This finding goes in line with the finding to the latter question, where participants indicated most commonly that in scenario 3 the vehicle didn't react as expected.

Table 5-5: Proportion of participants who indicated that they understood the task distribution between them and the vehicle.

Group	*Scenario 1*	*Scenario 2*	*Scenario 3*	*Scenario 4*	*Scenario 5*
L2H-on (multi-step)	85%	90%	70%	90%	90%
L2H-off (multi-step)	84%	100%	90%	95%	90%
L2H-off (clear-cut)	79%	95%	79%	95%	79%

Trust

To assess participants trust in automation, the questionnaire from Körber (2019) was assessed and analyzed. Furthermore, an overall score was calculated and considered. The one-way ANOVA results show no significant differences between the three groups for all subscales, except for "Intention of Developers" ($F(2, 55) = 3.38$, $p = .041$, $\eta^2_p = .11$). Post-hoc tests revealed that the two L2H-off groups differed significantly ($p_{Tukey} = .039$) with the L2H-off multi-step group having significantly higher ratings regarding "Intention of Developers". It appears that the L2H-off multi-step group rated the developers as more trustworthy and more oriented towards the well-being of the users. Overall, mean scores are rather average to high across all groups for all subscales (see Table 5-6).

Table 5-6: Mean values and standard deviations for the trust subscales. Values could range between 1 = I don't agree at all and 5 = I fully agree.

Group	Overall		*Reliability/ Competence*		*Understanding/ Predictability*	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on (multi-step)	3.70	0.60	3.26	0.63	4.11	0.64
L2H-off (multi-step)	3.92	0.49	3.37	0.63	4.30	0.35
L2H-off (clear-cut)	3.58	0.58	3.36	0.83	4.01	0.74

Group	*Familiarity*		*Intention of Developers*		*Propensity to Trust*		*Trust in Automation*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
L2H-on (multi-step)	2.88	1.59	4.48	1.07	3.77	0.80	3.70	0.94
L2H-off (multi-step)	3.37	1.92	5.00	0.90	3.42	0.67	4.08	0.90
L2H-off (clear-cut)	2.32	1.63	4.18	0.96	3.70	0.91	3.89	0.91

Acceptance

To assess participants acceptance towards the respective L2 function, the subscales "Effort Expectancy", "Facilitating Conditions" and "Perceived Safety" from Osswald et al. (2012) were assessed and analyzed. The (Welch) ANOVA results show no significant differences between the three groups, neither for "Effort Expectancy" ($F(2, 55) = 0.56$, $p = .574$), nor for "Facilitating Conditions" ($F(2, 34.98) = 2.57$, $p = .091$) nor for "Perceived Safety" ($F(2, 55) = 0.19$, $p = .825$). Overall, mean scores are average to high across all groups for all three subscales (see Table 5-7).

Table 5-7: Mean values and standard deviations for the acceptance subscales. Values could range between 1 = I don't agree at all and 7 = I fully agree.

Group	*Effort Expectancy*		Facilitating conditions		Perceived Safety	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on (multi-step)	6.13	0.99	4.97	1.27	4.74	1.16
L2H-off (multi-step)	6.29	0.70	5.33	0.90	4.63	1.06
L2H-off (clear-cut)	6.38	0.53	5.67	0.67	4.54	0.69

DMS Evaluation

Participants were asked to evaluate the DMS they experienced throughout the study. This was done using the eight single items already used in Study 1. For each item a one-way (Welch) ANOVA was performed. The results show no significant difference between the three groups for any item.

5.3.3.2 Eyetracking Data

Number eyes-off road glances >2s

The one-way ANOVA revealed no significant difference for the number of eyes-off road glances above 2 s when driving L2 ($F(2, 47) = 1.31$, $p = .279$). The DMS based on visual attention in the L2H-off groups did thus not influence the number of long gazes averted from the road (see Figure 5-6). However, the first warning was only issued after 5 s of inattention to the road.

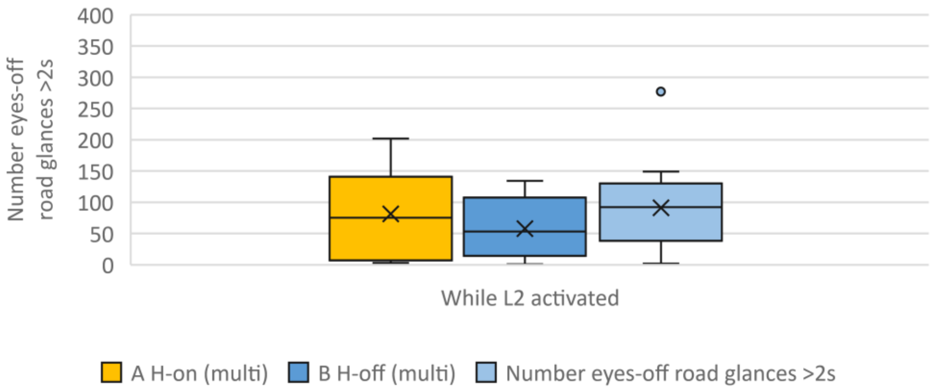


Figure 5-6: Boxplots for number of eyes-off road glances > 2s while driving with activated L2 function.

Visual attention ratio

The one-way (Welch) ANOVAs revealed no significant differences between the three groups regarding glance proportions for the three relevant areas of interest (road: $F(2, 31.68) = 0.87$, $p = .428$; instrument cluster: $F(2, 51) = 0.44$, $p = .645$; other: $F(2, 31.07) = 1.15$, $p = .329$) over the entire experimental drive (see Figure 5-7).

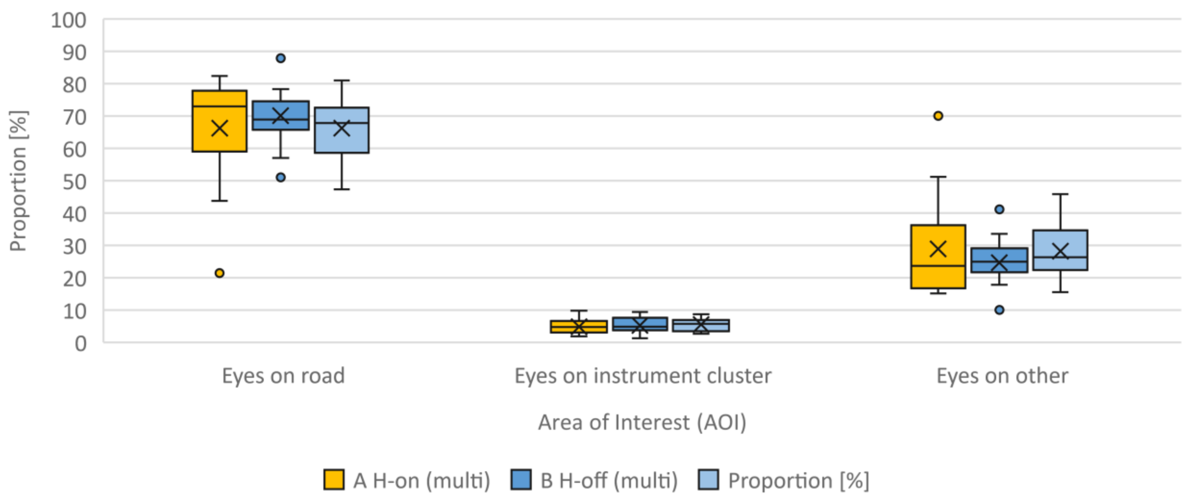


Figure 5-7: Boxplots for proportion of glances on the areas of interest road, instrument cluster/steering wheel and other.

5.3.4 Simulator Data

5.3.4.1 Overall

Number of H-off (HOR) and Eyes-off requests (EOR)

Statistically, only the two L2H-off groups were compared as the L2H-on group's DMS is too different (in terms of timing and requested behavior) from the L2H-off groups' DMS. However, Table 5-8 summarizes the mean values and standard deviations for all three groups. T-Tests revealed that there are significant differences between the two L2H-off groups regarding the mean number of warnings for warning stage 1 and warning stage 2 (stage 1: $t(22.63) = -2.87$, $p = .009$, $d = -.99$; stage 2: $t(16) = -2.38$, $p = .030$, $d = -.13$; stage 3: $t(16) = -1.38$, $p = .188$). The L2H-off clear-cut group experienced significantly more warnings in these two stages than the L2H-off multi-step group. However, when looking at the mean values, the warning frequency was overall rather low for all three groups (see Table 5-8).

One descriptive result is that there are very few “Warning stage 3” (DDCR) occurrences over all three groups. This shows that most participants reacted after the first or second warning stage. DMS with three-staged warnings appear to be appropriate for motivating users to redirect attention to the road or hands to steering control (driving with hands-on when using L2H-on functions and driving eyes-on when using L2H-off functions). Possibly, even 2 stages could be sufficient.

Table 5-8: Mean values and standard deviations for the number of DMS warnings.

Group	Warning stage 1		Warning stage 2		Warning stage 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on (multi-step)	1.05	1.78	0.32	1.00	0.00	0.00
L2H-off (multi-step)	1.33	1.33	0.00	0.00	0.00	0.00
L2H-off (clear-cut)	3.47	2.79	1.35	2.34	0.18	0.53

Mean reaction time to H-off (HOR) and Eyes-off requests (EOR)

A Welch ANOVA revealed that there are significant differences between the L2H-off multi-step group and the other two groups regarding their mean reaction times to HOR or EOR signals ($F(2, 27.03) = 12.98$, $p < .001$, $\eta^2 = .201$; see Figure 5-8). Post-hoc Games Howell t-tests reveal that the L2H-off multi-step group reacts significantly faster (L2H-on multi-step: $M = 2.41$, $SD = 1.91$; L2H-off multi-step: $M = 1.12$, $SD = 0.42$; L2H-off clear-cut: $M = 4.83$, $SD = 3.26$) than the L2H-off clear-cut group ($p < .001$) and the L2H-on multi-step group ($p = .047$).

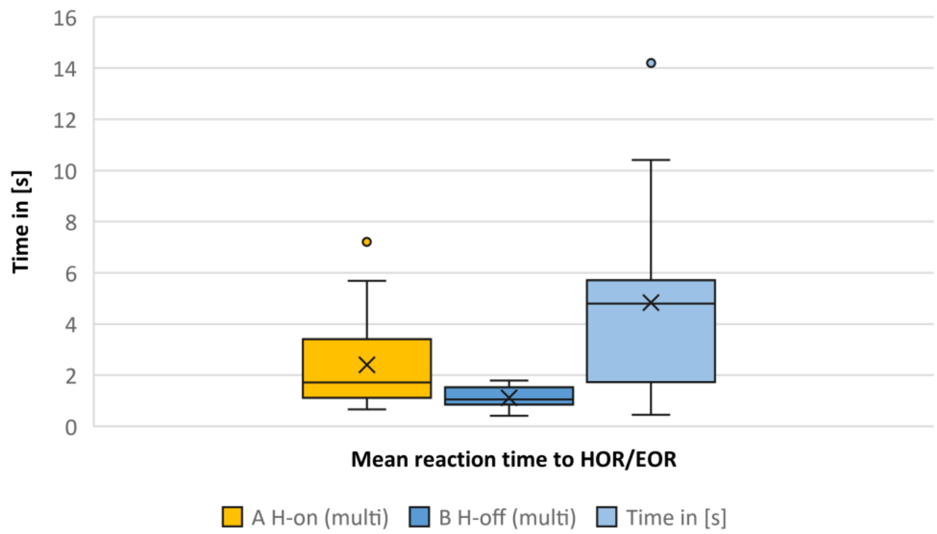


Figure 5-8: Boxplots showing the mean reaction time to HORs and EORs.

Time Hands-free driving

An ANOVA revealed a significant difference ($F(2, 53) = 94.55$, $p < .001$, $\eta_p^2 = .78$) between the proportion of hands-free driving while driving with an activated L2 function for the L2H-on multi-step group (M = 16.52%, SD = 15.77%), L2H-off multi-step group (M = 87.43%, SD = 24.29%), and L2H-off clear-cut group (M = 93.58%, SD = 16.35%). Post-hoc Tukey-corrected t-tests were conducted, which showed a significantly lower proportion for the L2H-on multi-step group than for the L2H-off multi-step group ($p < .001$) as well as the L2H-off clear-cut group ($p < .001$), but no significant difference between the L2H-off groups ($p = .594$).

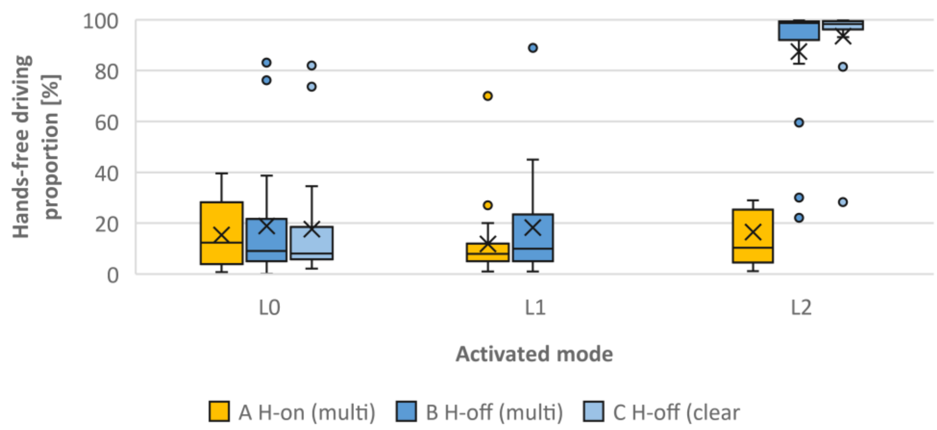


Figure 5-9: Boxplots showing the proportion of hands-free driving in the respective mode for the three experimental groups.

When looking at the L2H-on and L2H-off multi-step functions, data show that hands-free driving also occurred while driving manually (L2H-on multi-step group: M = 15.44, SD = 12.96%; L2H-off multi-step group: M = 18.89%, SD = 24.60%; L2H-off clear-cut group: M = 17.65%, SD = 22.65%; see Figure 5-9) and in L1 mode (L2H-on multi-step group: M = 11.89%, SD =

14.45%; L2H-off multi-step group: *M* = 18.28%, *SD* = 21.23%; see Figure 5-9). Having a closer look at the multi-step groups, data show that some participants took their hands off the steering wheel when driving L1. This could be an indicator for mode confusion. However, hands-free driving also occurred while driving manually (L0). This in turn suggests that this occurrence might probably not be an indicator for mode confusion, but either a normal behavior during driving or noise in the detection of hands-on steering wheel.

Number of attempted activations of L2, although not available

An ANOVA showed no significant difference ($F(2, 52) = 0.08$, $p = .928$) regarding the number of failed L2 activations between the three groups. Descriptively, the number of failed L2 activation attempts was very low for all three groups (L2H-on multi-step: *M* = 0.37, *SD* = 1.04; L2H-off multi-step: *M* = 0.28, *SD* = 0.67; L2H-off clear-cut: *M* = 0.28, *SD* = 0.75).

Number and description of collisions

No incidents occurred throughout the study over all groups.

5.3.4.2 Scenario 1, 3 and 4

Scenario 1, 3 and 4 had in common that an FDCR was issued, to which participants had to react. In considering these scenarios, the focus is on the response behavior of all three groups to the FDCRs and on the continued use of ACC after FDCRs within the multi-step groups.

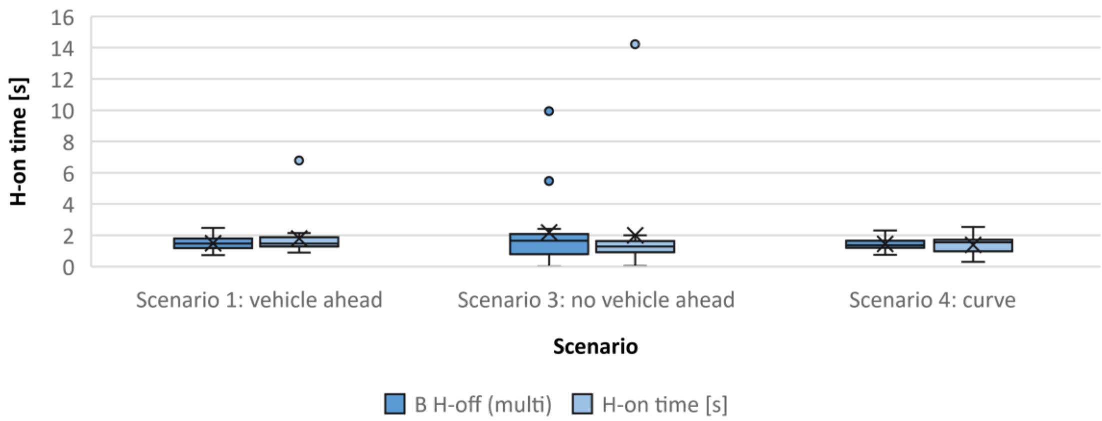


Figure 5-10: Boxplots showing the hands-on time after FDCRs for those participants who did not have their hands on the steering wheel at FDCR.

Function direct control requests (FDCR)

Hands-on time

When analyzing hands-on time, only participants who did not yet have their hands on the steering wheel at the time of the FDCR were considered (scenario 1: $n_{multi\text{-}step} = 16$; $n_{clear\text{-}cut} = 18$; scenario 2: $n_{multi\text{-}step} = 15$; $n_{clear\text{-}cut} = 16$; scenario 4: $n_{multi\text{-}step} = 15$; $n_{clear\text{-}cut} = 18$). T-tests revealed that there are no significant differences between the L2H-off multi-step and the L2H-off clear-cut groups regarding their hands-on reaction times after a FDRC was given (Scenario 1: $t(32) = -.94$, $p = .356$; Scenario 3: $t(29) = .04$ $p = .969$; Scenario 4: $t(31) = .17$, $p = .869$; see Figure 5-10). Both groups put their hands equally fast back onto the steering wheel after the FDCR was presented.

Time to first input

A one way (Welch) ANOVA was performed for each scenario. The analysis of intervention times revealed that there are no significant differences between the groups for Scenario 1 ($F(2, 53) = 1.017$, $p = .369$), Scenario 3 ($F(2, 49) = .103$, $p = .903$) and Scenario 4 ($F(2, 41) = .34$, $p = .717$) (see Figure 5-11). All three groups reacted equally fast by giving an active input.

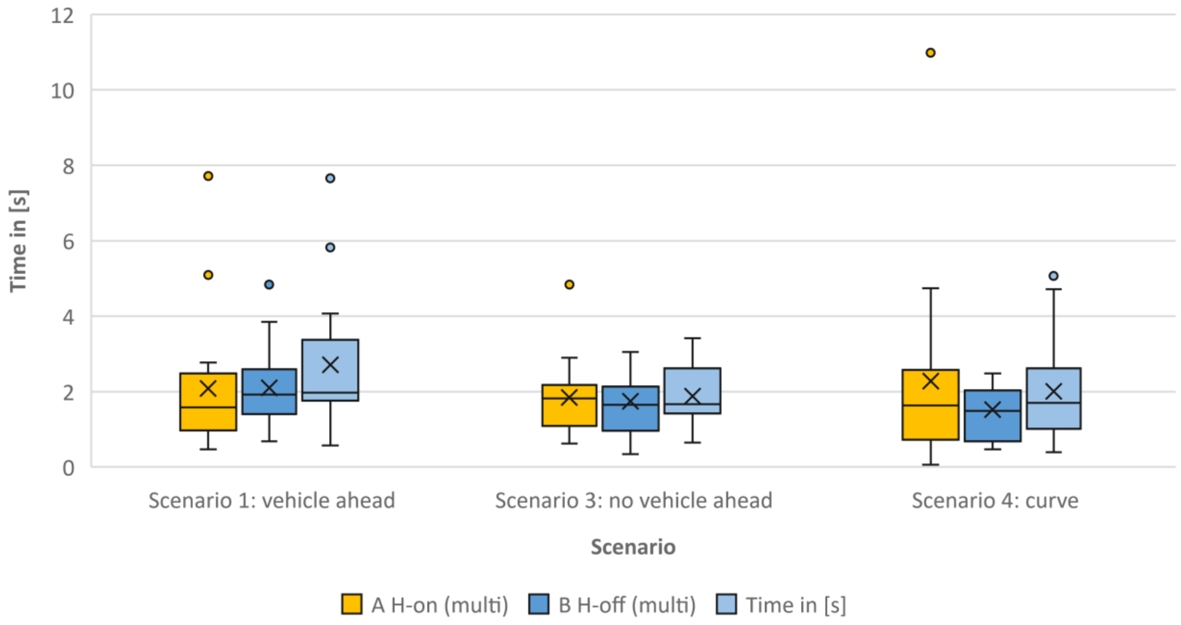


Figure 5-11: Boxplots showing the time to first input after FDCRs.

Table 5-9 gives an overview of the mean times to first intervention. Active input occurred 1.53 s to 3.07 s after the FDCR.

Table 5-9: Mean values and standard deviations of time to first intervention.

Group	Scenario 1		Scenario 3		Scenario 4	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on (multi-step)	2.09	1.78	1.84	1.07	3.07	4.12
L2H-off (multi-step)	2.10	1.06	1.74	1.02	1.53	1.03
L2H-off (clear-cut)	2.71	1.69	1.88	0.75	2.00	1.36

This could partly explain the significant difference. Table 5-10 gives an overview of the first interventions following after the FDCR per group. It shows that in all three scenarios and in all three groups, the first response to the FDCR was almost always a steering intervention.

Table 5-10: Overview of the first interventions following after the FDCR per experimental group.

	Scenario 1			Scenario 3			Scenario 4		
	A H-off (multi)	B H-off (multi)	C H-off (clear)	A H-off (multi)	B H-off (multi)	C H-off (clear)	A H-off (multi)	B H-off (multi)	C H-off (clear)
Steering	13	17	15	10	17	16	8	13	8
Braking	4	2	1	3	2	2	9	6	9
Accelerating	1	0	3	1	0	1	0	0	2

ACC usage

To assess whether people experiencing a multi-step L2 function make (appropriate) use of the ACC (L1 mode), ACC usage after FDCR was analyzed. In Scenario 1, it would be appropriate to leave ACC active as the vehicle adapts to the changed speed limit when following the vehicle ahead. In Scenario 3 and 4, participants should detect that leaving the ACC active (due to no control over set-speed) would lead to an inappropriate speed and they should cancel ACC via braking or the button. Table 5-11 shows how many participants in each group had the ACC activated at the beginning (FDCR) and at the end of the scenario (FDCR + 10 s). The overview shows that more participants deactivated ACC in Scenario 3 and 4 compared to Scenario 1 and that overall more participants of the L2H-off multi-step group deactivated ACC. However, some participants in both groups left the ACC activated in Scenario 3 and 4, resulting in an inappropriate speed. This could have several reasons.

Leaving ACC active could be interpreted as an indication for mode confusion. However, keeping the results of the video-based interviews in mind, it appears that some participants rather experienced an expectation mismatch than mode confusion: As they expected the ACC to react automatically, they may have taken longer to detect the inappropriate speed. Overspeeding was therefore unintentional. Secondly, as some participants indicated that they couldn't remember the situation, they may not have noticed the new speed limit and weren't aware of their own inappropriate speed. The changed speed limit was not displayed in the HMI, so the information regarding the changed speed limit was not available if participants missed the road sign. Due to the missing lateral dynamics of the static simulator, the inappropriate speed was probably further difficult to detect, especially in Scenario 4, where in reality over-speeding would have been physically noticeable by higher lateral acceleration. An additional third explanation could be that participants may have noticed the new speed limit, but actively decided to leave ACC activated, which would be an indicator for intentional inappropriate usage.

Table 5-11: Participants who had the ACC activated at the beginning (FDCR; L2 deactivation) and at the end of the scenario (FDCR + 10s).

Group	Scenario 1		Scenario 3		Scenario 4	
	Begin	*End*	*Begin*	*End*	*Begin*	*End*
L2H-on (multi-step)	78%	72%	72%	28%	72%	17%
L2H-off (multi-step)	84%	58%	68%	16%	63%	5%

5.3.4.3 Scenario 2 and 5

In Scenarios 2 and 5, no FDCR was given. Participants had to notice themselves that they needed to intervene. In Scenario 2, steering input was needed, while in Scenario 5, braking input was necessary to avoid a collision.

Hands-on time

T-tests revealed that there are no significant differences between the L2H-off multi-step and the L2H-off clear-cut groups regarding their hands-on times when approaching an obstacle (Scenario 2: $t(27) = -.30$, $p = .766$; Scenario 5: $t(31) = -1.44$, $p = .159$; see Figure 5-12). Both groups put their hands equally fast back onto the steering wheel (Scenario 2: L2H-on multi-step: $M = -8.30$, $SD = 9.52$; L2H-off multi-step: $M = -7.33$, $SD = 7.76$; Scenario 5: L2H-on multi-step: $M = -17.47$, $SD = 5.31$; L2H-off multi-step: $M = -14.78$, $SD = 5.37$).

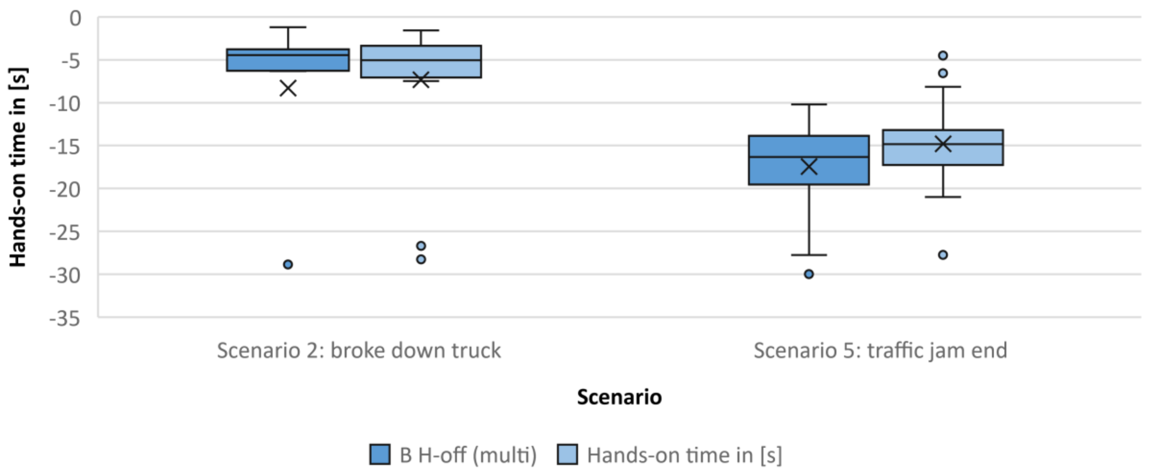


Figure 5-12: Boxplots showing the hands-on time. The origin of the y-axis (0 s) represents the point in time when the obstacle would be hit.

In both scenarios, outliers can be observed. One interpretation could be that participants simply put their hands on the steering wheel without any concrete reason. Another interpretation could be, that both scenarios could be seen some time in advance and could therefore be anticipated which could then in turn lead to early hands-on times.

Time to first input

An ANOVA was performed for each scenario to assess whether the three groups differed regarding their time to first input. For both Scenario 2 ($F(2, 46) = .75$, $p = .476$) and Scenario 5 ($F(2, 49) = .11$, $p = .901$), no significant differences could be found. All three groups reacted equally fast to the respective obstacle (Scenario 2: L2H-on multi-step: $M = -3.08$, $SD = 1.39$; L2H-off multi-step: $M = -3.55$, $SD = 3.92$; L2H-off clear-cut: $M = -4.49$, $SD = 4.56$; Scenario 5: L2H-on multi-step: $M = -16.11$, $SD = 4.99$; L2H-off multi-step: $M = -16.19$, $SD = 4.66$; L2H-off clear-cut: $M = -15.60$, $SD = 2.18$) (see Figure 5-13).

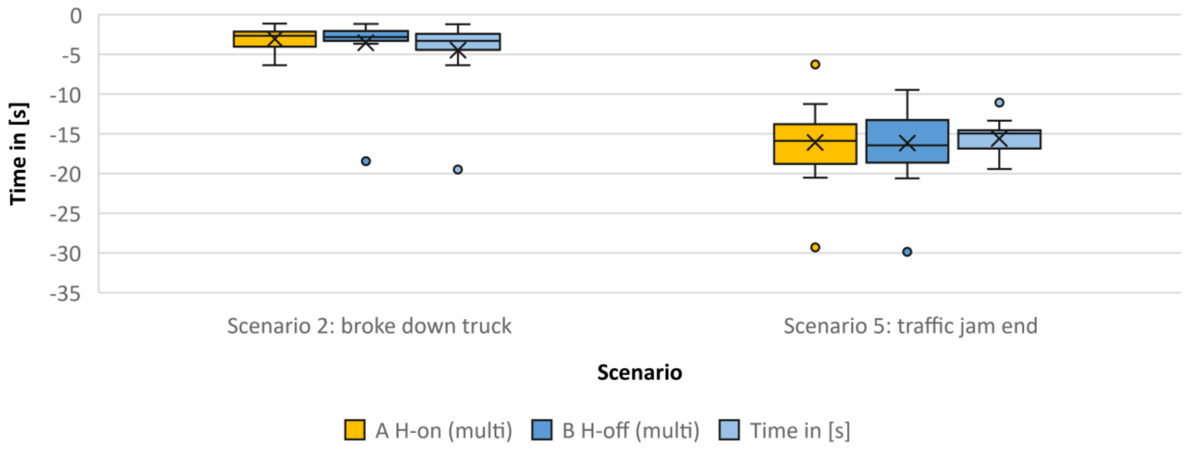


Figure 5-13: Boxplots showing the time of first input. The origin of the y-axis (0 s) represents the point in time when the obstacle would be hit.

Minimum lateral distance

For Scenario 2, minimum lateral distance to the truck was analyzed to rate the evasive maneuver conducted in response to the emerging collision object in the ego lane (i.e., the truck). The minimum lateral distance was calculated based on the distance between the center of the ego vehicle and the center of the truck. A one-way ANOVA revealed no significant differences (F(2, 49) = .08, p = .923) between the three groups (L2H-on multi-step: M = 4.51, SD = 0.90; L2H-off multi-step: M = 4.64, SD = 1.01; L2H-off clear-cut: M = 4.56, SD = 0.78), indicating that all three groups seem to have avoided the obstacles with a similar distance (see Figure 5-14).

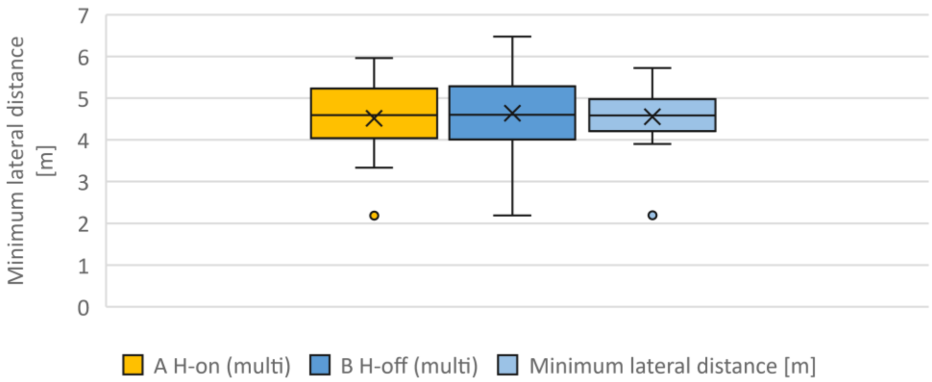


Figure 5-14: Boxplots showing the minimum lateral distance to the truck.

Minimum frontal distance

For Scenario 5, minimum distance (front ego vehicle to rear front vehicle) to the nearest vehicle of the traffic jam ahead was analyzed. A one-way ANOVA revealed no significant differences (F(2, 52) = .52, p = .599) between the three groups (L2H-on multi-step: M = 21.21, SD = 8.43; L2H-off multi-step: M = 19.76, SD = 6.72; L2H-off clear-cut: M = 22.47, SD = 9.31), indicating that all three groups seem to have avoided the obstacles with a similar distance (see Figure 5-15).

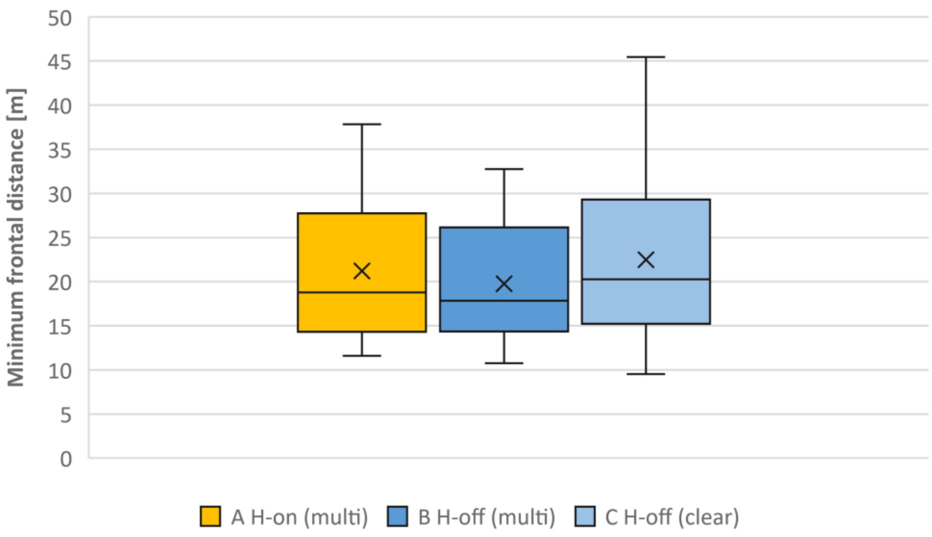


Figure 5-15: Boxplots showing the minimum frontal distance to the traffic jam end.

5.3.5 Summary / Conclusion

Study 2 focuses on the research subject “Mode Confusion” (CQ4). As described beforehand, there is only little to no research on the occurrence of mode confusion when switching between different L2 function designs (H-on and H-off). One possible concern was that individuals may have a reduced awareness of their responsibilities when using an L2H-off function, as they have less contact with the steering wheel and thus might feel less directly involved in the driving task. Therefore, the current study aimed to assess whether hands-free L2 functions increase mode confusion and if so, which countermeasures should be implemented on the function design side to counteract this. The following main research questions were addressed:

- RQ1: Are there differences between L2H-on and L2H-off functions with regard to mode confusion?
- RQ2: Do clear-cut transitions (L0 - L2 and vice versa) increase mode awareness/decrease mode confusion compared to successive/multi-step transitions (L0 – L1 – L2 and vice versa)?
- RQ3: Which transitions (L0 - L1; L0 - L2; L1 – L2 and vice versa) cause mode confusion?

Regarding RQ1, we assessed whether people using L2H-off functions are more prone to mode confusion based on the concerns mentioned beforehand. Regarding RQ2, it was assumed that clear-cut modes of either on or off should increase mode awareness/decrease mode confusion as there are less transitions the driver may go through and less task variability. The third research question was assessed rather exploratory. However, insights from the Expert Study (Section 4.2) suggested that there might be more confusion when switching between L1 and L2 modes.

The study results will first be interpreted with regards to the main RQs and with focus on CQ4. Subsequently, the results are discussed and interpreted in light of the other four projects CQs.

Main research questions of the study

In sum, the results of the study provide no evidence that mode confusion occurred when using a L2H-off function compared to when using a L2H-on function. For both L2 functions assessed within the current study a rather good to very good understanding of system functionality, system limits and driver responsibilities was observed, which is an essential prerequisite for mode awareness. Video-based interview data indicates that most participants did not get confused at all, as the majority responded that vehicle behavior matched with their expectations and that they knew which tasks they had to perform as drivers and which tasks the vehicle would continue to perform. Nevertheless, few participants indicated that FDCRs were confusing when there were no visible or obvious reasons for them to occur. Furthermore, interview results suggest that some participants experienced rather automation expectation mismatches than mode confusion, when it comes to adapting speed automatically or to avoiding obstacles. However, these confusions occurred for a small proportion of participants and over all three groups and are therefore not specific to a certain L2 design. Furthermore, looking at SUS ratings and evaluation of perceived complexity no differences could be found. In fact, SUS ratings were very good for both L2H-off functions and the L2H-on function investigated and perceived complexity was medium to low, again indicating that all three function designs were rather not subjectively confusing. There were also no differences in the eyetracking data to suggest that one group exhibited mode inappropriate gaze behavior.

When looking at the objective simulator data, there were very few L2 activation attempts when the function was currently not available, indicating that participants of the multi-step groups knew in particular when they could activate L2. However, it should be mentioned, that the L2 function was available frequently, so there was not an overly large number of opportunities for erroneous activation attempts overall. The hands-free driving times are also plausible and give no reason to conclude mode confusion for the multi-step groups. Furthermore, all participants were able to handle Scenarios 2 and 5 safely. When it comes to reactions to FDCRs, no significant differences could be found between the three groups.

RQ2 targeted potential differences between clear-cut and multi-step groups. In fact, there were mostly no differences between these groups. When looking at the eyetracking data, no differences could be found that would suggest a deviating or inappropriate gaze behavior in any of the groups. Both function designs were perceived equally low to medium complex and usable. There are also no statistical differences in terms of system and driver role understanding. However, on a single-item basis, it appears that the majority of participants experiencing the multi-step function was uncertain whether they would fall back onto L1 after an FDCR. This finding is interesting as participants did experience this transition at least four times throughout the experiment. This finding could be an indication that mode confusion may occur in these situations and especially when it comes to those transitions. This finding should be kept in mind when designing L2 functions and HMI.

Neither hands-on times nor time to first intervention after an FDCR differed between the multi-step and the clear-cut groups, indicating that both groups prepared to take over lateral guidance equally fast.

In sum, there is no evidence that L2H-off systems lead to (more) mode confusion. Furthermore, multi-step systems seem to be rated and experienced equally little complex, usable and understandable compared to clear-cut systems. However, the results regarding inappropriate ACC usage and uncertainties regarding the fallback onto L1 might indicate that (unintentional) errors and mode confusion could occur more frequently in these systems due to their multi-layered design.

CQ1 Hands-off = mind-off?

Based on the current study results there are no indications that L2H-off functions lead to users being less engaged (mind-off) compared to L2H-on functions. There were no differences with respect to eyetracking data, indicating that visual attention was equal across groups. Furthermore, coping behavior (e.g., time to input or minimum distances) in Scenarios 2 and 5, where participants had to detect obstacles and resulting intervention needs themselves, gave no indications for L2H-off users to be less attentive than L2H-on users.

CQ2 Prolonged transition times

Based on the current study results, there are no indications for significantly prolonged transition times for L2H-off functions compared to L2H-on functions. Regarding mean reaction times to HOR and EOR, it actually shows that the L2H-off multi-step group reacted significantly faster than the L2H-on multi-step group. Regarding reaction time to FDCRs, no differences could be found. Overall, the study found no evidence that the L2H-off groups experienced a disadvantage with respect to transition times.

CQ3 Foreseeable misuse

No NDRT was offered or admissible during L2 use in this study, limiting the explanatory power towards CQ3. Based on the questionnaire results regarding trust in automation, however, the data of the current study provide no evidence that the use of L2H-off functions leads to greater or more frequent misuse compared to L2H-on functions. The data suggest that trust is given, but not high, probably indicating that misuse may be rather unlikely when using L2H-on and L2H-off functions.

CQ5 Safety

Regarding CQ5, the data on coping with Scenario 2 and 5 are probably the most interesting of this study. In both scenarios, no collisions occurred. Furthermore, results regarding time to first input, lateral or frontal minimum distance indicate that both scenarios were handled well by participants experiencing a L2H-on function and by participants experiencing a L2H-off function.

5.3.6 References

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R. & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29* (4), 552-557 doi:10.7334/psicothema2016.383

Boos, A., Feldhütter, A., Schwiebacher, J., & Bengler, K. (2020). Mode errors and intentional violations in visual monitoring of level 2 driving automation. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1-7). doi:10.1109/ITSC45102.2020.9294690

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189(194*), 4-7.

Consumer Reports (2020). *Active Driving Assistance Systems: Test Results and Design Recommendations*. Retrieved from: https://data.consumerreports.org/wp-content/uploads/2020/11/consumer-reports-active-driving-assistance-systems-november-16-2020.pdf

Körber, M.(2019). Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (Vol. 823, pp. 13–30). Cham: Springer International Publishing. doi:10.1007/978-3-319-96074-6_2

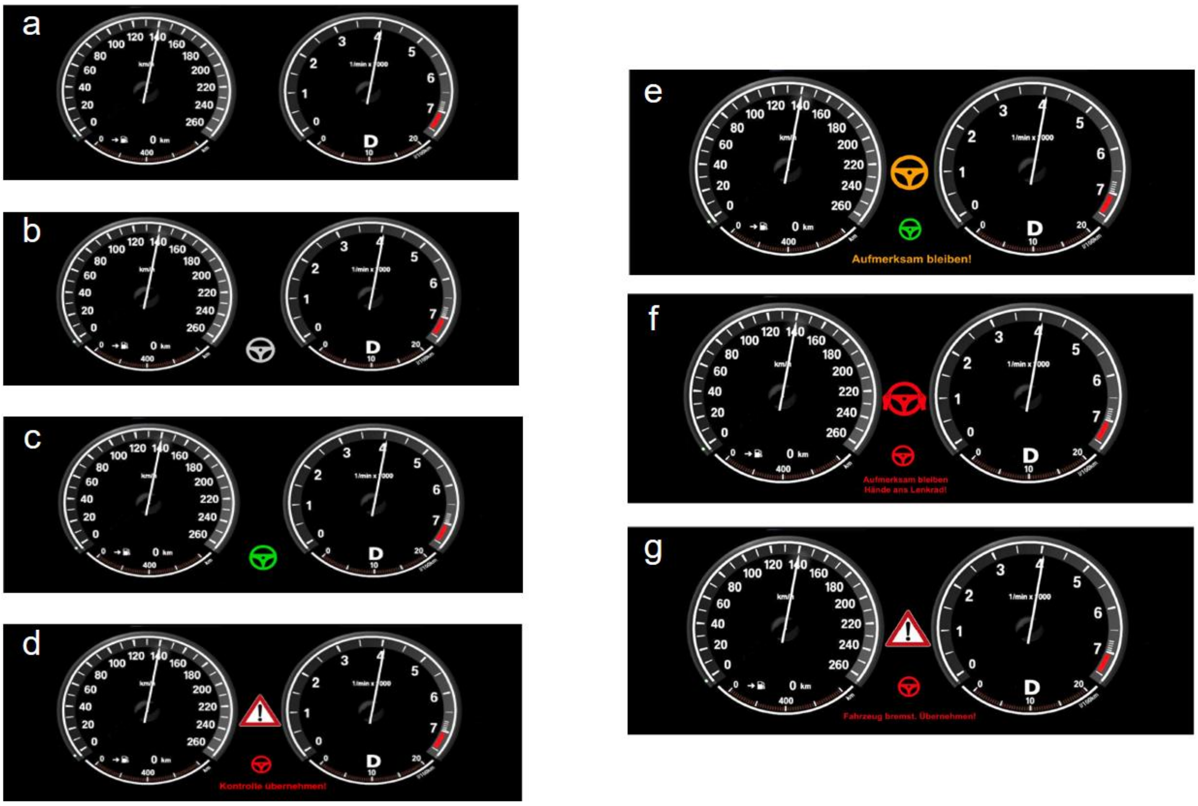
Kurpiers, C., Biebl, B., Mejia Hernandez, J., & Raisch, F. (2020). Mode awareness and automated driving—What is it and how can it be measured?. *Information, 11*(5), 277. doi:10.3390/info11050277

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces andInteractive Vehicular Applications (*pp. 51-58).

5.3.7 Appendix

5.3.7.1 Appendix I

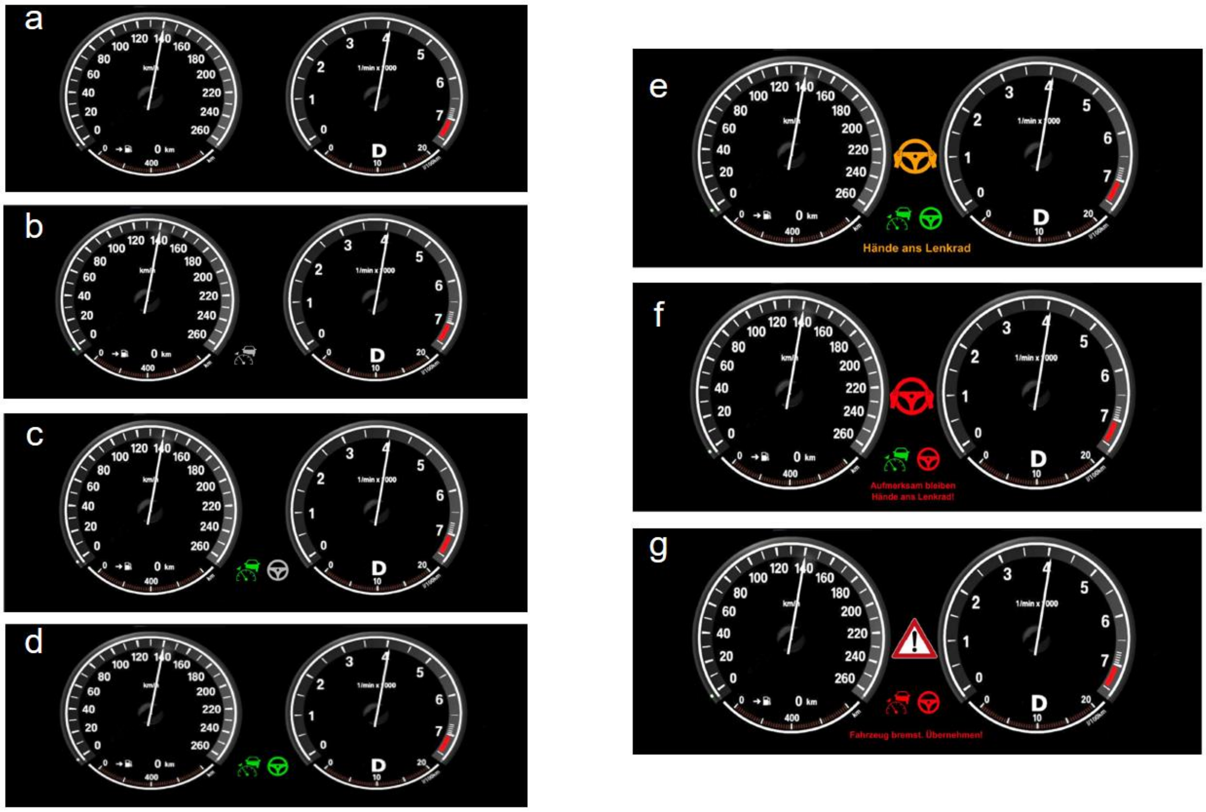
The following pictures summarize the HMI displays of the L2H-off clear-cut function.



(a) L2 function not available; (b) L2 function available; (c) L2 activated; (d) FDCT; (e) DMS EOR; (f) DMS HOR; (g) DDCT

5.3.7.2 Appendix II

The following pictures summarize the HMI displays of the L2H-on multi-step function.



(a) L1 function not available; (b) L1 function available; (c) L1 activated and L2 function available; (d) L2 function activated; (e) DMS HOR; (f) DMS HOR; (g) DDCT

5.3.7.3 Appendix III

Items to assess system understanding:

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte Fahrerassistenzsystem.

	Nicht zutreffend	Zutreffend	Unsicher
Das Level 2 System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer.	○	○	○
Das Level 2 System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden.	○	○	○
Wenn das Level 2 System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit.	○	○	○
Das Level 2 System passt die Geschwindigkeit an langsamer vorausfahrende Fahrzeuge an, außer diese stehen oder die Differenzgeschwindigkeit ist zu hoch.	○	○	○
Das Level 2 System erkennt immer, wenn es eine Situation nicht meistern kann.	○	○	○
Das Level 2 System lenkt selbstständig innerhalb des eigenen Fahrstreifens.	○	○	○
Wenn das Level 2 System aktiv ist, hält es einen voreingestellten Mindestabstand zum vorausfahrenden Fahrzeug ein.	○	○	○
Wenn das Level 2 System aktiv ist, steuert es die Geschwindigkeit des Fahrzeugs entsprechend der eingestellten Geschwindigkeit.	○	○	○
Wenn das Level 2 System aktiv ist, führt es Fahrstreifenwechsel durch, ohne dass der Fahrende dabei selbst lenken muss.	○	○	○
Wenn ich, während das Level 2 System aktiviert ist, bremse, fallen alle Assistenzfunktionen weg und ich muss selbstständig lenken sowie Geschwindigkeit und Abstand regeln, bis ich diese wieder aktiviere.	○	○	○
Wenn ich, während das Level 2 System aktiviert ist, lenke (um bspw. einem Hindernis auszuweichen), fährt das Fahrzeug mit der eingestellten Höchstgeschwindigkeit weiter.	○	○	○
Wenn ich, während das Level 2 System aktiviert ist, beschleunige, reduziert das Fahrzeug die Geschwindigkeit anschließend wieder auf die eingestellte Höchstgeschwindigkeit, wenn ich den Fuß wieder vom Gaspedal nehme.	○	○	○
Um das Level 2 System zu aktivieren, muss ich den Knopf einmal drücken.	○	○	○
Wenn, während das Level 2 System aktiviert ist, das Signal „Bitte übernehmen" im Display erscheint, fallen alle Assistenzfunktionen weg und ich muss selbstständig lenken sowie Geschwindigkeit und Abstand regeln.	○	○	○
Es gibt einen Unterstützungsmodus, bei dem das Fahrzeug die Beschleunigung regelt aber der Fahrende den selbstständig lenkt.	○	○	○
Es gibt einen Unterstützungsmodus, bei dem das Fahrzeug die Lenkung übernimmt aber der Fahrenden selbstständig die Beschleunigung regelt.	○	○	○

Items to assess driver role understanding:

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte Fahrerassistenzsystem.

	Nicht zutreffend	Zutreffend	Unsicher
Ich muss das Level 2 System stets überwachen, wenn das System aktiviert ist und mich vergewissern, dass das System die Fahraufgabe entsprechend den aktuellen Bedingungen (z. B. Geschwindigkeitsbegrenzungen, Abstand zu anderen Verkehrsteilnehmern, Straßenbedingungen) durchführt.	○	○	○
Ich darf mich mit anderen Tätigkeiten, wie z.B. E-Mails schreiben, beschäftigen, wenn das Level 2 System aktiviert ist.	○	○	○
Ich muss zu jeder Zeit bereit sein die Fahraufgabe übernehmen zu können.	○	○	○
Ich kann jederzeit in die Automation eingreifen.	○	○	○
Ich darf mich von der Überwachung des Verkehrsraums abwenden, wenn das L2 System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind.	○	○	○
Ich muss dem Straßenverkehr die gleiche Aufmerksamkeit widmen wie beim Fahren ohne Fahrerassistenzsysteme (z. B. Straßenzustand, Interaktion mit anderen Fahrern).	○	○	○
Ich muss die gleiche oder eine ähnliche Sitzposition wie beim Fahren ohne Fahrerassistenzsysteme einnehmen.	○	○	○

5.4 Experimental Study 3

Documentation by P. Dautzenberg (Institut für Kraftfahrzeuge, RWTH Aachen University)

Similar to Study 1, Study 3 focuses on the functional design aspect “Attentiveness Alert (AR)” and on challenge 1 (hands-off = mind off?). While Study 1 investigates whether there are general differences between L2H-off functions with eyes-on request (EOR) and L2H-on functions with hands-on request (HOR) with regards to user behavior and safety, Study 3 focuses on the timing of AR and potentially resulting differences with regards to user behavior and safety.

5.4.1 Research Questions & Hypotheses

Study 3 aims to answer two main research questions (RQ):

- RQ1: Are there differences between L2H-off functions with eyes-on requests (EOR) and L2H-on functions with hands-on request (HOR) with regards to attention and user behavior?
- RQ2: Are there differences between L2H-off functions with differently timed eyes-on requests (EOR) with regards to attention and user behavior?

Regarding RQ1, it was assumed that users of an L2H-off function with EOR should show at least as good/safe driving performance and handling of ODD limits as users of an L2H-on function with HOR. Regarding RQ2, it was assumed that a DMS that cautions the driver to pay visual attention to the road after 3 seconds eyes-off road may enable the driver to notice obstacles or silent system failures earlier, especially during difficult-to-anticipate, time-critical ODD limits, than a DMS that cautions the driver for the first time after 5 seconds of inattention (see e.g. Euro NCAP, 2022; Victor et al., 2018; Simons-Morton et al., 2014; Schneider et al., 2022). The time intervals of 3 s and 5 s were selected based on literature and existing field solutions.

5.4.2 Method

5.4.2.1 Participants

In total N = 61 persons (72% male) with an average age of *M* = 33 years (*SD* = 14, Range: 22–66 years) participated in the study. Participants were randomly assigned to one of the three experimental groups. A description of the three groups can be found in Table 5-1.

Table 5-1: Description of the three experimental groups.

Group	Function	Sample size	Gender	Age	ACC experience	L2 experience
A	L2H-on	21	2 female 19 male	*M* = 36 *SD* = 16	43%	19%
B	L2H-off 5s	20	10 female 10 male	*M* = 32 *SD* = 11	60%	10%
C	L2H-off 3s	20	5 female 15 male	*M* = 32 *SD* = 13	45%	20%

To ensure that the survey results were not biased by a lack of affinity for technology, the Affinity for Technology Interaction Scale (ATI Scale; Franke, Attig, & Wessel, 2019) was used. The results showed that there were no significant differences between the experimental groups and an overall high tendency with regard to technology affinity.

5.4.2.2 Simulator and Mock-Up

Study 3 was conducted by ika at RWTH Aachen University in the same static driving simulator and using the mock-up of fka GmbH as described in Study 2.

5.4.2.3 Level 2 Function Design

For the current study, one L2H-on and two L2H-off functions were implemented. Each of the three functions was experienced and evaluated by one of the aforementioned groups. The L2H-on as well as the L2H-off functions tested in this study were implemented according to the clear-cut principle of Study 2 (see Section 5.3.2.3; transitions between L0 and L2 only; no L1 functions implemented). As in Study 2, brake interventions lead to an immediate deactivation of the L2 function and a fallback to L0. Steering input or acceleration can override the L2 function for a short time. A modification compared to Study 2 is that the speed could also be regulated via a hard key on the steering wheel with which participants could change set speed in 10 km/h steps. This enables a change of set speed without deactivating the L2 function. Again, for all functions, thresholds for the deactivation of L2 automation were ≥ 5.7° steering input (difference to current steering wheel angle) and 10% brake pedal travel.

5.4.2.4 DMS and HMI

The DMS was implemented in the same way as in Study 2. The HMI displays were also adopted from Study 2. The only change was a more prominent color display when L2 is activated (green frame) and when warnings (DMS or FDCR) are displayed (red filling; see Figure 5-1).

Figure 5-1: Improved HMI concept with more prominent color display when L2 is activated compared to the HMI concept of Study 2 (green frame; left) and when warnings are displayed (red filling; right)

For the L2H-off functions, the timing of the first warning stage of the AR was varied. Based on literature and existing field solutions, 3 s (see e.g. Euro NCAP, 2022; Victor et al., 2018; Simons-Morton et al., 2014; Schneider et al., 2022) and 5 seconds (e.g., see US vehicles tested in SP3.4) were assumed and implemented as variations in the study. For the Level 2 H-on system, the first warning stage of the HOR was set to 15 seconds (see Study 1 and 2). Figure 5-2 summarizes all three warning cascades.

L2H-off	Request Design					
HMI	1. Warning Stage		2. Warning Stage		3. Warning stage	
	3 s	Eyes-on Request (visual + acoustic alert) „Aufmerksam bleiben“	+3 s	Hands-on Request (visual + acoustic alert) „Aufmerksam bleiben. Hände ans Lenkrad!“	+5 s	DMS direct controrl Request (visual + acoustic alert) „Fahrzeug bremst. Übernehmen!“
Driver's task		Eyes on road		Eyes on road + hands on wheel		Take-over driving task

Adapted warning cascade to be assessed within Study 3.

L2H-off	Request Design					
HMI	1. Warning Stage		2. Warning Stage		3. Warning stage	
	5 s	Eyes-on Request (visual + acoustic alert) „Aufmerksam bleiben“	+3 s	Hands-on Request (visual + acoustic alert) „Aufmerksam bleiben. Hände ans Lenkrad!“	+5 s	DMS direct controrl Request (visual + acoustic alert) „Fahrzeug bremst. Übernehmen!“
Driver's task		Eyes on road		Eyes on road + hands on wheel		Take-over driving task

L2H-on	Request Design					
HMI	1. Warning Stage		2. Warning Stage		3. Warning stage	
	15 s	Hands-on Request (visual + acoustic alert) „Hände ans Lenkrad.“	20 s	Hands-on Request (visual + acoustic alert) „Hände ans Lenkrad.“	25 s	DMS direct controrl Request (visual + acoustic alert) „Fahrzeug bremst. Übernehmen!“
Driver's task		Hands on wheel		Hands on wheel		Take-over driving task

Warning cascade which were assessed within Study 3 which were also implemented in Study 1 and Study 2.

Figure 5-2: DMS warning cascades implemented and assessed in the current study.

5.4.2.5 Driving Scenarios

The driving duration was approx. 45 minutes. The highway consisted of two lanes in each direction with a Level of Service B. Since Study 3 follows on from Study 1, the same four (system limit) scenarios were implemented (cf. Section 5.2.2.4). Differences will be presented in the following. Scenario 1 and 3 are almost identical to Study 1. The ego vehicle does not automatically adapt to a new indicated speed limit. As there is no system reaction or function direct control request (FDCR), the driver must detect the necessity to intervene him/herself. One difference compared to Study 1 was that the new speed limit (which was to be detected by the driver) was not 100km/h but 80km/h to increase the affordance to intervene. Another difference was that shortly before and during scenario 1 the NDRT (see 5.4.2.6) was presented, while no NDRT was offered in scenario 3. Scenario 2 and 4 were also almost the same as with Study 1. However, the anticipation time available to the drivers to detect potential obstacles/taking control necessities was varied to assess the influence of DMS timing in situations of different time-criticality. In Scenario 2, the ego vehicle's lane is blocked by roadworks. The necessity to intervene is indicated by speed reduction and road signs indicating the roadworks ahead and the ending of the lane. The first sign was given 600m before the construction site. Therefore, the necessity to intervene can be anticipated. Moreover, the vehicle ahead changes lane (lane change duration ~4 s) so that the obstacle becomes fully visible to the driver 6 seconds before collision. If participants do not intervene, a FDCR is given at a time-to-collision (TTC) of 2.7 seconds. This timing corresponds to that of existing forward collision warnings (Nusholtz et al., 2013). In Scenario 4, the ego vehicle's lane is blocked by a breakdown vehicle rolling out on the right lane at approx. 30 km/h. The vehicle ahead changes lane (lane change duration ~4 s) so that the obstacle becomes fully visible to the driver 4 seconds before collision. As with scenario 2, a FDCR is given at a TTC of 2.7 seconds, if participants do not intervene beforehand. In contrast to scenario 2, there are no indications of the need for driver intervention other than the lane change of the vehicle ahead. Therefore, the necessity to intervene cannot be anticipated. The order in which scenario 2 and 4 are presented was counterbalanced across participants to avoid order effects. Figure 5-3 shows the presentation of the four scenarios within the entire testing period, as well as the availability of the L2 function and the time intervals during which the non-driving related task (NDRT, see 5.4.2.6) was offered.

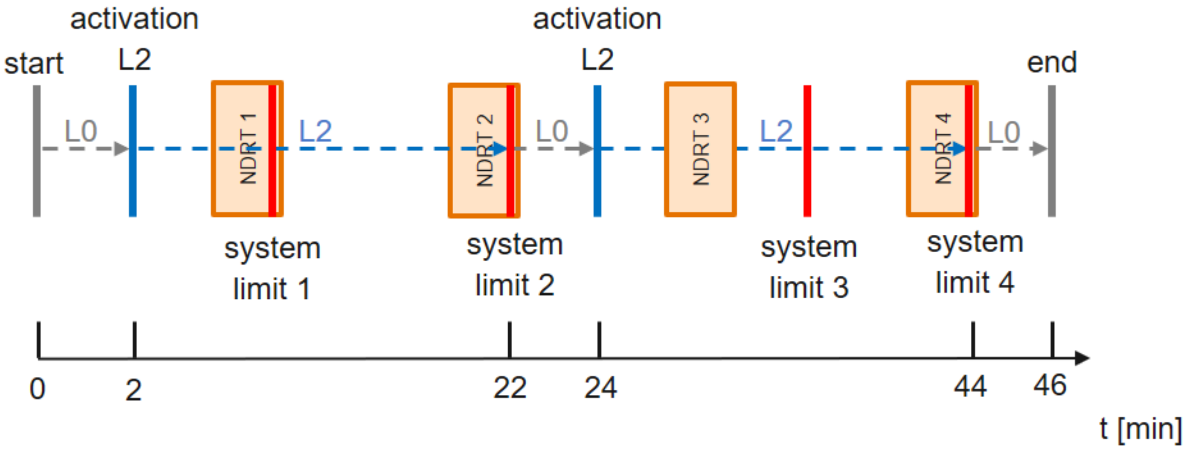


Figure 5-3: Experimental track incl. timing of scenarios, NDRT presentation and L2 availability.

5.4.2.6 Non Driving Related Task (NDRT)

To investigate if a certain AR timing (first warning stage after 3 s vs. 5 s) is better suited to enable the driver to react to obstacles (that are difficult vs. easy to anticipate due to the time available to detect the obstacle), visual distraction of the driver needs to be influenced/induced in a way that is reasonable and admissible for the driving context. Therefore, a (well interruptible) task (in line with requirements defined by ESoP, 2005) is needed that (partially) engages the visual attention of the driver. For this purpose, we used a NDRT that was also used in a similar variation in Study 1 (cf. Section 5.2.2.3.2). Participants were asked to read texts (~ 60 words) which were presented on a ten inch 16-by-9 display positioned on the dashboard directly above the center stack for 30 seconds. Subsequently, they were presented with questions about the texts they just read. Participants had 10 seconds to read and answer the question. The time available for reading (30 s) and answering (10 s) was limited to partially control potential levels of distraction. However, two major changes from Study 1 were that the responses to the visually presented questions were to be given verbally (not via keystroke) and that the task was not presented throughout the entire experimental period, but rather selectively at four time points for a period auf 05:30 minutes each (see Figure 5-3).

It was up to the participants how strongly they wanted to engage in the NDRT depending on whether they felt safe. They were offered additional 5€ to their announced 35€ incentive if they performed particularly well on the NDRT. However, if safety-relevant situations occurred or their driving behavior was classified as unsafe, they were informed not to receive the additional incentive. What the participants did not know was that they would receive the full 40€ incentive in any case. Participants were informed of this after the study.

5.4.2.7 Study Design

A between-subject design with the factor *Level 2 Function Design* (L2H-on with HOR vs. L2H-off with EOR after 5 s vs. L2H-off with EOR after 5 s) was applied. The variation of L2 functions resulted in three experimental groups. As dependent variables both subjective questionnaire ratings (see Table 5-2) as well as behavior data gathered via eyetracking and simulator were collected and assessed. Table 5-3 lists all subjective and objective data collected and analyzed within Study 3.

Table 5-2: Subjective questionnaire data.

Questionnaires and single items		CQ
Subjective rating of monitoring performance	• Single item: “How attentively did you monitor system behavior and surrounding traffic when the L2 function was activated?” (see also Study 1)	1
DMS evaluation	• 8 items (see also Study 1 and 2)	3, 4, 5
Mode confusion	• System and Driver role understanding (see also Study 2, Appendix) • Single item: “Was it always clear to you whether the L2 function was activated or not, or what support the L2 function provides?” (see also Study 2)	4
Trust in Automation	TiA (Körber, 2019) - subscales: • Understanding/Predictability • Trust in Automation	3, 5
Acceptance	CTAM (Osswald et al., 2012) – subscales: • Behavioral intention to use • Perceived safety	3, 5
NDRT involvement	NDRT questionnaire (adapted by Metz et al., 2014) (see also FOT and Study 1)	3

Table 5-3: Objective data.

Objective data		Unit	CQ
Eyetracking	• Number of eyes-off road glances > 2s • Attention ratio: eyes-on road, eyes-on instrument cluster/steering wheel, eyes-on NDRT, eyes on other	• Number • Percentage	1
Simulator	• Number of HOR and EOR • Reaction frequency to changed speed limit	• Number • Percentage	1
	• Hands-free driving while driving L2 • Number of attempted activations of L2, although not available	• Percentage • Number	4
	• Reaction time to HOR and EOR • Time to braking/deceleration input • Number and description of incidents • Minimum TTC • Minimum lateral distance	• Time in s • Time in s • Number • TTC in s • Distance in m	2, 5

5.4.2.8 Procedure

Participants were welcomed, asked to read and sign all legal documents, and informed about the general procedure. Participants were then asked to answer demographic and mobility behavior related questions. After that, participants were instructed to read the instruction manual, explaining the L2 function and the NDRT. Subsequently, participants answered the items regarding system and driver role understanding, after which participants were brought to the driving simulator and started a familiarization drive (about 10 to 15 minutes) to get used to the simulator and the L2 function. The experimenter motivated participants to try different ways to override or deactivate the L2 function. Within the familiarization drive, participants also experienced the NDRT and a FDCR. After the familiarization drive, the experiment drive (about 45 minutes) started. During the experimental drive, the four scenarios were presented as described beforehand. When the NDRT was presented, the experimenter noted the number of responses given by the participants, regardless of their correctness. After the experimental drive, participants were asked to answer the post-drive questionnaires (regarding system and driver role understanding, mode confusion, monitoring behavior, DMS evaluation, trust, and acceptance). Finally, participants were informed about the study goals, incentivized and seen off. The total duration of the experiment was about 1:30 h per participant.

5.4.2.9 Data Analysis

The data was analyzed using SPSS 27 (IBM Statistics). Normality distribution and equality of variances were analyzed using the Shapiro-Wilk test and the Levene’s test. If these assumptions for parametric tests were not met, we used the non-parametric alternative. We performed

ANOVA regardless of the violation of the normal distribution, as it appears to be robust to the violation of this condition (Blanca, Alarcón, Arnau, Bono & Bendayan, 2017). In case of violation of equality of variances, Welch ANOVA was performed. If not reported otherwise, a one-way ANOVA was conducted for all subjective and objective measures. The significance level was set to $\alpha = .05$.

5.4.3 Results

5.4.3.1 Subjective Data

Subjective rating of monitoring performance

Participants were asked to rate their own monitoring performance after the experimental drive via the single item: “How attentively did you monitor system behavior and surrounding traffic when the L2 function was activated?”. The corresponding ANOVA revealed no significant difference between the three groups ($F(2, 58) = 0.75$, p = .475). The mean values (L2H-on: $M = 3.95$, $SD = 1.25$; L2H-off 5s: $M = 3.60$, $SD = 1.50$; L2H-off 3s: $M = 4.10$, $SD = 1.09$) indicate that participants of all three groups tended to rate their own monitoring performance as average to very good (see Figure 5-4).

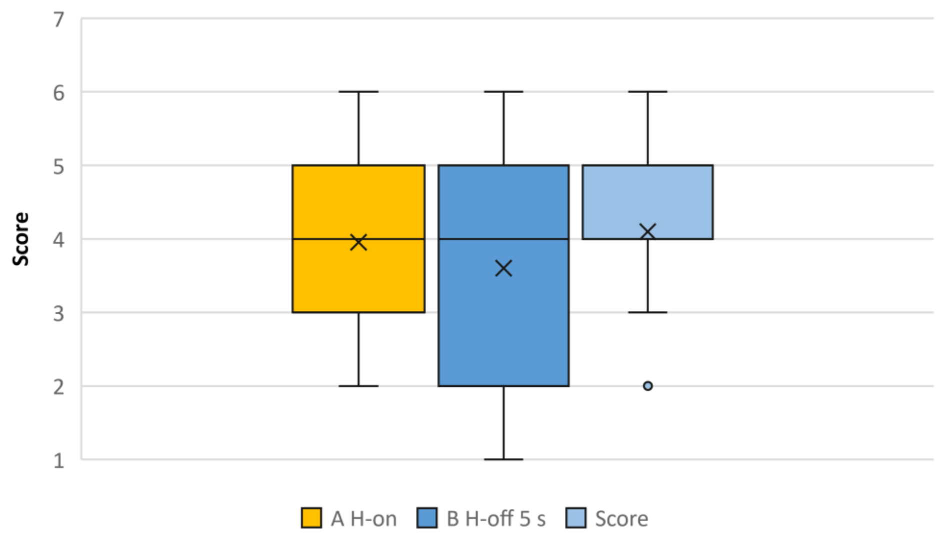


Figure 5-4: Boxplots for the self-rating of the monitoring performance (1 = inattentive to 7 = always attentive).

DMS Evaluation

Participants were asked to evaluate the DMS they experienced throughout the study. This was done using the eight single items already used in Study 1 and Study 2. For each item, a one-way (Welch) ANOVA was performed. The results show significant differences between the three groups for the items DMS01 (“The warnings came too frequently”; $F(2, 35.31) = 9.46$, p = .001), DMS04 (“I found the warnings annoying”; $F(2, 33.87) = 5.45$, $p = .009$), and DMS05

("I could understand why a warning occurred"; *F*(2, 58) = 3.32, p = .043). Post-hoc Games-Howell and Tukey-corrected t-tests were conducted, which showed that the L2H-on group differed significantly from the L2H-off 5 s group for all three items (DMS01: *p* = .045; DMS04: *p* = .030; DMS05: *p* = .034) and from the L2H-off 3 s group for the DMS01 item (*p* = .001). The results indicate that the L2H-on group's DMS was perceived as less intrusive and that the L2H-on group more often understood why a warning occurred, at least compared to the L2H-off 5s group (see Figure 5-5). This finding could be attributed to the fact that the L2H-on DMS reacts later (after 15 s) and thus inherently less frequently.

The two L2H-off groups did not differ significantly in their ratings. It appears that there are no significant differences between a DMS that gives a first warning after 5 s and a DMS that prompts a first warning after 3 s of gaze aversion when looking at participants' subjective evaluation regarding frequency, annoyance, understanding or fulfilment of purpose.

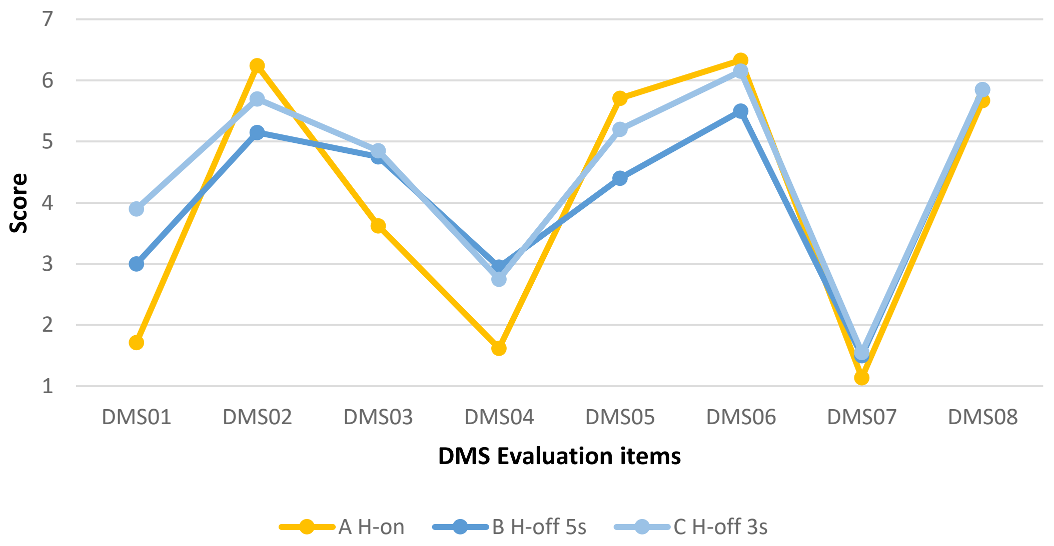


Figure 5-5: Mean values of DMS evaluation per item (1 = I don't agree at all to 7 = I totally agree).

Mode confusion

Participants were asked to answer questions that would assess their understanding of the L2 function, its limits and functionality as well as their understanding of their role and responsibilities as the driver. These questions were asked before and after the experimental drive. This was done via almost the same items already used in Study 2 (see Appendix). A two-way ANOVA was performed with the factors *Level 2 function design* (L2H-on vs. L2H-off 5s vs. L2H-off 3s) and *Assessment point* (pre vs. post). For system understanding, no main effects for the two factors (Function design: *F*(2,116) = .020, *p* = .981; Assessment point: *F*(2,116) = .007, *p* = .933) nor an interaction could be found (*F*(2,116) = .104, *p* = .901). Overall, the mean values indicate a good to very good system understanding (see Figure 5-6).

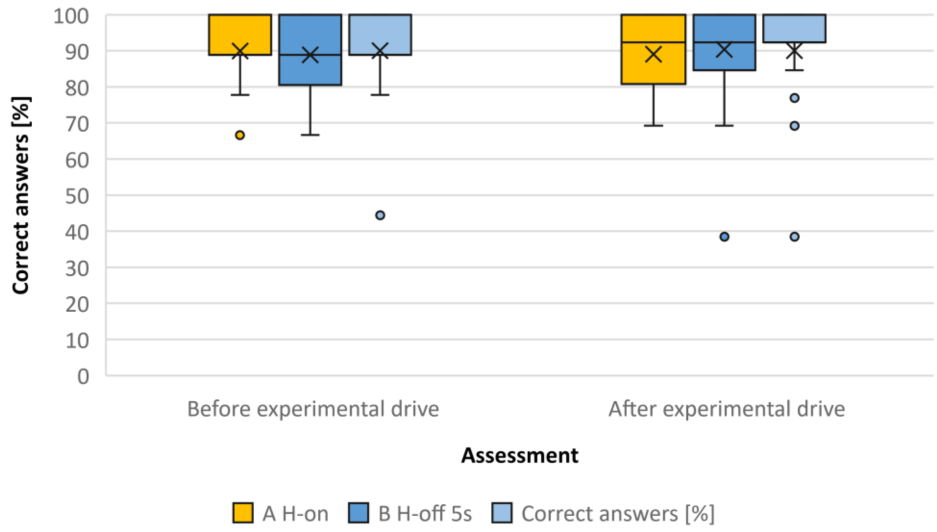


Figure 5-6: Boxplots for the proportion of correct answers targeting the system understanding before and after the experimental drive.

For driver role understanding again no main effects for the two factors (Function design: $F(2,86) = 1.26$, $p = .288$; Assessment point: $F(2,86) = .55$, $p = .458$) nor an interaction could be found ($F(2,116) = 1.11$, $p = .333$). Overall, the mean values indicate a good to very good driver role understanding (see Figure 5-7).

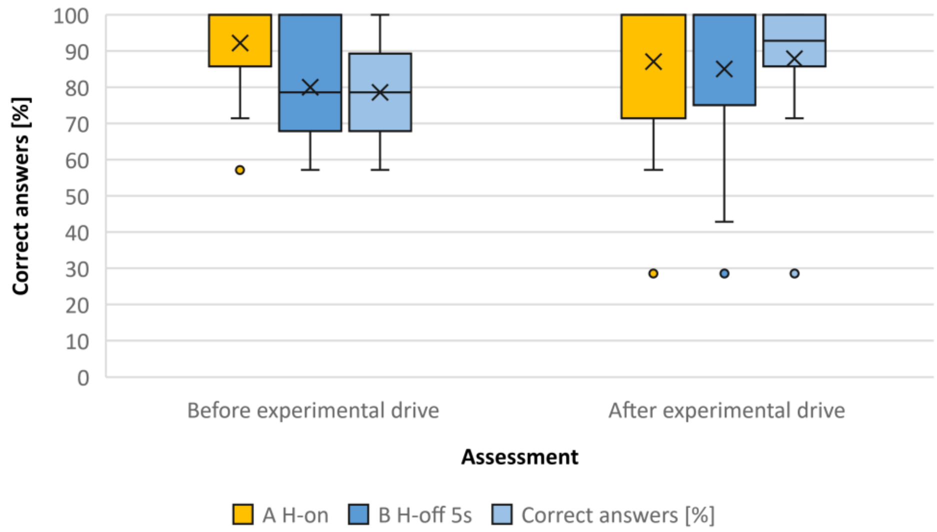


Figure 5-7: Boxplots for the proportion of correct answers to items targeting the driver role understanding before and after the experimental drive.

Furthermore, participants were asked to answer the single item: "Was it always clear to you whether the L2 function was activated or not, or what support the L2 function provides?" (see also Study 2). The majority of participants indicated that it was clear to them whether the L2 function was activated or not (L2H-on: 90%; L2H-off 5 s: 95%; L2H-off 3 s 100%). Additionally, participants were asked if they had to actively react to speed limits while driving to check for appropriate mode awareness in scenario 1 and 3 in particular. The majority of participants indicated correctly that they had to actively react to speed limits throughout the experimental drive (L2H-on: 100%; L2H-off 5 s: 100%; L2H-off 3 s: 95%).

Trust

To assess participants' trust in automation, the subscales "Understanding/Predictability" and "Trust in Automation" from Körber (2019) were assessed and analyzed. Furthermore, an overall score was calculated and considered. The ANOVA results show no significant differences between the three groups, neither for "Understanding/Predictability" ($F(2, 58) = 1.54$, $p = .223$), nor for "Trust in Automation" ($F(2, 58) = .056$, $p = .945$). Overall, mean scores indicate average to high trust across all groups for both subscales (see Table 5-4).

Table 5-4: Mean values and standard deviations for the TiA subscales (Körber, 2019). The scale ranges between 1 = I don't agree at all and 5 = I fully agree.

Group	Overall		*Understanding/ Predictability*		Trust in Automation	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on	4.04	0.64	4.46	0.46	3.62	0.97
L2H-off 5s	3.85	0.82	4.18	0.67	3.53	1.07
L2H-off 3s	3.85	0.81	4.18	0.69	3.52	1.09

Acceptance

To assess participants acceptance towards the respective L2 function, the subscales "Behavioral Intention to Use" and "Perceived Safety" from Osswald et al. (2012) were assessed and analysed. The (Welch) ANOVA results show no significant differences between the three groups, neither for "Behavioral Intention to Use" ($F(2, 33.82) = 1.19$, $p = .314$), nor for "Perceived Safety" ($F(2, 58) = .28$, $p = .755$). Overall, mean scores are average to high across all groups for both subscales (see Table 5-5).

Table 5-5: Mean values and standard deviations for the acceptance subscales. Values could range between 1 = I don't agree at all and 7 = I fully agree.

Group	*Behavioral Intention to Use*		Perceived Safety	
	M	*SD*	*M*	*SD*
L2H-on	5.08	1.21	5.71	0.98
L2H-off 5s	4.82	1.59	5.03	1.89
L2H-off 3s	4.80	1.19	5.28	1.97

NDRT involvement and intention

When looking at the number of tasks worked on throughout the study, there are no significant differences between the three groups ($F(2, 58) = .07$, $p = .932$), indicating that participants of all three groups were equally involved.

After the experimental drive participants were asked to fill out a NDRT-related questionnaire (adapted by Metz et al., 2014, see also e.g., Study 1) to assess if they could imagine to engage in any of the activities described. The respective ANOVAs, however, revealed no significant differences regarding intended involvement in NDRTs between the three L2 groups, indicating that no group plans on engaging stronger in any (inappropriate) activities while driving with an activated L2 function.

5.4.3.2 Eyetracking Data

Number eyes-off road glances > 2 s

The one-way ANOVA revealed a significant difference for the number of eyes-off road glances above 2 s when using a L2 function ($F(2, 51) = 3.49$, $p = .038$, $\eta_p^2 = .12$). Post-hoc Tukey-corrected t-tests were conducted, which showed that the significant difference existed between the L2H-off 5 s and the L2H-off 3 s groups ($p = .036$). The L2H-off group with the 3 s DMS had thereby significantly less eyes-off road glances above 2 s, which could be attributed to the shorter timed DMS. The L2H-on group did not differ significantly from any other group (see Figure 5-8).

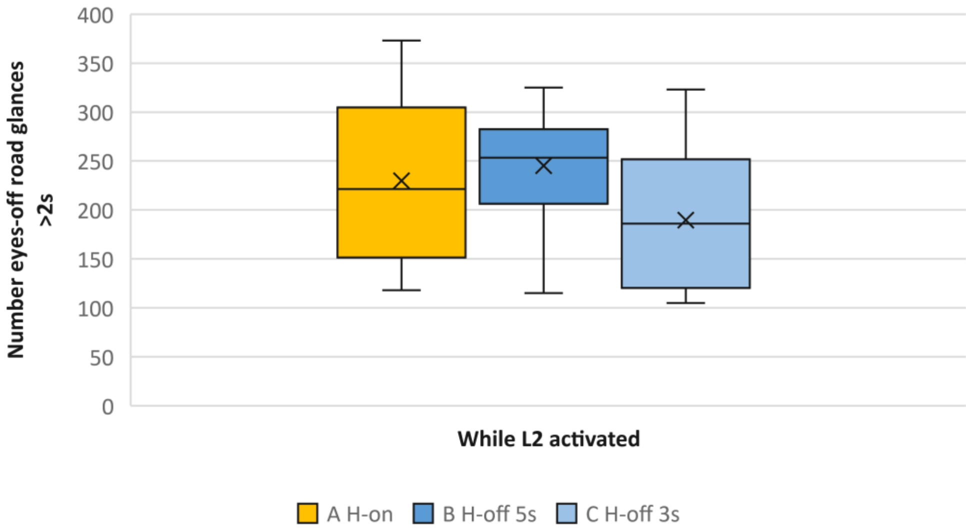


Figure 5-8: Boxplots for number of eyes-off road glances above 2 s while driving with an activated L2 function.

Visual attention ratio

The one-way ANOVA revealed a significant difference for the category “eyes-on other” ($F(2, 52) = 6.10$, $p = .004$, $\eta_p^2 = .20$). Post-hoc Tukey-corrected t-tests were conducted, which revealed that significant differences existed between the LH-on group and the two LH-off groups (H-on vs. H-off 5 s: $p = .003$; H-on vs. H-off 3 s: $p = .043$). This finding indicates that the L2H-on group attributed significantly more visual attention on areas other than the road (instrument cluster or NDRT; other, undefined areas) compared to both L2H-off groups, which could be attributed to the DMS which guides participants' visual attention repeatedly and regularly back to the driving relevant areas (see Figure 5-9).

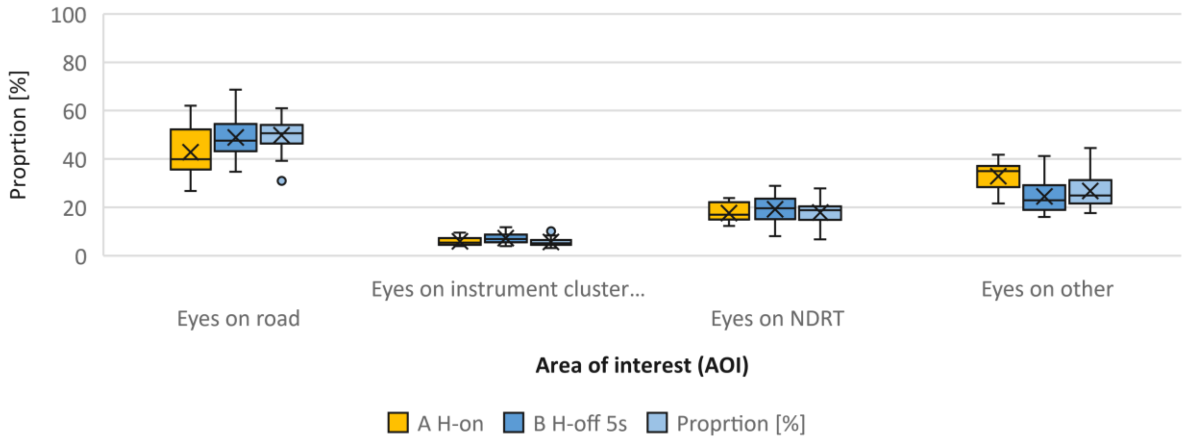


Figure 5-9: Boxplots for proportion of visual attention on the areas road, instrument cluster, NDRT and other.

5.4.4 Simulator Data

5.4.4.1 Overall

Number of H-off (HOR) and Eyes-off requests (EOR)

Statistically, only the two L2H-off groups were compared as the L2H-on group's DMS is too different (in terms of timing and requested behavior) from the L2H-off groups' DMS. However, Table 5-6 summarizes the mean values and standard deviations for all three groups. T-Tests revealed that there are no significant differences between the two L2H-off groups regarding the number of warnings over all three warning stages (Stage 1: $t(33) = -0.51$ $p = .615$; Stage 2: $t(33) = -0.55$, $p = .588$; Stage 3: $t(33) = -0.30$, $p = .764$). This result indicates that a DMS that warns after 3 s of gaze aversion does not automatically lead to significantly more warnings.

Another descriptive result is that there are few to none "Warning Stage 3" occurrences over all three groups. This shows that most participants reacted after the first or second warning stage. DMS with three-staged warnings appear to be sufficient for motivating users to behave appropriately (driving with hands-on when using L2H-on functions and driving eyes-on when using L2H-off functions). Possibly, even 2 stages could be sufficient.

Table 5-6: Mean values and standard deviations for the number of HOR and EOR received by each group.

Group	Warning stage 1		Warning stage 2		Warning stage 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
L2H-on	1.05	1.16	0.24	0.70	0.00	0.00
L2H-off 5s	6.88	9.04	5.65	8.98	0.41	0.80
L2H-off 3s	9.50	19.33	8.44	19.15	0.50	0.92

Mean reaction time to H-off (HOR) and Eyes-off requests (EOR)

A Welch ANOVA revealed no significant difference ($F(2, 31.255) = 1.161$, $p = .326$) between the L2H-on and the L2H-off groups when looking at the mean reaction time to HORs or EORs (Figure 5-10). Participants of all groups reacted equally fast (L2H-on $M = 2.05$, $SD = 1.23$; L2H-off 5 s: $M = 2.89$, $SD = 2.26$; L2H-off 3 s: $M = 2.75$, $SD = 2.32$).

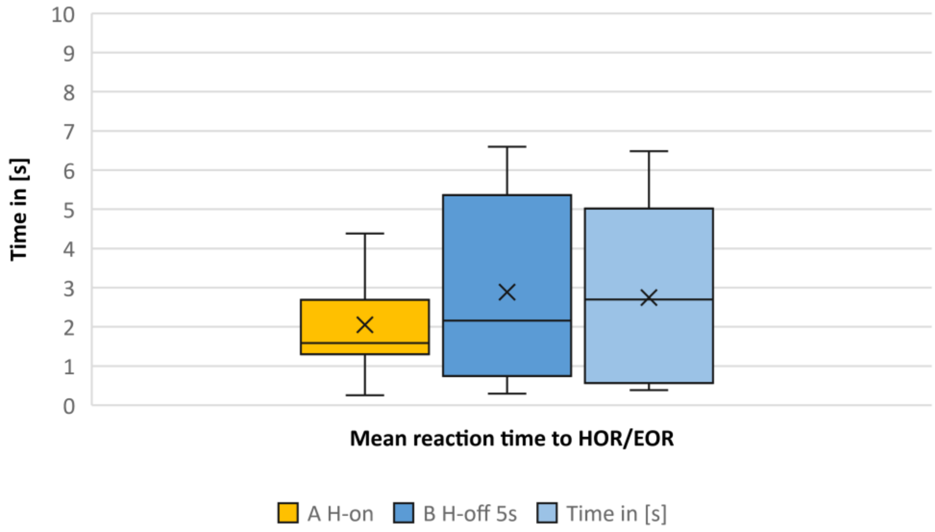


Figure 5-10: Boxplots showing the mean reaction time to HORs and EORs.

Time Hands-free driving

An ANOVA revealed a significant difference ($F(2, 56) = 74.726$, $p < .001$, $\eta_p^2 = .727$) between the proportion of hands-free driving while driving with activated L2 function in relation to the total time driven in level 2 of the H-on group ($M = 16.13\%$, $SD = 19.60\%$), H-off 5s group ($M = 81.66\%$, $SD = 26.85\%$), and H-off 3s group ($M = 89.96\%$, $SD = 16.34\%$) (Figure 5-11). Post-hoc Tukey-corrected t-tests were conducted, which showed a significant lower proportion for the L2H-on group than for the L2H-off 5s group ($p < .001$) and the L2H-off 3s group ($p < .001$), but no significant difference between the L2H-off 5s and L2H-off 3s group ($p = .453$).

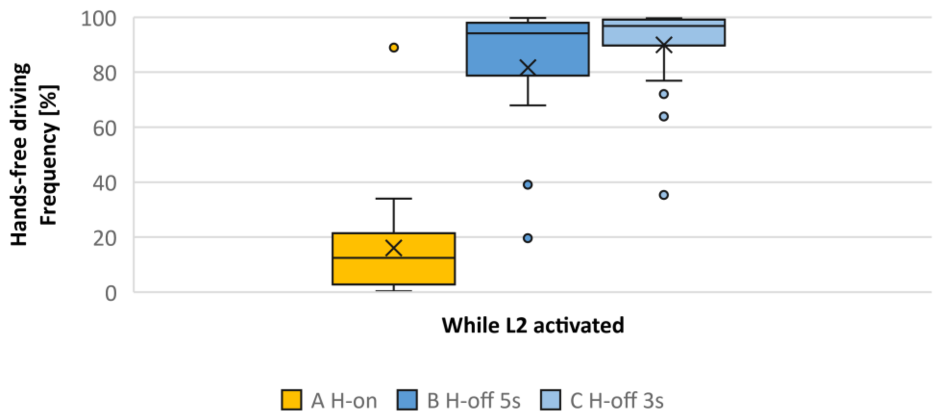


Figure 5-11: Boxplots showing the proportion of hands-free driving for the three experimental groups while driving with activated L2 function.

Number of attempted activations of L2, although not available

A Welch ANOVA was performed due to missing homogeneity of variances. The analysis showed no significant difference ($F(2, 31.354) = 2.102$, $p = .139$) regarding the number of failed level 2 activations between the three groups. Descriptively, the number of failed L2 activations attempts was low for all three groups (L2H-on group: $M = 2.14$, $SD = 3.12$; L2H-off 5s group: $M = 1.11$, $SD = 1.97$; L2H-off 3s group: $M = 0.70$, $SD = 1.03$).

Number and description of collisions

In Scenario 1, 2 and 3 no collisions occurred. In the cut-out scenario (scenario 4), 19 participants experienced the FDCR 2.7 s before collision (L2H-on: $n = 7$; L2H-off 5 s: $n = 4$; L2H-off 3 s: $n = 8$). 13 of these participants could handle the situation after the FDCR without colliding with the broke-down vehicle. This result indicates that the timing of the FDCR was appropriate for the majority of participants to assist in coping. Six participants, however, were unable to successfully manage the situation, resulting in respective six collisions over all three groups. Speed at collision and participants behavior before or during the collision is summarized in Table 5-7.

Table 5-7: Description of incidents.

Group	**Number of incidents**	**Speed at incident (km/h)**		**Behavior before/while incident (qualitative analysis based on video data)**
		M	*SD*	
L2H-on	$n = 1$	107.29	0.00	VP17 (male, 66 years): involved in NDRT, reacts to FDCR (before collision) with braking, appears very surprised
L2H-off 5 s	$n = 3$	94.02	16.99	VP44 (female, 29 years): reacts to FDCR (before collision) by putting the hands back to the steering wheel and braking VP47 (male, 27 years): reacts to FDCR (before collision) by putting the hands back to the steering wheel and braking appears not surprised VP50 (female, 28 years): reacts to FDCR (before collision) by putting the hands back to the steering wheel, braking and steering
L2H-off 3 s	$n = 2$	75.13	33.26	VP19 (male, 27 years): reacts to FDCR (before collision) by putting the hands back to the steering wheel and braking VP58 (male, 66 years): involved in NDRT, reacts to FDCR (before collision) by putting the hands back to the steering wheel with braking and steering, appears very surprised

Descriptively, collision occurred more often in the L2H-off groups than in the L2H-on group. One assumption might be that this could be due to the time needed to put the hands back on to the steering wheel. However, there are observations that could lead to rejecting this conclusion. One observation is that all participants of the L2H-off groups experiencing a collision managed to put their hands back on to the steering wheel before collision. Furthermore, the data show that 35 of the 40 participants experiencing the L2H-off functions and even the majority of participants experiencing the FDCR in this respective scenario were able to respond successfully. Overall, driving hands-free might be less the challenge, but the available anticipation time might be. Therefore, it could be considered if the FDCR should be presented earlier than 2.7 s to enable more participants to react successfully. In sum, collision occurrence seems to be no systematic, group or function related event, but rather an individual coping problem.

5.4.4.2 Scenario 1 and 3

Scenarios 1 and 3 primarily focused on participants noticing the indicated change in speed, the absence of automatic vehicle response and the need to intervene themselves. The focus is therefore on longitudinal behavior (braking and/or deceleration).

Results show that in both scenarios, some participants of all three groups missed to react accordingly (by braking or set-speed adaptation) to the changed speed limit and the missing vehicle reaction. Descriptively, while over 50% of the L2H-on group responded to the respective system limit, it was slightly less in the two L2H-off groups (see Table 5-8). Either, the L2H-off groups may not have benefitted from the gaze-based DMS in either scenario (unintentional misbehavior) or L2H-off participants are less inclined to actively react to the maladapted set-speeds (intentional misbehavior/misuse).

Table 5-8: Frequency with which participants reacted to the changed speed limit with braking or deceleration.

Group	Scenario 1	Scenario 3
L2H-on	62%	67%
L2H-off 5s	50%	39%
L2H-off 3s	40%	60%

A Welch ANOVA was performed for scenario 1 and an ANOVA for scenario 3 to assess whether the three groups differed in their reaction time (braking/deceleration input). Both analyses showed no significant difference (scenario 1: $F(2, 13.129) = .615$, $p = .555$; scenario 3: $F(2, 30) = .039$, $p = .961$). Participants who reacted with braking or deceleration did so at similar rates across all three groups (see Figure 5-12).

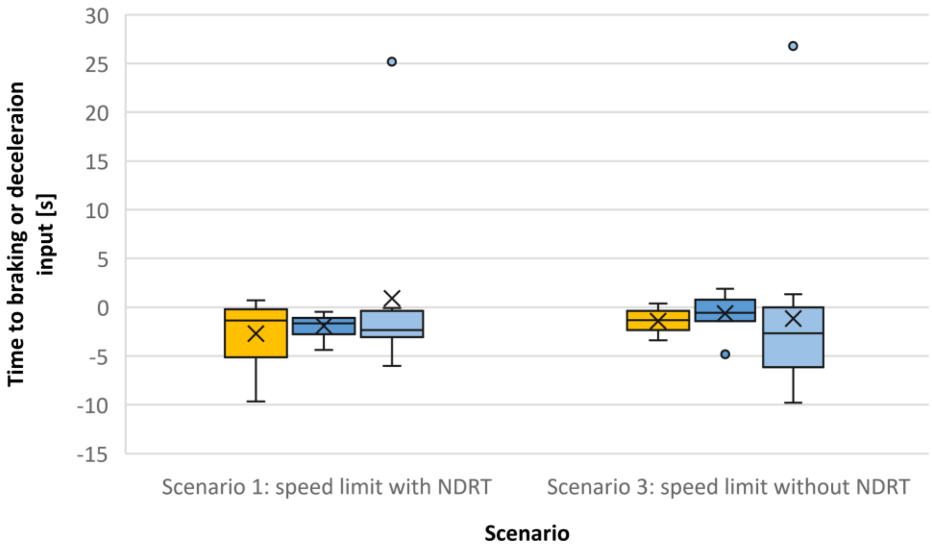


Figure 5-12: Boxplots showing the time to braking or deceleration input. The number 0 on the y-axis represents the point in time when the sign with the new speed limit was passed.

5.4.4.3 Scenario 2 and 4

Scenarios 2 and 4 primarily focused on participants reacting to blocked lanes with longer (roadworks with first indication 600 m beforehand) and shorter (cut-out with 4 s reaction time) anticipation time. As steering input is necessary for successfully handling this scenario, not only longitudinal, but also lateral distance is focused on.

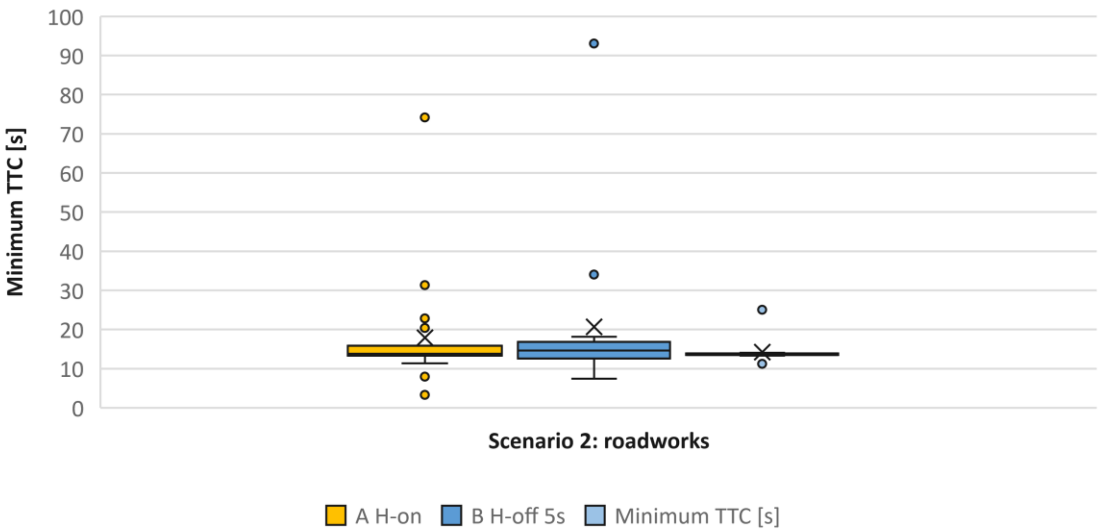


Figure 5-13: Boxplots showing the minimum TTC in scenario 2

Minimum TTC

For scenario 2 and scenario 4, minimum TTC was analyzed. The minimum TTC refers to the foremost point of the ego vehicle and the rearmost point of the front vehicle. For scenario 2, no significant differences (*F*(2, 21.389) = 1.154, *p* = .334) were found between the three groups (L2H-on group: *M* = 17.90, *SD* = 15.20; L2H-off 5s group: *M* = 20.64, *SD* = 20.84; L2H-off 3s group: *M* = 14.21, *SD* = 2.78), indicating that all three groups reacted equally fast to the blocked lane by either braking and/or steering (see Figure 5-13).

For scenario 4, no significant differences (*F*(2, 48) = 1.564, *p* = .220) were found between the three groups either (L2H-on group: *M* = 1.29, *SD* = 0.80; L2H-off 5s group: *M* = 1.63, *SD* = 0.96; L2H-off 3s group: *M* = 1.09, *SD* = 0.56), indicating that all three groups reacted equally fast to the breakdown car by either braking and/or steering (see Figure 5-14).

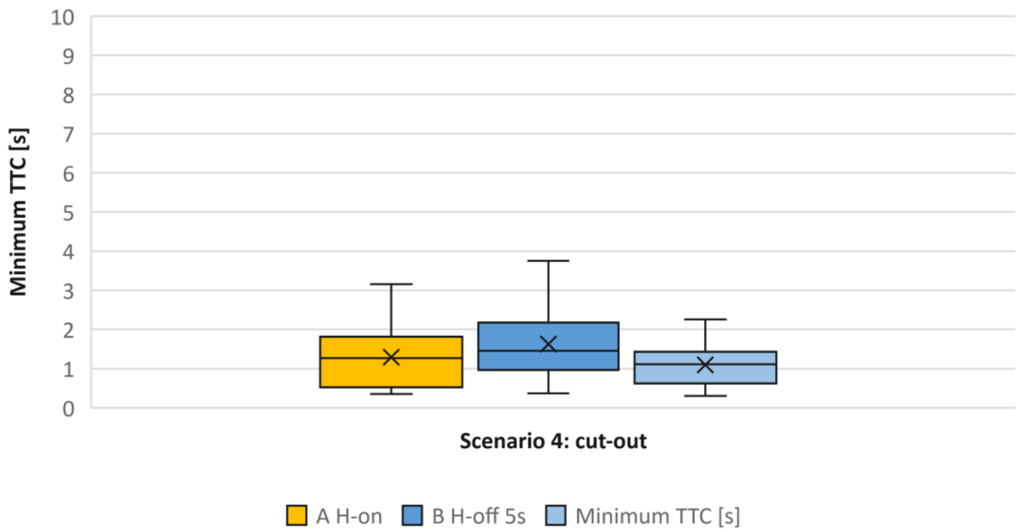


Figure 5-14: Boxplots showing the minimum TTC in scenario 4.

Overall, it appears that participants seem to have reacted earlier in scenario 2 than in scenario 4. This finding can be attributed to the different anticipation time given in the respective situation. While in scenario 2 the first indication for the upcoming necessary lane change was given 600 m (~18 s) before end of lane, is was given just 4s before colliding with the brokedown car in scenario 4.

Minimum lateral distance

For scenario 2 and scenario 4, minimum lateral distance to the respective objects blocking the lane were analyzed. Minimum lateral distance means the actual distance between ego vehicle and obstacle (outside to outside). For both scenarios, no significant differences were found between the three groups (scenario 2: *F*(2, 48) = 1.564*, p* = .220; scenario 4: *F*(2, 47) = 0.289, *p* = .750), indicating that in both scenarios all three groups seem to have steered around the obstacles with a similar distance (see Figure 5-15). The outliers in scenario 4 show the participants who caused a collision with the breakdown vehicle. Speed during the intervention period (start of intervention + 5 s) were taken into account. For Scenario 2, these range between 80 and 120 km/h. For Scenario 4, speed ranges between 87 and 120 km/h with one person driving remarkably slower with 36 km/h.

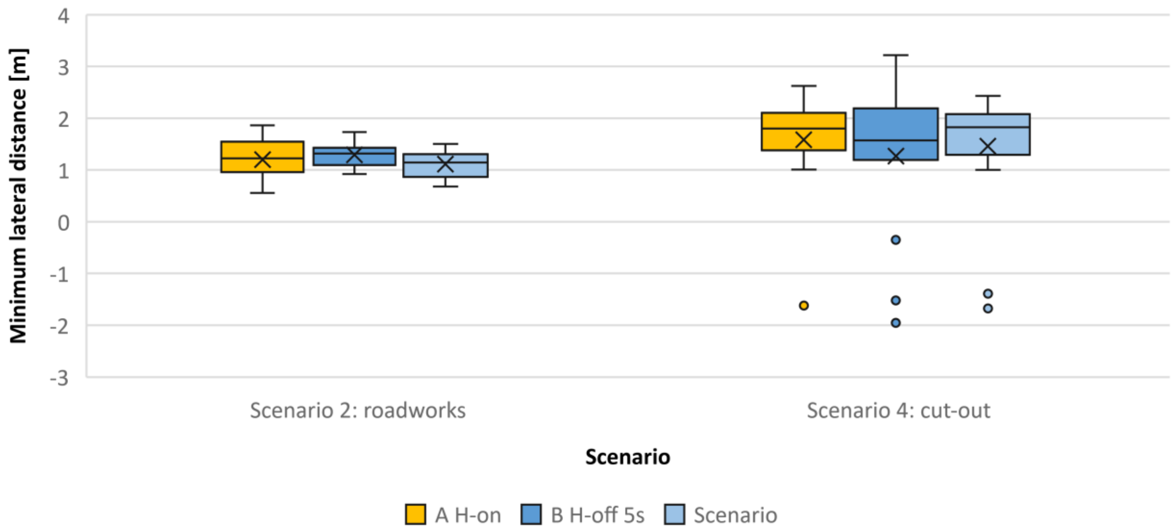


Figure 5-15: Boxplots showing the minimum lateral distance to the respective objects blocking the lane.

5.4.5 Summary / Conclusion

Similar to Study 1, this study (Study 3) focuses on the research questions DMS timing ("Attentiveness Alert (AR)") and the first challenge "hands-off = mind-off?" (CQ1). Based on the results of Study 3 the following main research questions should be answered:

- RQ1: Are there differences between L2H-off functions with eyes-on requests (EOR) and L2H-on functions with hands-on request (HOR) with regards to attention and user behavior?
- RQ2: Are there differences between L2H-off functions with differently timed eyes-on requests (EOR) with regards to attention and user behavior?

Regarding RQ1, it was assumed that users of an L2H-off function with EOR should show at least as good/safe driving performance and handling of ODD limits as users of an L2H-on function with HOR. Regarding RQ2, it was assumed that a DMS that cautions the driver to pay visual attention to the road after 3 seconds eyes-off road may enable the driver to notice obstacles or silent system failures earlier, especially during difficult-to-anticipate, time-critical ODD limits, than a DMS that cautions the driver for the first time after 5 seconds of inattention (see e.g. Euro NCAP, 2022; Victor et al., 2018; Simons-Morton et al., 2014; Schneider et al., 2022).

The study results will first be interpreted with regards to the main RQs. Subsequently, the results are discussed and interpreted in light of the projects CQs.

Main research questions of the study

Although participants rated their monitoring performance equally well, eyetracking data revealed some differences in terms of visual attention between the groups. Data show that participants using the L2H-on function had more visual attention on other areas than road, instrument cluster or NDRT, compared to the two L2H-off group. This result indicates that gaze based (eyes on) DMS comparatively prevented or decreased visual attention to those areas irrelevant to the driving task. Furthermore, it was found that there were less eyes-off road glances above 2 s when monitored by the DMS reacting after 3 s of gaze aversion compared to the DMS intervening after 5s. One additional observation at this point is that the design of three warning stages seems appropriate/sufficient as there were almost no warnings after stage 2 given. Eventually, two warning stages could also be sufficient.

Another important finding is that both L2H-off groups used the hands free driving option when L2 was activated and actually had their hands off the steering wheel. However, there were almost no significant differences regarding the timing of actions in each scenario across the three groups.

Regarding reactions to changed speed limits in scenario 1 and 3, recorded data show that there are participants in all three groups who did not react to the silent system failure, i.e. missing speed adaptation by the system. There are two explanations. On the one hand, participants might not have noticed the changed speed limit and the missing vehicle response to that change. In this case, the DMS mechanisms of both L2H-on and L2H-off functions would not have been sufficient to draw the attention of the drivers to this circumstance. On the other hand, participants might have noticed the change in speed limit and the missing vehicle response, but actively decided not to react, since surrounding traffic or the integrated ACC did not make a direct intervention necessary. This assumption is strengthened by the fact that (almost) all participants across all three groups indicated that they were aware that they had to react to speed limits actively at times throughout the study. Therefore, the missing reaction of some participants might not be attributed to DMS design but to intentional behavior contrary to traffic regulations.

Scenario 2 shows that the L2H-off 3 s group reacted later than the L2H-off 5 s group. One explanation might be that participants experiencing the L2H-off 3 s function are earlier aware of the ending lane and react later but better prepared to the blocked lane while participants experiencing the L2H-off 5 s function react more directly as they might have recognized the actual end of lane later. However, a detailed analysis of input quality would be necessary to support this hypothesis. Importantly, regardless of reaction timing, all participants were able to sovereignly handle the situation.

In scenario 4, no differences in reaction time could be found. In sum, six collisions occurred in this scenario. However, it should be kept in mind that this scenario was intentionally designed to be time-critical as well as difficult to anticipate and thus represents a rather exceptional situation, which demands high driver readiness and skills. This was also apparent in the handling of this type of scenario in Study 1, where even drivers without driver assistance (L0; manual driving) triggered an emergency braking maneuver instead of intervening in time. Based on video analysis, individual participant examination and the fact that collisions occurred across all groups, we come to the conclusion that those events are not systematic for one L2 design, i.e. not function related, but reflect rather individual coping problems in combination with a challenging scenario design.

In sum, the three groups did not differ in takeover performance or in coping with system limits, indicating that performance was equally well when using L2H-on and L2H-off functions. This finding extends to the subjective data, as all three functions were equally trusted and accepted. However, it can be assumed that gaze based DMS giving EOR (as used with L2H-off functions) draw users visual attention more strongly to the primary driving task. For the 3 s DMS there were significantly less glances off road above 2 s compared to the 5 s DMS. Besides this result, no differences between the two L2H-off DMS designs could be found, indicating that both systems work and are assessed equally well. Regarding future function design, both timings could thus be pursued. However, giving the first EOR after 3 s might involve drivers' visual attention slightly more, but would need to be analyzed closely with regard to potential side effects on acceptance (disuse) and the reliability of DMS alerts (false alarm rate).

CQ1 Hands-off = mind-off?

Based in the current study results, there are no indications that L2H-off functions lead to users being less engaged (mind-off) compared to L2H-on functions. In fact, the gaze based (eyes on) DMS seems to positively influence participants' attention by regularly directing drivers' gaze to the road and the relevant surroundings. Indications for this conclusion come from e.g., the eyetracking data. Data shows that the L2H-on group with the hand posture based (hands on) DMS shows descriptively less eyes on road glances compared to the two L2H-off groups working with the gaze based DMS. Furthermore, we could find less eyes off road glances > 2 s for the L2H-off group experiencing the gaze based DMS that initiates a first AR after 3 s. Further indications for participants experiencing L2H-off functions reacting equally fast to system limits compared to L2H-on users indicate similar cognitive readiness to intervene (see next paragraph on CQ2).

CQ2 Prolonged transition times

Based on the current study results there are no indications for significantly prolonged transition times in interaction with L2H-off functions compared to L2H-on functions. E.g., when looking at the mean reaction times to HORs and EORs, no differences could be found between the L2H-off groups and the L2H-on group. Furthermore, there were no differences in reaction time to changed speed limits or when confronted with the broke down car in scenario 4, indicating that even in hard-to-anticipate and time-critical situations, the response time using L2H-off functions is the same as for L2H-on functions. The reaction time results of scenario 2 (roadworks) revealed that participants experiencing the 3 s DMS reacted significantly later in the roadworks scenario compared to the group experiencing the 5 s DMS. However, no collisions occurred in this respective scenario in neither group. Therefore, the different reaction times do not indicate a more or less successful coping behavior.

CQ3 Foreseeable misuse

This study did not explicitly investigate the potential for misuse. Overall, however, data of the current study provides no evidence that the use of L2H-off functions leads to greater or more frequent misuse compared to L2H-on functions. This conclusion is based on the questionnaire results regarding intended involvement in NDRTs, where no statistical difference could be found between the three L2 groups. Overall, descriptive results indicate that neither users of L2H-on, nor users of L2H-off functions plan on engaging stronger in any (inappropriate) activities while driving with activated L2 automation. Furthermore, questionnaire results show that trust and perceived safety are given, but not high, probably indicating that misuse may be rather unlikely when using L2H-on and L2H-off functions.

CQ4 Mode confusion

In sum, the results of this study provide no evidence that mode confusion occurred (more frequently) when using a L2H-off function compared to using a L2H-on function. For all L2 functions assessed within the current study a rather good to very good understanding of system functionality, system limits and driver responsibilities was observed, which is an essential prerequisite for mode awareness.

CQ5 Safety

Regarding CQ5, the data on coping with scenario 2 and 4 are probably the most interesting together with the findings on visual attention discussed for CQ1. In scenario 2 (roadworks), which can be anticipated earlier compared to scenario 4, no collisions occurred. In the rather difficult to anticipate and time critical scenario 4 (cut-out), six collisions occurred over all three groups. As collisions occurred in each group and the respective participants differed regarding e.g., age or behavior before and during the collision, those events don’t seem to be systematic, group or function related. They are more likely to reveal individual coping problems. Furthermore, we found no significant differences between L2H-on and L2H-off groups in terms of minimum TTC or minimum lateral distance, indicating similar coping behavior.

5.4.6 References

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29* (4), 552-557. doi:10.7334/psicothema2016.383

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction, 35*(6), 456-467. doi:10.1080/10447318.2018.1456150

Körber, M. (2019). Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (Vol. 823, pp. 13–-30). Cham: Springer International Publishing. doi:10.1007/978-3-319-96074-6_2

Metz, B., Landau, A., & Just, M. (2014). Frequency of secondary tasks in driving – Results from naturalistic driving data. *Safety science* (68), 195-203. doi:10.1016/j.ssci.2014.04.002

Nusholtz, G. S., Hsu, T. P., Gracián, M. A., Prado, J. A., Cedeño, R., Santillán, N. M., & Herrera, D. (2013). Forward Collision Warning Timing in Near Term Applications. *SAE International Journal of Transportation Safety, 1*(2), 467-479. doi:10.4271/2013-01-0727

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, (*pp. 51-58). doi:10.1145/2390256.2390264.

European Statement of Principle (ESoP) on human machine interface (2005). 12*th* World Congress on Intelligent Transport Systems, San Francisco, USA.

Schneider, N., Ahrens, L., & Pruksch, A (2022). Controllability of lateral drift failures while driving with SAE Level 2 Advanced Driver Assistance Systems. *14. Workshop Fahrerassistenz und automatisiertes Fahren*.

Simons-Morton, B. G., Guo, F., Klauer, S. G., Ehsani, J. P., & Pradhan, A. K. (2014). Keep your eyes on the road: Young driver crash risk increases according to duration of distraction. *Journal of Adolescent Health*, *54*(5), 61-67. doi:10.1016/j.jadohealth.2013.11.021

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human factors, 60*(8), 1095-1116. doi:10.1177/001872081878816

5.4.7 Appendix

Items to assess driver role understanding:

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte **Fahrerassistenzsystem** .

	Nicht zutreffend	Zutreffend	Unsicher
Ich muss das Level 2 System stets überwachen, wenn das System aktiviert ist und mich vergewissern, dass das System die Fahraufgabe entsprechend den aktuellen Bedingungen (z. B. Geschwindigkeitsbegrenzungen, Abstand zu anderen Verkehrsteilnehmern, Straßenbedingungen) durchführt.	○	○	○
Ich darf mich mit anderen Tätigkeiten, wie z.B. E-Mails schreiben, beschäftigen, wenn das Level 2 System aktiviert ist.	○	○	○
Ich muss zu jeder Zeit bereit sein die Fahraufgabe übernehmen zu können.	○	○	○
Ich kann jederzeit in die Automation eingreifen.	○	○	○
Ich darf mich von der Überwachung des Verkehrsraums abwenden, wenn das L2 System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind.	○	○	○
Ich muss dem Straßenverkehr die gleiche Aufmerksamkeit widmen wie beim Fahren ohne Fahrerassistenzsysteme (z. B. Straßenzustand, Interaktion mit anderen Fahrern).	○	○	○
Ich muss die gleiche oder eine ähnliche Sitzposition wie beim Fahren ohne Fahrerassistenzsysteme einnehmen.	○	○	○

Items to assess system understanding

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte **Fahrerassistenzsystem**.

	Nicht zutreffend	Zutreffend	Unsicher
Das Level 2 System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer.			
Das Level 2 System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden.			
Wenn das Level 2 System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit.			
Das Level 2 System passt die Geschwindigkeit an langsamer vorausfahrende Fahrzeuge an, außer diese stehen oder die Differenzgeschwindigkeit ist zu hoch.			
Das Level 2 System erkennt immer, wenn es eine Situation nicht meistern kann.			
Das Level 2 System lenkt selbstständig innerhalb des eigenen Fahrstreifens.			
Wenn das Level 2 System aktiv ist, hält es einen voreingestellten Mindestabstand zum vorausfahrenden Fahrzeug ein.			
Wenn das Level 2 System aktiv ist, steuert es die Geschwindigkeit des Fahrzeugs entsprechend der eingestellten Geschwindigkeit.			
Wenn das Level 2 System aktiv ist, führt es Fahrstreifenwechsel durch, ohne dass der Fahrende dabei selbst lenken muss.			
Wenn ich, während das Level 2 System aktiviert ist, bremse, fallen alle Assistenzfunktionen weg und ich muss selbstständig lenken sowie Geschwindigkeit und Abstand regeln, bis ich diese wieder aktiviere.			
Wenn ich, während das Level 2 System aktiviert ist, lenke (um bspw. einem Hindernis auszuweichen), fährt das Fahrzeug mit der eingestellten Höchstgeschwindigkeit weiter.			
Wenn ich, während das Level 2 System aktiviert ist, beschleunige, reduziert das Fahrzeug die Geschwindigkeit anschließend wieder auf die eingestellte Höchstgeschwindigkeit, wenn ich den Fuß wieder vom Gaspedal nehme.			
Wenn, während das Level 2 System aktiviert ist, das Signal „Bitte übernehmen" im Display erscheint, fallen alle Assistenzfunktionen weg und ich muss selbstständig lenken sowie Geschwindigkeit und Abstand regeln.			

5.5 Anchor Study (Study 4)

Documentation by J. Josten, D. Schwarze (fka GmbH)

Apart from the qualitative user survey in the USA (cf. Section 4.3), all studies within this project described so far included participants without prior experience with L2H-off functions, due to being conducted in Germany. One goal of this fourth and last simulator study (i.e., Anchor Study) is thus to investigate whether differences in usage behavior might be expected from drivers who are experienced in the interaction with EOD-based (EOD = eyes-on detection) L2 functions. Since such functions are already available in the USA, a parallel study was set up in the fka SV driving simulator in Santa Clara, CA, USA, and in the driving simulator at fka GmbH in Aachen, Germany, to compare differently experienced user samples in the exact same driving setting. This anchor study specifically assessed two research foci:

RQ1: Are L2H-off and L2H-on systems comparable with regard to:

a. Detection of intervention needs by the driver (controllability of system failures, i.e. lateral drift)?
b. Controllability of system-initiated deactivations (function direct control request (FDCR) timing at operational design domain (ODD) limit)?

RQ2: Can findings be generalized over different samples regarding

a. prior experience with L2 systems (none vs. EOD vs. hands-on detection (HOD))?
b. cultural differences (US vs. DE sample)?

The first research question (RQ1.a) aims to evaluate to what extent L2H-off and L2H-on functions are comparable in the users' detection and compensation of a failure in lateral assistance. This scenario is based on a study by Schneider, Ahrens and Pruksch (2022) who investigated the detection and compensation of silent failures in the steering system when using a L2 system hands-free on a test track. In the study by Schneider et al. (2022), not all participants intervened in time, i.e. before leaving the lane. The authors found the same behavioral pattern in all of these cases, namely a preparation to intervene by moving the hands towards the steering wheel with a subsequent retraction of the hands instead of providing active steering input. An additional finding was that drivers who did not intervene expressed high levels of trust in the function. The authors conclude "that the problem is not the driver's ability to handle limited lateral failure dynamics when driving hands-free but rather a cognitive misattribution of the systems capability which is build up by experience of the system and user expectations" (Schneider et al., 2022, p. 190). As the study by Schneider et al. (2022) did not include a direct comparison between different L2 designs, this study adopts their research focus to broaden the scope towards functional design (possibly influencing the preparedness to detect and compensate steering failures) and prior experience (possibly influencing the level of trust in and expectations towards L2 functions).

The second question (RQ1.b) aims at interventions in case of upcoming, visible ODD limits that require a driver intervention. The question here is if drivers detect the ODD limit in time and how they chose to handle the system limit. As a fallback option in case drivers do not deactivate the system on their own accord, two different FDCR timings will be investigated (2 s vs. 4 s). The FDCR timings were selected based on literature and project internal discussions, in which a time of at least 2.7 s – 3 s was deemed necessary for successful reactions (see also Damböck et al., 2012; Mok et al., 2015). It is expected that drivers who are frequently reminded to visually monitor the road scene (as common for L2H-off functions) should realize the need to intervene and be able to compensate any potential disadvantage resulting from monitoring hands-free.

The rather qualitative second focus of this study shall provide an indication how well requirements derived for function design might be generalized over different driver populations (RQs 2a and 2b). However, the impact of experience in this investigation has to be considered with regard to different aspects. Firstly, drivers do not interact with their own L2 system, but with the prototypical L2 systems under test, e.g. regarding DMS criteria or HMI design. This reduces the potential for a direct transfer of experience, as drivers will first need to establish a basic similarity between their own and the function under test. Further, drivers were not interacting with the function on familiar routes where they know how the function normally behaves or when they can likely trust the function to carry out longitudinal and lateral control on most days without any need for driver intervention. This might further reduce differences between experienced and novice users in monitoring behavior. The US survey has shown (cf. Section 4.3) that at least some L2 users strategically adapt their interaction with the function depending on function knowledge and the current driving situation. This effect will likely be reduced by the experimental setup and prototypical function. To allow for a maximized transfer of knowledge between daily life and prototypical functions, consideration was given to the realization of a realistic function design (e.g. activation of L2 via ACC-mode, speed adaptations via hard keys) and system instructions.

5.5.1 Method

To address research foci and answer the related questions, sample criteria and function design were varied within this study. Function design (L2H-on function with HOD and L2H-off function with EOD) was varied between subjects and four subsamples were recruited, differing in prior ADAS experience, especially with L2 functions (experienced versus novice), and in experience with different traffic conditions, by comparison of a US and German sample. L2 experience was defined as having at least one month of usage experience to be able to include enough users of relatively new L2H-off functions. In Germany, extended ACC experience was also accepted as a substitute for L2 experience. An overview of the samples can be found in Figure 5-1. Furthermore, participants were required to have five years of driving experience and a general interest in technology or assistance systems to balance the attitude towards technology between experienced and novice users in Germany. An appropriate age and gender distribution of the participants was intended, similar to prior studies of this project.

For the study conducted at fka GmbH in Aachen, the same simulator setup as in Study 2 and 3 was used (see description in Section 5.3.2). The simulator used for testing in Santa Clara, CA, at fka SV Inc. similarly covered an angular range of 220° around the driver onto a circular screen surface with a diameter of five meters. In contrast to the German setup, the rearview mirror was not covered during the study. The vehicle used in the US simulator setting was a BMW F13 6 series that has been modified for use in the driving simulator. The driver display was 295 mm wide and 112 mm high. The screen used for NDRT display was 155 mm wide and 95 mm high. The same software (Virtual Test Drive, VTD) was used for simulation purposes. Adaptations to the track were made for the US tests regarding visual alignment with US highways including signage and lane markings.

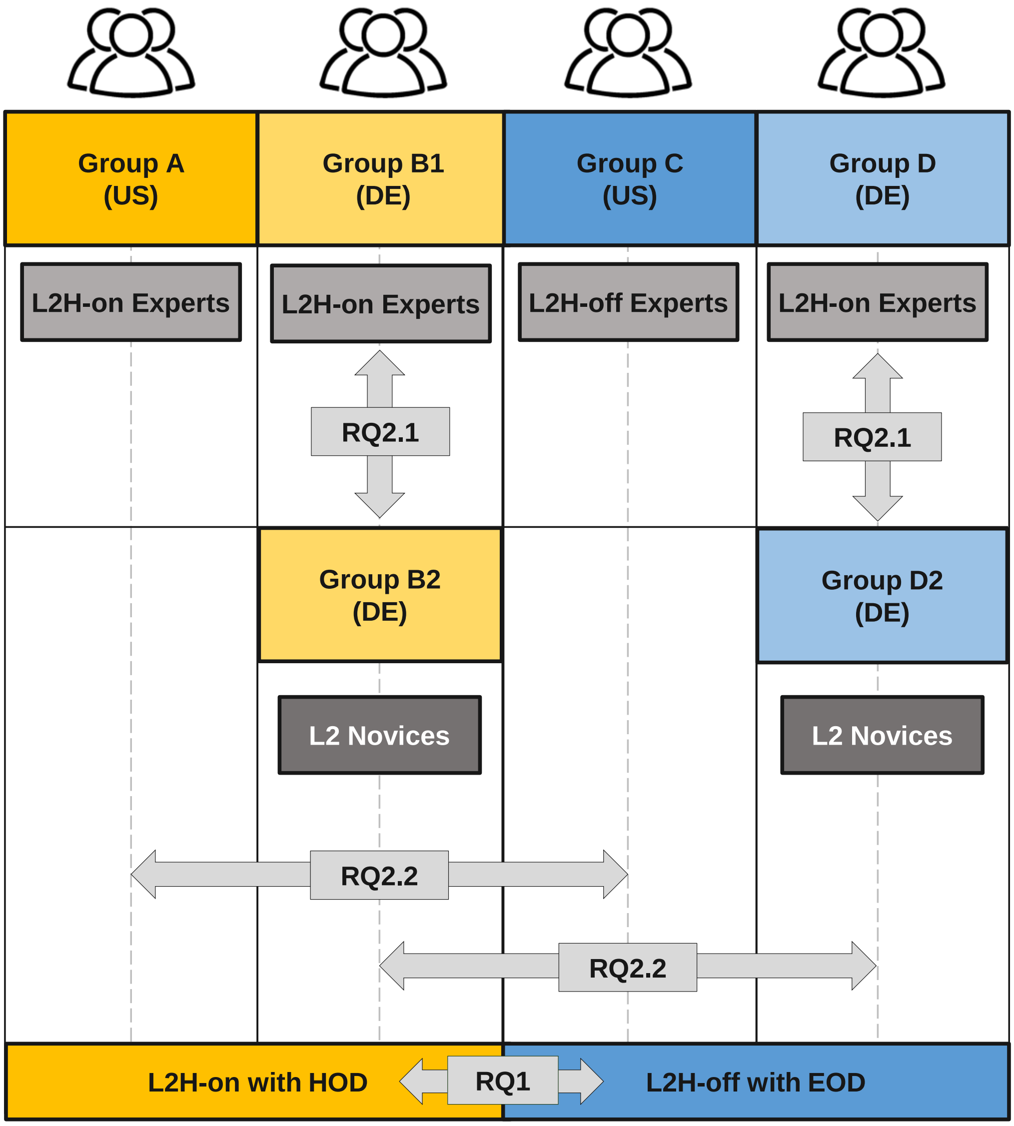


Figure 5-1: Overview over experimental groups and L2 functions in relation to research questions addressed (EOD = eyes-on detection; HOD = hands-on detection).

5.5.1.1 Simulator and system set-up

Hold-of-wheel detection was realized with the same equipment as in Study 2 and Study 3. Detection of contact to the wheel was thus not always possible for light or little contact. For hand posture detection at the beginning of transitions, video analysis was conducted to verify hand positions. Due to the availability of video data, depending on camera angle and seating position as well as recording issues, not all data sets could be verified. Gaze tracking for the L2H-off function was realized in the same manner as in Study 2 and 3 (see Section 5.3.2.4). Gaze data was recorded in both simulator settings using Tobii Glasses (see Section 5.3.2.4).

The multi-step system from Study 2, including an unassisted (manual), ACC and L2 mode (Section 5.3.7.2), was used for this study (see Figure 5-2). Standby of assistance modes was indicated by grey icons. ACC was always in standby mode. Assistance had to be activated by a button press on the steering wheel. Lateral support could be activated within the L2 ODD when longitudinal support (ACC) was active. Set-speed was 120 km/h (DE; US: 75 mph), but could be adapted manually. Distance settings could not be adapted. A medium distance setting was chosen, similar to Study 2 and 3.

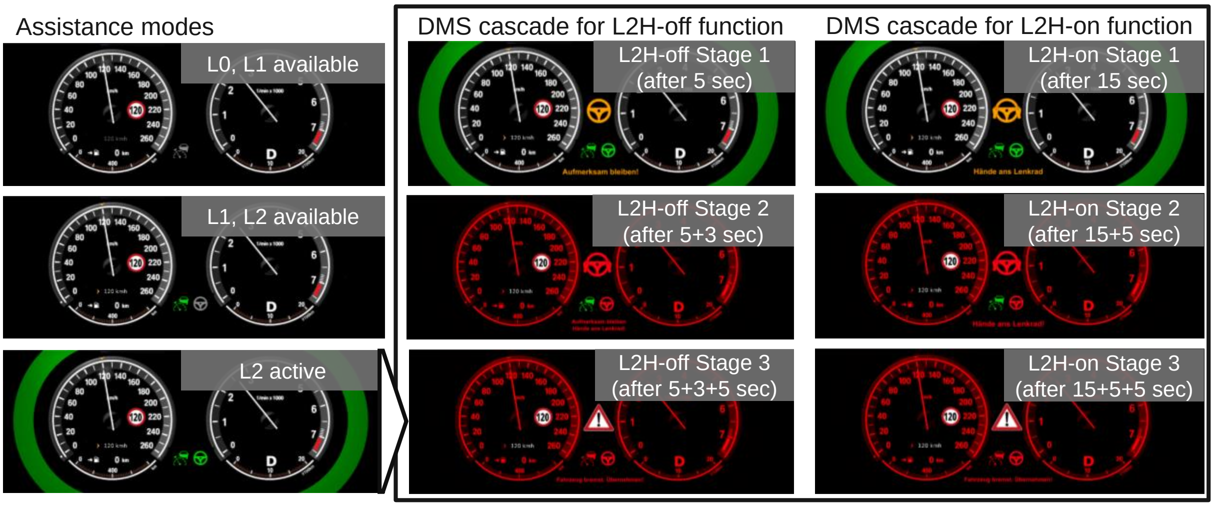


Figure 5-2: Overview on the (German) HMI design used within the study for all assistance modes (left) and for different DMS stages (from top to bottom) during L2 use for the two L2 functions (center: L2H-off; right: L2H-on).

5.5.1.2 Scenarios and procedure

The study took about 90 minutes to complete. The overall driving time per participant was about 40 minutes. The highway consisted of two lanes in each direction, predominantly with a Level of Service B (Transportation Research Board, 2000). The participants were instructed to obey the traffic rules and to keep on the right lane unless there was reason to change lanes or overtake. Given a continuous activation of the L2 function where possible, as instructed, manual (L0) and ACC (L1) use was heavily limited during the drive, restricting the potential of comparisons to other assistance modes (see Figure 5-3).

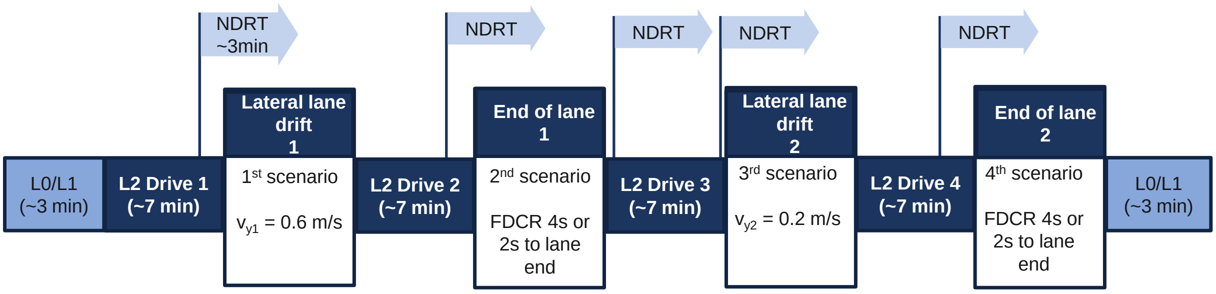


Figure 5-3: Schematic overview on the track design including the four scenarios.

Each participant encountered all four relevant measurement scenarios in the drive: two silent failures (lane drift), which differed regarding the time to contact with the barrier without driver intervention, and two ODD limits (lane end), which were indicated by road signs well ahead (see Figure 5-4).

The lane drift was designed as a challenging handling situation with no indication of the failure by the function itself but only by visual cues taken from lane positioning. The ODD limit (lane end) required the driver to both overrule the system and conduct a lane change or to take direct control after issuance of a FDCR issued in balanced order either 2 s or 4 s before lane end. No lead vehicle was present to block the view onto the upcoming situation. If the driver re-activated the L2 function after lane change, the FDCR was issued on the left lane (see Figure 5-4, bottom) at the same distance to the lane-merging point to indicate the ODD limit. For analysis, only the first intervention was analyzed and reactions to the FDCR on the left lane were not analyzed as no immediate change to control input was required in that case.

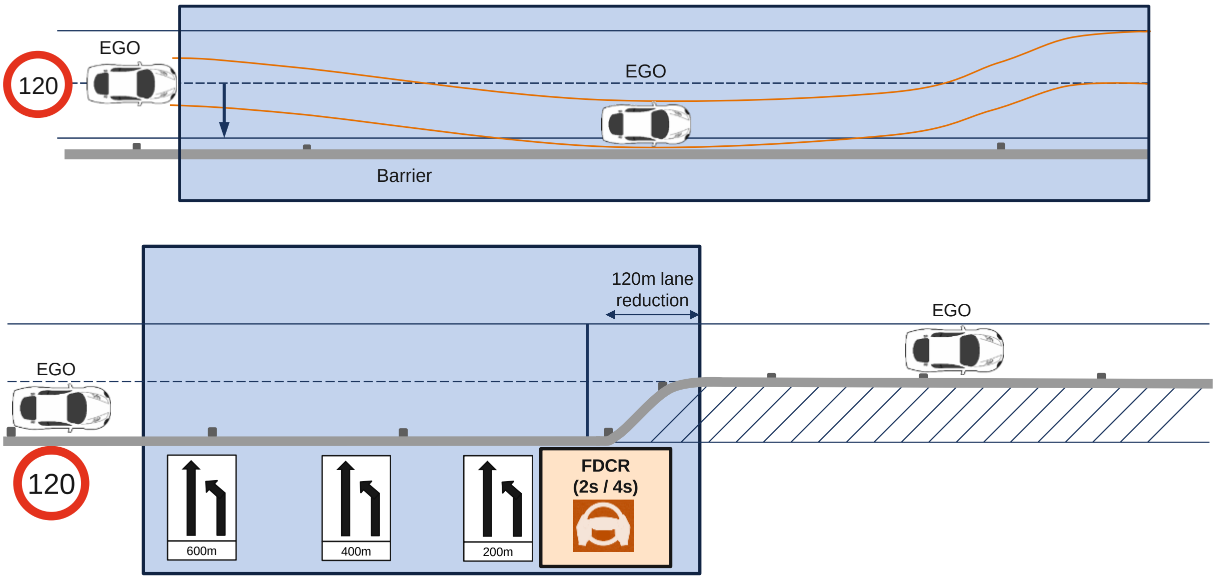


Figure 5-4: Top: Schematic overview of Scenario 1 and 3 which differed in drift velocity. The section of interest for analysis begins at drift onset.
Bottom: Schematic overview of Scenario 2 and 4 (not to scale). The section of interest for analysis of the lane end scenario begins at the first road sign.

The order of the lane drift scenarios was not balanced as the more time-critical scenario was considered to be of higher interest and thus was always presented first to exclude practice or expectation effects. The drift scenario was adapted from Schneider et al. (2022). The drift was designed in a continuous, jerk-free manner to enable a valid investigation of handling performance in this scenario, even in a static driving simulator. The time remaining between drift onset and crossing into the hard shoulder was 3.51 s in the first scenario (0.6 m/s drift velocity) and 6.3 s in the third scenario (0.2 m/s drift velocity), if drivers did not intervene. The hard shoulder of the highway was designed less wide than usual (referenced to German highways; width = 1 m) to create a subjective urgency for intervention.

A visual-verbal NDRT was offered for five defined intervals of 3.5 minutes each during the drive. The NDRT was offered between 2.7 and 1.5 minutes before each system failure or ODD limit and once outside of any measurement scenarios to measure the effect on gaze behavior with different DMS solutions during normal operation. Verbal answers were counted by the experimenter, but not checked for correctness. The task implementation as well as instruction were the same as in Study 3 (Section 5.4.2.8). Participants were asked to read texts (about 60 words) which were presented for 30 seconds on the center information display (CID). Subsequently, they were presented with a question about the text they read. Participants had 10 seconds to read and answer the question verbally. The time available for reading (30 s) and answering (10 s) was limited to partially control potential levels of distraction throughout the drive.

5.5.1.3 Metrics

Metrics used for data evaluation in Study 4 were aligned with the metrics of other studies within the project. An overview can be found in Table 5-1. An overview on non-standardized scales used in the study can be found in the Appendix to this chapter.

For the lane drift (Scenario 1 and 3), driver behavior was referenced to the onset of the function failure and the distance to the lane marking / barrier. For the lane end (ODD limit; Scenario 2 and 4), driver behavior was referenced to the first road sign indicating the ODD limit or the FDCR respectively the lane end. A small bonus (e.g., 5€ in addition to 35€ for German samples) was promised to participants if they performed well in combination of the NDRT and driving task. The bonus was paid in addition to the general financial compensation to all participants regardless of actual performance.

General procedure (e.g., the number and order of questionnaires), the familiarization drive, instructions before the drive as well as manuals providing information on functional limits, HMI and DMS / admissible driver behavior during use were synchronized with prior studies (Study 3 and Study 4, see Section 5.3 and Section 5.4)). The familiarization drive included a section of manual driving, L1 and L2 function activations, driver-initiated deactivations as well as an NDRT practice trial. An interview was conducted with participants after the drive to gather information on situation understanding and the choice of action. Participants watched a short, prototypical video of each of the situations (lane drift and land end). Interviews were only conducted if participants remembered to have encountered such a situation at least once during their drive based on the video shown after the test drive.

Table 5-1: Overview on subjective and objective metrics used in Study 4. Where feasible, the attribution to the project's five challenges and questions (CQ) is indicated.

Subjective assessment
• Attitude towards technology (ATI; Franke, Attig, & Wessel, 2019)
• Acceptance (CTAM, Osswald et al., 2012; subscales: effort expectancy, attitude towards using technology, behavioral intention to use, perceived safety; CQ3 / CQ5)
• Trust (TiA, Körber, 2019; subscales: understanding/predictability, trust in automation; CQ3 / CQ5)
• Subjective monitoring quality (7-point scale, 1 item; CQ1)
• Estimated NDRT engagement (6-point scale; adapted by Metz et al., 2014; CQ3)
• DMS evaluation (7-point scale; 8 items; CQ3 / CQ4 / CQ5)
• System understanding (3-point scale; 9 items pre and 12 items post; CQ4)
• Role understanding (3-point scale; 7 items pre and post; CQ4)
• Interview data, e.g., reasoning behind driver actions, awareness of system limits
Visual attention (CQ1)
• Attention ratio [%] on defined areas of interest (on road, HMI / steering wheel, NDRT)
• Long eyes-off road glances [count] (duration > 2s)
• Fixations [s] / [count] during NDRT engagement (on road, NDRT)
Vehicle-related metrics
• Driving time / ADAS usage [s]
• Mode at measurement point (L0 / L1 / L2)
• Hands-on wheel ratio [%] (measured by capacitive steering wheel sensor) (CQ3)
• Hands-on (reaction) time [s] (after hands-free driving; capacitive sensor) (CQ2)
• Intervention time [s] (CQ2) First active input as observed from the following: o Steering (in case of overrule): 5.6° / after FDCR (i.e., no function overrule necessary): difference of at least 1° to angle at TOR o Brake: 10% pedal travel / after FDCR: 1.14% pedal travel o Accelerator: 20% pedal travel o Button press on steering wheel
• Maximum lateral acceleration [m/s²] within 5s after driver intervention (CQ5)
• Anticipation of system limit (deactivation before FDCR [count]) (CQ2)
• Distance to lane end at lane change [m] (CQ2 / CQ5)
• Type of driver-initiated deactivations [count]
• Type of driver intervention after FDCR [count]
• Termination of DMS requests [s] (CQ2) by gaze attribution to road or hold of steering control
• Number and type of DMS requests received [count] (CQ1)
Other metrics
• Sample criteria, e.g., age, gender, prior ADAS experience (CC, ACC, L2)
• Hand posture (8-point scale, see FOT; level of control, hands-on versus hands-free) (CQ1 / CQ2 / CQ5)
• Number of NDRT responses [count] (CQ3)

5.5.1.4 Sample

Overall, N = 85 participants took part in the study of which n = 7 had to be excluded from data analysis due to limited sensitivity of the DMS, technical difficulties with the simulation or the failure to comply with instructions (e.g., no smartphone use during testing). One additional participant did not finish the study due to difficulties with the simulation software. One participant (L2H-on US) did only activate the L2 function once during the drive, i.e., did not re-activate the function after the first scenario with subsequent deactivation. This participant is only included in the subjective analysis, where feasible, and in the analysis of handling success of Scenario 1 (fast lane drift). Only participants with complete data sets and at least 70% valid gaze data points were included in the analysis of gaze data. Deviating case numbers due to missing values for subjective ratings are reported where necessary.

Recruiting in Germany was carried out via the in-house database of fka, considering gender, age and prior experience with ADAS as well as technical affinity. The US sample was recruited by a recruiting agency under consideration of demographic characteristics and prior L2 experience as well as from the participant pool of the US survey (see Section 4.3). US participants were slightly older than German (DE) participants (M_{US} = 40 years, SD_{US} = 10; M_{DE} = 30 years, SD_{DE} = 9). Overall, 33 % of the sample were female, with total numbers ranging between n = 1 (L2H-off US) and n = 8 females (L2H-on DE) per subsample. DE subsamples did not differ in age (M_{both} = 30 years, SD_{Exp} = 11, SD_{Nov} = 8), but slightly more females were included in the novice than expert sample (n = 8/22 versus n = 3/18). Table 5-2, Figure 5-5, and Figure 5-6 provide an overview on prior ADAS and driving experience in the different subsamples.

Table 5-2: Experience with different ADAS systems and driving experience for the four subsamples in this study.

	L2H-on US	**L2H-on DE**	**L2H-off US**	**L2H-off DE**
ACC experience	*n* = 19/19 (100%)	*n* = 8/20 (40%)	*n* = 17/17 (100%)	*n* = 10/20 (50%)
L2H-on experience	*n* = 16/19 (84%)	*n* = 9/20 (45%)	*n* = 4/17 (24%)	*n* = 10/20 (50%)
L2H-off experience	*n* = 3/19* (16%)	*n* = 0/20 (0%)	*n* = 13/17 (76%)	*n* = 0/20 (0%)
Driver license (years)	*M* = 20, *SD* = 8	*M* = 12, *SD* = 9	*M* = 26, *SD* = 13	*M* = 12, *SD* = 10
Distance driven in the last 12 months (median of categories)	15,0001-30,000 miles	5,001-10,000 km	8,001-15,000 miles	5,001-10,000 km

*Note: Driving experience is reported in different units for DE and US samples. *Users reported to use their system only hands-on.*

US participants were required to indicate the according vehicle make and model or function name for verification of prior experience. In addition, verification of L2 experience in difference to L1 experience was established based on filter questions during recruiting based on the support provided by the function during use, e.g., lane keeping, distance to preceding vehicle and set speed. The sample-split by experience, including ACC experience, is evident in the DE subsamples with n = 22 ADAS novices and n = 18 ADAS experienced drivers. Three drivers took part in the study albeit not meeting the minimum timespan in driving experience of at least 5 years set by the recruiting procedure. However, most drivers have regular driving experience with n = 42 drivers driving (almost) daily compared to n = 20 drivers driving several times per month or less, mostly in the DE samples.

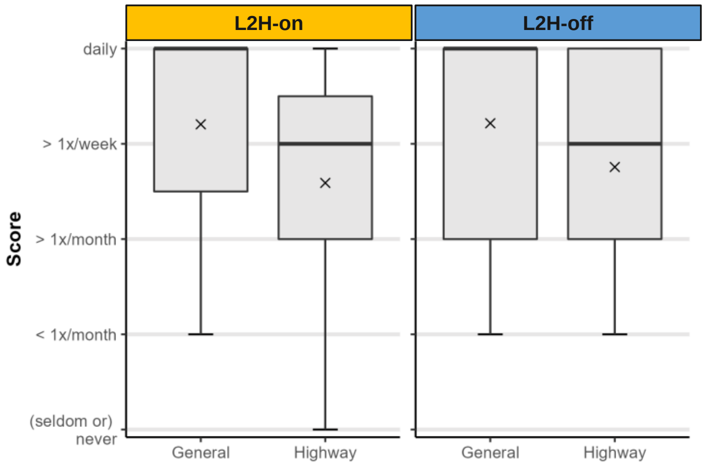


Figure 5-5: Driving experience in the last 12 months of the participant groups in general and on highways ($n_{H\text{-}on}$ = 39, $n_{H\text{-}off}$ = 37), measured on a 5-point categorical scale.

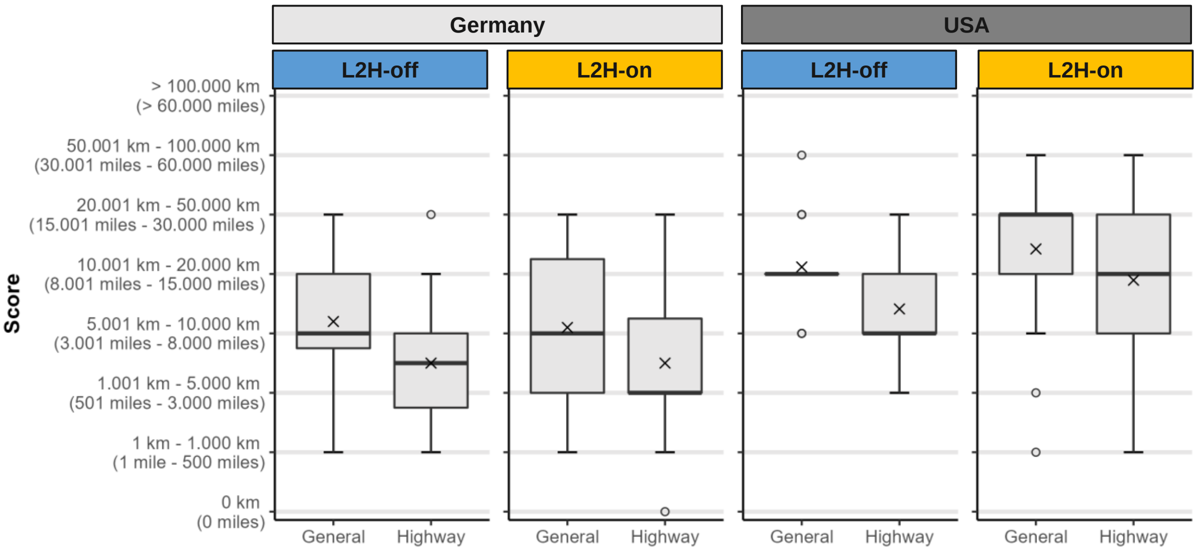


Figure 5-6: Mileage driven within the last 12 months by the participants in the two L2 function groups in the two countries of testing ($n_{L2H\text{-}off\ DE}$ = 20, $n_{L2H\text{-}on\ DE}$ = 20, $n_{L2H\text{-}off\ US}$ = 17, $n_{L2H\text{-}on\ US}$ = 19). The rating categories were kept comparable between countries, but rounded for the US sample (miles instead of kilometers) for comprehensibility.

The two L2 groups did not differ regarding their technical affinity (ATI; Franke et al., 2019; U = 662.5, Z = -0.61, p = .538; $M_{\text{H-on}}$ = 4.74, SD = 0.76, n = 39; $M_{\text{H-off}}$ = 4.84, SD = 0.75, n = 37). However, ADAS experienced drivers showed a higher technical affinity than ADAS inexperienced drivers (U = 280.5, Z = -2.25, p = .024, r = 0.36; M_{Exp} = 5.17, SD = 0.46; M_{Nov} = 4.58, SD = 0.89), despite efforts undertaken in the recruiting process to screen for a positive attitude towards technology in the ADAS inexperienced group.

5.5.2 Results

Data was analyzed using IBM SPSS Statistics 28 for subjective data and R Version 2022.07.0 for objective data. The level of α was set to α = .05. Similar to Study 1 (Section 5.2.2.7), non-parametric tests were used where necessary. Results will be reported with a primary focus on differences between the two different L2 functions under test, i.e., split into an *H-on* group and an *H-off* group. For some metrics, results are additionally reported separately for the German subsample to test for differences between ADAS experienced (*Exp*) and inexperienced (*Nov*) drivers. The stability in findings over different samples (*US* and *DE* subsample) as a side focus of this investigation will be considered primarily descriptively.

5.5.2.1 Acceptance and trust ratings

Acceptance ratings (CTAM, overall mean; Osswald et al., 2012) revealed no difference between L2 functions, $U = 651$, $Z = -0.34$, $p = .733$ ($M_{\text{H-on}} = 5.34$, $SD = 0.96$, $n = 39$; $M_{\text{H-off}} = 5.47$, $SD = 0.87$, $n = 35$; 7-point scale), with equally high ratings not only for both functions but in both countries of testing. However, acceptance differed significantly between ADAS experienced and inexperienced users in the DE subsample, $U = 260$, $Z = -2.07$, $p = .038$, $r = 0.33$, with slightly higher ratings for the experienced subsample ($M_{\text{Exp}} = 5.74$, $SD = 0.63$; $M_{\text{Nov}} = 5.15$, $SD = 0.87$).

Trust in automation after the test drive (TiA, mean of subscales *trust in automation* and *understanding/predictability*; Körber, 2019) was significantly higher for the L2H-off group, $U = 484$, $Z = -2.32$, $p = .020$, $r = 0.27$ ($M_{\text{H-on}} = 3.41$, $SD = 0.65$, $n = 39$; $M_{\text{H-off}} = 3.74$, $SD = 0.43$, $n = 36$; 5-point scale). The absolute level of trust was however not considered indicative of overtrust. Trust did neither differ descriptively between countries ($M_{\text{US}} = 3.47$, $SD = 0.64$, $n = 35$; $M_{\text{DE}} = 3.66$, $SD = 0.51$, $n = 40$) nor significantly between the German experience groups, $U = 249.5$, $Z = -1.42$, $p = .156$.

5.5.2.2 System and driver role understanding

The percentage of correct answers to function-related questions was rather high in all groups, before as well as after experiencing the function with mean values above 80% correct ratings (9 items pre-drive, 12 items post drive; “unsure” was rated as incorrect, see Appendix) and mean correctness per item always above 60%. Interestingly, average agreement to the necessity to keep at least one hand on the steering wheel after activation of the L2 system decreased after use in the L2H-on group by 18%. The highest change in the L2H-off group concerned a higher believe after use (mean increase: 17.5%) that the L2 system will always detect when it cannot handle a situation, despite experiencing a silent lane drift twice in the course of the drive. Experienced ADAS users in the DE sample scored higher before and after the drives regarding a correct system understanding in comparison to novice ADAS users (pre: $M_{\text{Exp}} = 90.20\%$, $SD = 10$; $M_{\text{Nov}} = 88.33\%$, $SD = 10.23$; post: $M_{\text{Exp}} = 91.18\%$, $SD = 9.68$; $M_{\text{Nov}} = 84.58\%$, $SD = 13.76$).

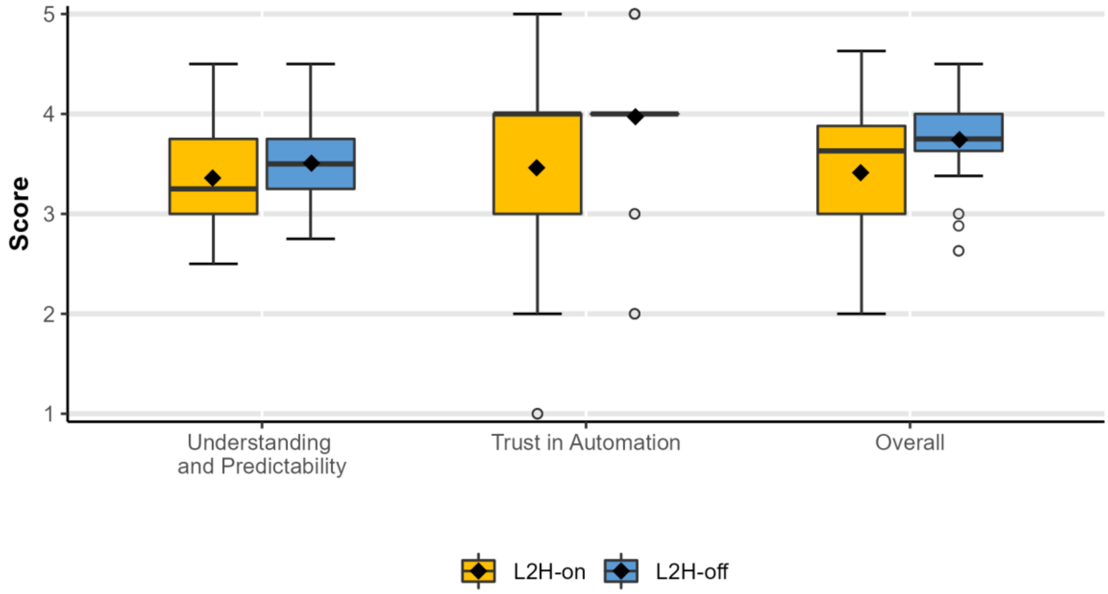


Figure 5-7: Trust in automation (TiA, Körber, 2019; $n_{H\text{-}off}$ = 36, $n_{H\text{-}on}$ = 39) with separate scores for the two TiA subscales used and an aggregated score (1 = *stimme gar nicht zu / do not agree*; 5 = *stimme voll zu / completely agree*)

Likewise, role understanding was similarly high with mean correctness above 80% in all groups both before and after use of the function. The largest change (18%) in the L2H-on group resulted from a higher agreement after use to the statement that the seating position should be the same during L2 use as in manual driving. The largest change (9%) in the L2H-off group resulted from a higher agreement after use to the statement that the same amount of attention needs to be attributed to the road in L2 and manual driving. Experienced ADAS users in the DE sample scored higher before and after the drives regarding a correct role understanding in comparison to novice ADAS users (pre: M_{Exp} = 91.43%, SD = 14.57; M_{Nov} = 78.57%, SD = 15.97; post: M_{Exp} = 95.24%, SD = 8.52; M_{Nov} = 85.71%, SD = 13.22).

Regarding the frequency in different NDRT that drivers would engage in while using the respective L2 function in their daily life (never to very frequently; 6-point scale), the highest frequency of potential interactions in both L2 groups was frequently *talking with a fixated mobile device*, i.e. a primarily auditory activity. Other frequent activities were the *interaction with passengers* and *vehicle related inputs*, both activities that can presumably be aligned well with the driving task. Differences between L2 groups were rather small, albeit slightly higher tendencies in the L2H-off group towards primarily visual-motoric NDRTs.

5.5.2.3 Usage time and hands-off percentage during use

Apart from the aforementioned participant failing to activate the L2 function continuously (total driving time in L2: 7 minutes), driving time in L2 mode ranged between 26 and 39 minutes per participant. Total driving time in manual mode (L0), being required only until activating the ACC at the beginning of the drive, ranged between 8 seconds and 9 minutes. L1 use (ACC mode), used alternatively to L0 after the very beginning of the drive, was required after system limits and especially during the road works scenario and ranged between 3 to 10 minutes per participant. Due to the differences in time of use, comparisons to manual driving will not be reported. Comparisons between ACC and L2 use need to consider potential effects of different times of use.

A similar hands-off percentage was found between experimental groups for L1 episodes, $U = 632$, $Z = -0.75$, $p = .452$. For L2 use, a significantly higher mean percentage of hands-off wheel was observed in the L2H-off group (85.23% vs. 23.77% of L2 usage time), $U = 44$, $Z = -6.98$, $p < .001$, $r = 0.81$, indicating that L2 users adapt their hand posture to the requirements of the function in use, either by instruction or by DMS request. Not all participants of the L2H-off group were using the function hands-free at the beginning of the measurement situation. Validation via video data of the steering wheel based detection data for $n = 33$ participants (due to availability of video data) indicated hands-free driving for over 50% of the L2H-off sample at drift onset / first road sign and – for the respective subsample – at FDCR. In line with the observed hands-off proportions, the L2H-off group reported a higher inclination to drive hands-free when using the respective function (7-point scale; $M_{\text{H-on}} = 3.34$, $SD = 2.04$, $n = 39$; $M_{\text{H off}} = 4.56$, $SD = 1.58$, $n = 36$).

5.5.2.4 NDRT engagement and visual attention

Participants were asked to complete a visual-verbal secondary task at defined intervals during the drive to invoke use cases for effects of different types of DMS designs. NDRT engagement, measured by the completeness of tasks (100% = 26 tasks), did not differ descriptively between groups with 74.73% solved tasks ($SD = 21.9$) in the L2H-off group and 78.44% solved tasks in the L2H-on group ($SD = 15.19$). Correctness of solutions was not recorded or analyzed. The analysis of fixations during one NDRT interval of 3.5 minutes revealed no significant differences in average duration on AOIs (NDRT screen and on road) between functions, $F(1,66) = 2.79$, $p = .100$, albeit descriptively longer fixations in the L2H-on group ($M_{\text{H-on}} = 0.61$ s, $SD = 0.24$, $n = 32$; $M_{\text{H-off}} = 0.53$ s, $SD = 0.27$, $n = 36$). The difference between AOIs was significant, $F(1,66) = 26.61$, $p < .001$, $\eta_p^2 = .13$, with longer average fixations towards the NDRT ($M_{\text{NDRT}} = 0.66$ s, $SD = 0.3$; $M_{\text{on road}} = 0.48$ s, $SD = 0.16$). The interaction between function and AOI was not significant, $F(1,66) = 1.44$, $p = .234$. The descriptive comparison of fixation counts revealed a higher number of fixations in the L2H-off group.

Figure 5-8 depicts the attribution of gaze to relevant areas of interest. For comparison, attention towards the road during L1 episodes was about 98% for L2H-off groups respectively 97% for the L2H-on groups. Note, however, that L1 usage was short and no NDRT was offered during L1 episodes. Analyzed over the entire time of L2 use, no difference in eyes-on road ratios was observed between L2 groups, $t(65.38) = -0.11$, $p = .91$. No differences were observed in regard to eyes-on instrument cluster or steering wheel, $U = 507$, $Z = -1.19$, $p = .234$, as well as eyes-on NDRT, $U = 756$, $Z = -1.74$, $p = .081$, albeit a tendency of more gaze attribution to the NDRT in the L2H-on group. Similar to the eyes-on road ratio, the mean number of prolonged eyes-off road events (gaze off AOI > 2 sec) over the complete L2 usage time was similarly high in both L2 groups, $U = 482.5$, $Z = -1.52$, $p = .128$ ($M_{\text{H-on}} = 138.23$, $SD = 35.77$, $n = 34$; $M_{\text{H-off}} = 150.08$, $SD = 50.96$, $n = 36$).

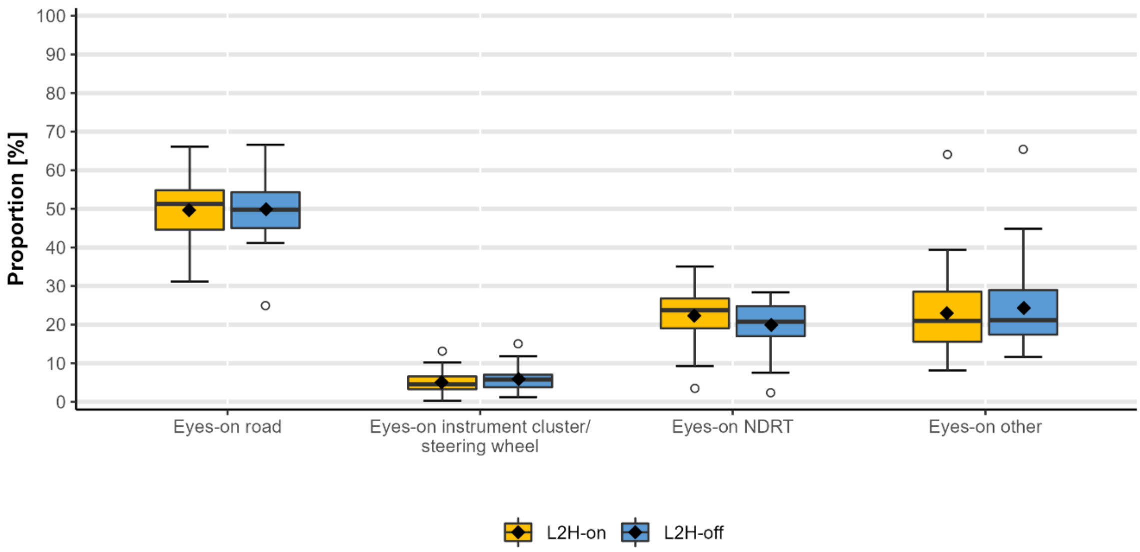


Figure 5-8: Gaze proportions during L2 use in comparison between L2H-on ($n = 32$) and L2H-off ($n = 38$) users for three different AOIs of interest (road, instrument cluster/steering wheel, NDRT screen) and elsewhere (other).

Novice and experienced ADAS users did not differ in their attention towards the road, $t(34.20) = 1.51$, $p = .140$ (see Figure 5-9). Furthermore, no differences were found in the attention to other areas of interest (instrument cluster/steering wheel: $U = 178$, $Z = -0.29$, $p = .770$; NDRT: $t(36.30) = 0.64$, $p = .528$). In comparison, the US sample, in both L2 groups, spend descriptively less attention to the road with about 5% difference in gaze proportions to the DE sample.

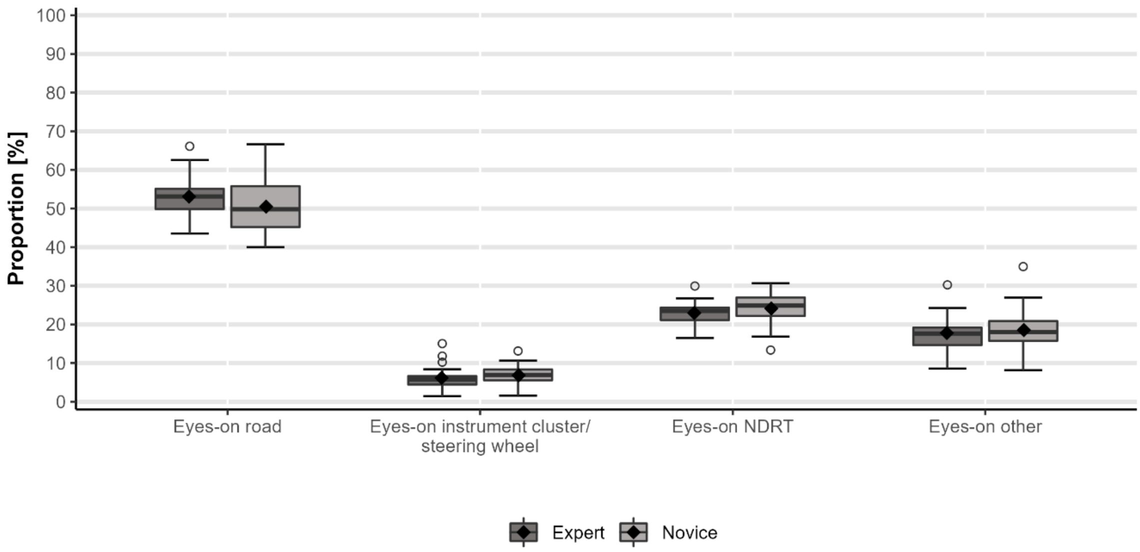


Figure 5-9: Gaze proportions during L2 use in comparison between ADAS experienced (Expert; $n = 19$) and ADAS inexperienced (Novice; $n = 20$) users for three different AOIs of interest (road, instrument cluster/steering wheel, NDRT screen) and elsewhere (other).

Subjective involvement in the driving task (7-point scale) during use was rated with median ratings of 5 except for the L2H-off DE sample (i.e., *Mdn* = 3; $M_{\text{H-on}}$ = 4.57, *SD* = 1.48, *n* = 39; $M_{\text{H-off}}$ = 4.20, *SD* = 1.39, *n* = 36). The subjective level of attention when supervising the function's behavior and the surrounding traffic was overall high ($M_{\text{H-on}}$ = 5.22, *SD* = 1.33, *n* = 39; $M_{\text{H-off}}$ = 5.03, *SD* = 1.38, *n* = 36) as was the reported general attentiveness during the drive with median ratings of 6 or more in all groups except the L2H-off DE sample (i.e., *Mdn* = 5; $M_{\text{H-on}}$ = 5.59, *SD* = 1.15, *n* = 39; $M_{\text{H-off}}$ = 5.21, *SD* = 1.26, *n* = 36). Experienced ADAS users reported a higher involvement (mean difference: 0.46) and higher attentiveness to supervision (mean difference: 0.71) than novice users.

Not unexpectedly, given alone the longer duration until a first request would be issued, the L2H-off group received a greater number of stage 1 requests from the DMS than the L2H-on group ($n_{L2H\text{-}off}$ = 142, single value $max_{L2H\text{-}off}$ = 52; $n_{L2H\text{-}on}$ = 78, single value $max_{L2H\text{-}on}$ = 18). This effect was also visible in the additional stages of DMS requests (Stage 2: $n_{L2H\text{-}off}$ = 93, $n_{L2H\text{-}on}$ = 29; Stage 3: $n_{L2H\text{-}off}$ = 13, $n_{L2H\text{-}on}$ = 4), with 34.51% of stage 1 EOR (eyes-on requests) and 62.82% of stage 1 HOR (hands-on requests) being terminated before the second request stage. This effect is contrasted by the significantly longer (but overall short) time taken to terminate requests (measured from onset of the first DMS alert) in the L2H-on group, (U = 725, Z = -3.15, p = .002, r = 0.4; $M_{\text{H-on}}$ = 1.89, *SD* = 0.9, *n* = 31 with at least one request received; $M_{\text{H-off}}$ = 1.32, *SD* = 0.92, *n* = 32 with at least one request received), a finding that is in line with the longer intervals implemented between DMS stages and the different response types needed to terminate requests. This effect might, however, also be influenced by a different sensitivity of the DMS systems, e.g. capacitive steering wheel and camera-based gaze tracking. Means of maximum reaction times to DMS requests did not differ between L2 functions, *U* = 431, *Z* = -0.88, *p* = .378 ($M_{\text{H-on}}$ = 3.6, *SD* = 3.14; $M_{\text{H-off}}$ = 5.01, *SD* = 4.27).

ADAS novices received a higher number of requests compared to ADAS experienced drivers (Stage 1: n_{Nov} = 74, n_{Exp} = 46; Stage 2: n_{Nov} = 44, n_{Exp} = 25; Stage 3: n_{Nov} = 4, n_{Exp} = 4), but a similar percentage of all requests (n_{Nov} = 40.54%; n_{Exp} = 45.65%) was terminated after the first stage. Experienced and novice users did not differ in the time taken to terminate requests (M_{Exp} = 1.51, *SD* = 0.98; M_{Nov} = 1.41, *SD* = 0.86). In line with objective findings, questions directed at the understandability of reasons for DMS requests as well as the requested action to terminate them revealed a high agreement in all groups (7-point scale; mean ratings above 5.50). Little agreement was stated to willfully ignoring DMS requests (mean ratings of 2.50 or lower). Albeit a higher frequency of requests in the L2H-off group, similarly low agreement to the statements '*I perceived the warnings to be annoying*' is given in all and to '*The warning came too frequently*' in most groups (7-point scale, 7 = high agreement; annoyance: mean ratings between $\text{M}_{\text{L2H-off US}}$ = 2.56, SD = 1.90, and $\text{M}_{\text{L2H-on US}}$ = 3.26, SD = 2.10; frequency: mean ratings between $\text{M}_{\text{L2H-off DE}}$ = 4.10, SD = 1.94, and $\text{M}_{\text{L2H-on US}}$ = 2.81, SD = 1.91). In all groups, the DMS is considered to raise subjective safety during the use of L2 functions (7-point scale, 7 = high agreement; mean ratings between $\text{M}_{\text{L2H-off US}}$ = 6.31, SD = 1.35, and $\text{M}_{\text{L2H-on US}}$ = 5.47, SD = 1.47). Slight agreement is given to the statement that more time would be spend on NDRT without the requests (7-point scale, 7 = high agreement; mean ratings between $\text{M}_{\text{L2H-off DE}}$ = 4.40, SD = 2.09, and $\text{M}_{\text{L2H-on US}}$ = 3.68, SD = 2.08).

5.5.2.5 Evaluation of interaction behavior at system limits

The interaction behavior at function limits will be reported separately for the silent failure (Scenario 1 and 3), presenting an unforeseen intervention need, and the lane end, presenting a foreseeable ODD limit (Scenario 2 and 4). For both types of scenario, the first driver intervention observed, regardless of the L2 function, was steering (see Table 5-3).

5.5.2.6 Driver interventions in the lane drift scenario

The silent failure was overruled in time by all but one participant (L2H-off DE, ADAS experienced), who did not intervene in time in the first, fast drift scenario and was excluded from further analyses for this scenario (see Table 5-3). The non-successful intervention in this case seems to be caused by a combination of little visual attention attributed to the road scene during the occurrence of the lane drift (average fixations to the road of about 66 ms) and a misunderstanding of the required driver actions in case of DMS requests. The L2H-on experienced participant, using a L2H-off function in the study, received an EOR shortly before the onset of the lane drift, but reacted to this with a small movement of the steering wheel as if to terminate an HOR. Gaze was attributed to the road only as a by-product of the hand movement for about 90 ms, but enough to terminate the EOR as implemented in this study. Only a second EOR, issued after contact with the barrier, lead to a discovery of the collision by the participant, who afterwards (in the interview) did not attribute this to the DMS, but to a “random discovery”.

All other participants overruled the lane drift in time to avoid contact with the barrier, implemented to provide a means for intervention necessity. However, less than half of the sample intervened before crossing the lane marking (point of reference: right vehicle boundary; see Table 5-3). A higher number of participants intervened before crossing the lane marking to the right in the second lane drift. Of the 49 participants who remembered both lane drift scenarios in the post-drive interview, n = 16 experienced the first drift scenario (n_{2nd} = 5) as more time critical. Overall, n = 18 indicated that visual cues such as the distance to the barrier or the lane markings were their main indicator to realize that the system was not behaving correctly. The majority of drivers indicated to have intervened immediately after realizing the misalignment within the lane, whereas n = 9 stated to have realized the drift, but decided to wait and see if the function could handle the situation itself. Of the ADAS inexperienced drivers, n = 2 reported to be confused by the function behavior, as the HMI signal suggested normal control. One L2H-off user reported to have placed the hands on the steering wheel when interacting with the NDRT after encountering the first drift scenario as a personal lesson learned.

Table 5-3: Type of driver interventions and intervention outcome. Upper table: Type of driver interventions and success in the lane drift scenario (silent failure). Lower table: Type of driver interventions and success in the lane end scenario representing an ODD limit.

	Lane Drift (0.6 m/s; Scenario 1)				**Lane drift (0.2 m/s; Scenario 3)**			
Type	**L2H-on (US)**	**L2H-on (DE)**	**L2H-off (US)**	**L2H-off (DE)**	**L2H-on (US)**	**L2H-on (DE)**	**L2H-off (US)**	**L2H-off (DE)**
Steer	19/19	20/20	17/17	19/20	18/18	20/20	17/17	20/20
Brake / Button	-	-	-	-	-	-	-	-
First reaction after collision	-	-	-	1/20	-	-	-	-
Maintains Lane	3/19	11/20	4/17	9/20	9/18	12/20	7/17	9/20
	First lane end (Scenario 2)				**Second lane end (Scenario 4)**			
Type	**L2H-on (US)**	**L2H-on (DE)**	**L2H-off (US)**	**L2H-off (DE)**	**L2H-on (US)**	**L2H-on (DE)**	**L2H-off (US)**	**L2H-off (DE)**
Steer	15/18	17/20	15/17	17/20	16/18	20/20	17/17	18/20
Brake	3/18	1/20	2/17	1/20	2/18	-	-	-
Button	-	1/20	-	2/20	-	-	-	2/20
First reaction after collision	-	1/20	-	-	-	-	-	-

Hands-on time, measured from drift onset for the L2H-off group, was faster for the first, faster drift than for the second, slower drift, $t(16) = -4.12$, $p < .001$, $d = 3.24$ (only participants driving hands-free, additionally excluding $n = 6$ participants with hands-on supervision of the function at onset of the lane drift in only one of the two encounters; $M_{first} = 4.59$, $SD = 1.10$; $M_{\text{second}} = 5.95$, $SD = 1.80$; $n = 17$). This difference in timing might be due to the slower drift velocity in the second scenario, leaving more time for participants to react to the lane drift, but also resulting in a later visible change in lane position. A comparison of ADAS experienced and inexperienced drivers in the DE subsample revealed earlier hands-on times for experienced drivers, although sample size for this comparison was restricted to those driving hands-free at measurement onset (1st lane drift: $M_{Exp} = 3.85$, $SD = 0.51$, $n_{\text{Exp}} = 5$; $M_{\text{Nov}} = 5.01$, $SD = 0.98$, $n_{\text{Nov}} = 7$; 2nd lane drift: $M_{Exp} = 5.45$, $SD = 1.49$, $n_{\text{Exp}} = 4$; $M_{\text{Nov}} = 7.41$, $SD = 1.13$; $n_{\text{Nov}} = 6$).

In line with hands-on times, interventions, mostly by steering (see Table 5-3), occurred significantly later in the second, slower drift scenario, $F(1,71) = 546.36$, $p < .001$, $\eta_p^2 = .71$ ($M_{\text{first}} = 4.37$, $SD = 0.81$; $M_{\text{second}} = 7.97$, $SD = 1.41$). Neither the difference between L2 functions, $F(1,71) = 0.29$, $p = .532$ ($M_{\text{H-on}} = 6.1$, $SD = 2.15$, $n = 38$; $M_{\text{H-off}} = 6.24$, $SD = 2.14$, $n = 35$), nor the interaction, $F(1,71) = 0.20$, $p = .657$, indicated differences in intervention times. US participants reacted slightly later than DE participants (see Figure 5-10). ADAS experience did not influence intervention times in the lane drift scenario, $U = 654$, $Z = -0.94$, $p = .348$ ($M_{\text{Exp}} = 5.52$, $SD = 0.69$, $n = 17$; $M_{\text{Nov}} = 5.95$, $SD = 1.22$, $n = 22$).

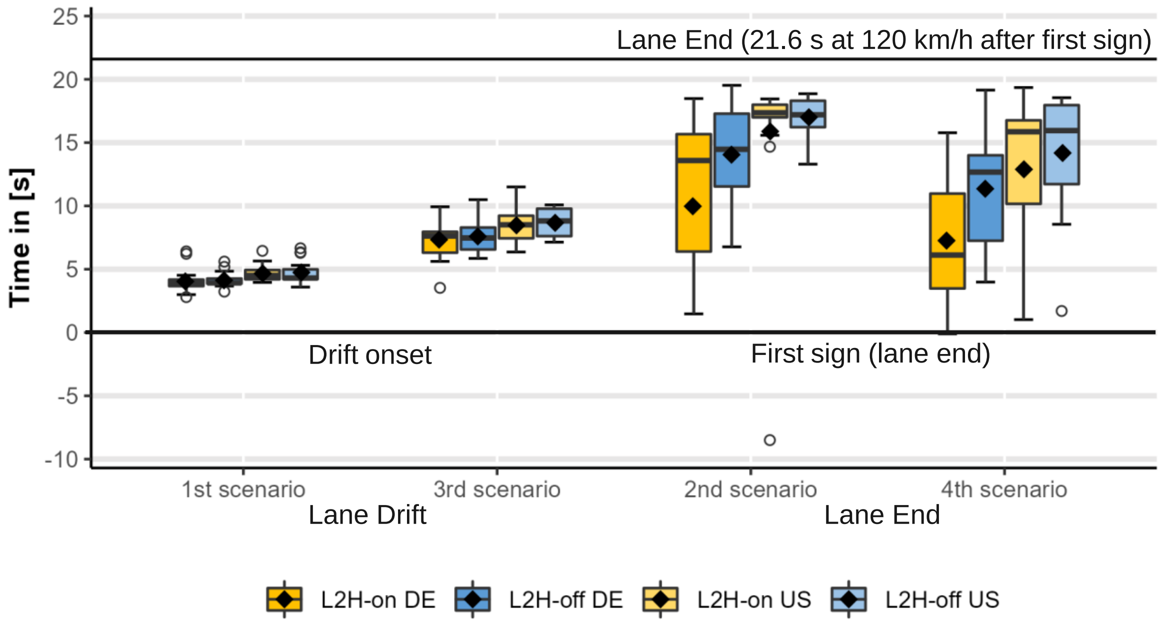


Figure 5-10: Intervention times in the four scenarios (1st: fast lane drift; 3rd: slow drift; 2nd: first ODD limit; 4th: second ODD limit) for L2H-on (($n_{lane\ drift}$ = 38; $n_{lane\ end}$ = 37) and L2H-off groups ($n_{lane\ drift}$ = 35; $n_{Lane\ end}$ = 37), each split into DE and US samples.

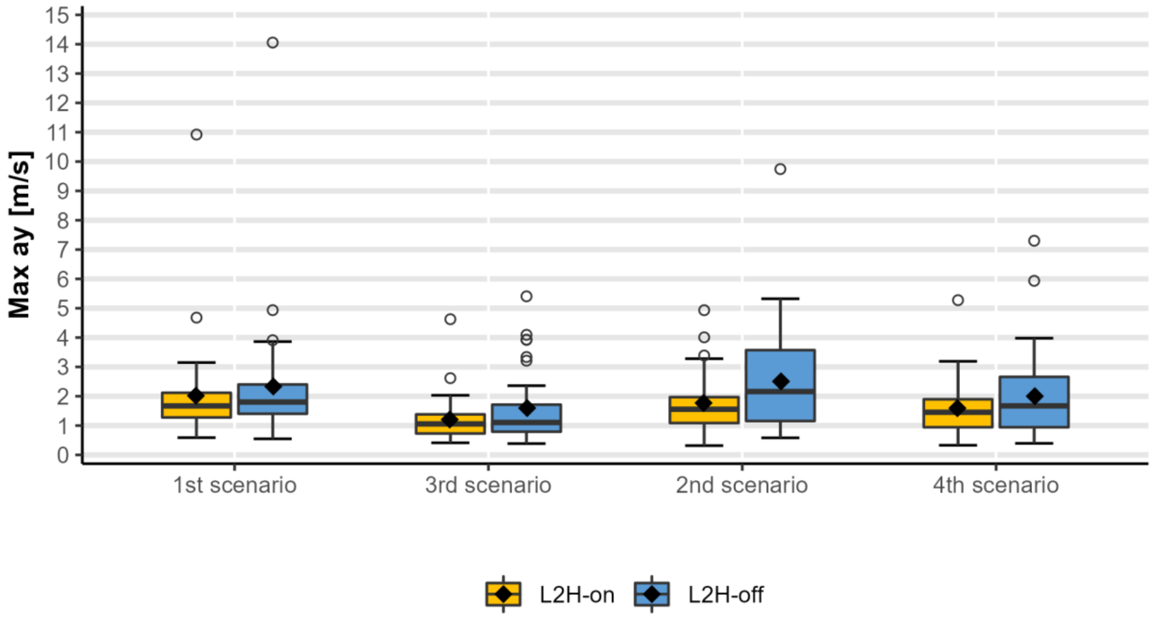


Figure 5-11: Maximum lateral accelerations in four scenarios (1st: fast lane drift; 3rd: slow drift; 2nd: first ODD limit; 4th: second ODD limit) for L2H-on ($n_{lane\ drift}$ = 38; $n_{lane\ end}$ = 37) and L2H-off groups ($n_{lane\ drift}$ = 35; $n_{lane\ end}$ = 37), each split into DE and US samples.

Maximum lateral accelerations within 5 s of the first driver intervention (overrule), analyzed only for relative comparison between groups due to the use of a static driving simulator, revealed a pattern in line with intervention times for the lane drift scenarios, i.e., higher lateral accelerations for faster mean interventions, but no difference between functions (see Figure 5-11). The prominent outliers in the first (lane drift) and second (lane end) scenario were re-evaluated by video analysis and can be explained with intense NDRT engagements, leading to an unexpected intervention need. Experience was not considered for this metric, as ADAS experience should not alter general handling abilities, but rather the timing and type of interventions, which did not differ. Longitudinal accelerations were not in the focus of analysis, as ACC remained in longitudinal control after FDCR.

5.5.2.7 Driver interventions in the lane end scenario

The lane change in face of the lane end, presenting an ODD limit of the function, was successfully handled by all but one participant (L2H-on DE, ADAS novice; bounding box crosses barrier), who was subsequently excluded from further analysis of this scenario. The non-successful intervention in this case was not caused by a misunderstanding of system limits, as the participant stated in the post-drive interview to have been aware that the system cannot change lanes on its own. Rather, a misunderstanding of the difference between DMS requests, requiring merely hold of steering control, and the FDCR at ODD limits, requiring in this case an immediate intervention, seems to have caused the late intervention. As illustrated by Figure 5-12, the participant realized the FDCR, but, instead of changing lanes, moved a second hand to the steering wheel while fixating the HMI until the FDCR terminates. After the system-initiated mode change to ACC, the gaze was re-directed to the road, followed by an immediate intervention. Before the FDCR, gaze was throughout directed at the NDRT, except for one quick glance towards the road. The participant thus missed all indications of the upcoming ODD limit.

Figure 5-12: Gaze direction (red circle) at three distinct time points during the non-successful handling of the ODD limit (lane end). A: Function direct control request (FDCR) at 4 s before lane end. B: FDCR end, system mode changes to ACC. C: First gaze towards road.

Hold of steering control (i.e., hands-on timing) in the L2H-off group was established faster in the second encounter of the lane end than in the first encounter, $W = 120$, $Z = -2.06$, $p = .039$, $r = -0.48$ (excluding participants with hands-on supervision of the function when passing the first sign announcing the lane end, amongst others $n = 5$ participants with hands-on supervision in only one of the two lane end scenarios; $M_{\text{first}} = 7.16$, $SD = 12.18$; $M_{\text{second}} = 1.78$, $SD = 8.03$; $n = 17$). A comparison of ADAS experienced and inexperienced drivers in the L2H-off group in the DE subsample revealed descriptively earlier hands-on times for experienced drivers, but sample size for this comparison was small including only those driving hands-free at measurement onset both times (1st encounter: $M_{Exp} = 5.89$, $SD = 7.82$, $n_{\text{Exp}} = 4$; $M_{\text{Nov}} = 10.36$, $SD = 8.13$, $n_{\text{Nov}} = 5$; 2nd encounter: $M_{Exp} = -0.40$, $SD = 2.02$, $n_{\text{Exp}} = 4$; $M_{\text{Nov}} = 5.19$, $SD = 8.32$, $n_{\text{Nov}} = 5$).

In line with hands-on reaction times, interventions, mostly by steering (see Table 5-3), occurred significantly later in the second encounter (4th scenario), $W = 2397$, $Z = -5.44$, $p < .001$, $r = 0.63$

(M_{first} = 14.13, *SD* = 6.90; M_{second} = 11.33, *SD* = 5.68). In contrast to the lane drift scenario, L2H-on users intervened slightly (by tendency, not significantly) earlier than L2H-off users, *U* = 2273.50, *Z* = -1.78, *p* = .07 ($M_{H\text{-}on}$ = 11.42, *SD* = 7.89, *n* = 37; $M_{H\text{-}off}$ = 14.03, *SD* = 4.27, *n* = 37). The variance in intervention times was larger in the less time-critical lane end scenario compared to the lane drift. As can be seen on the right hand side of Figure 5-10, this tendency is mostly due to differences between L2 functions in the second encounter (4th scenario) in the German subsample. However, the difference in intervention times did not have any implications regarding the quality or safety of the intervention, as almost 10 s driving time until lane end remained on average in both groups at the moment of intervention. Furthermore, most drivers reacted even earlier than in the first encounter and no incident arises out of late interventions, rendering any difference in intervention time meaningless. In the L2H-on group, one driver overrules the function especially early in the first encounter, even before passing the first road sign. As can be seen in Figure 5-10, US participants invented later than German participants in both lane end scenarios. ADAS experience did influence intervention times in the lane end scenario only by tendency, *F(1,37)* = 3.52, *p* = .068, with earlier interventions in the group of experienced drivers (M_{Exp} = 9.89, *SD* = 5.20, *n* = 18; M_{Nov} = 12.29, *SD* = 4.74, *n* = 22).

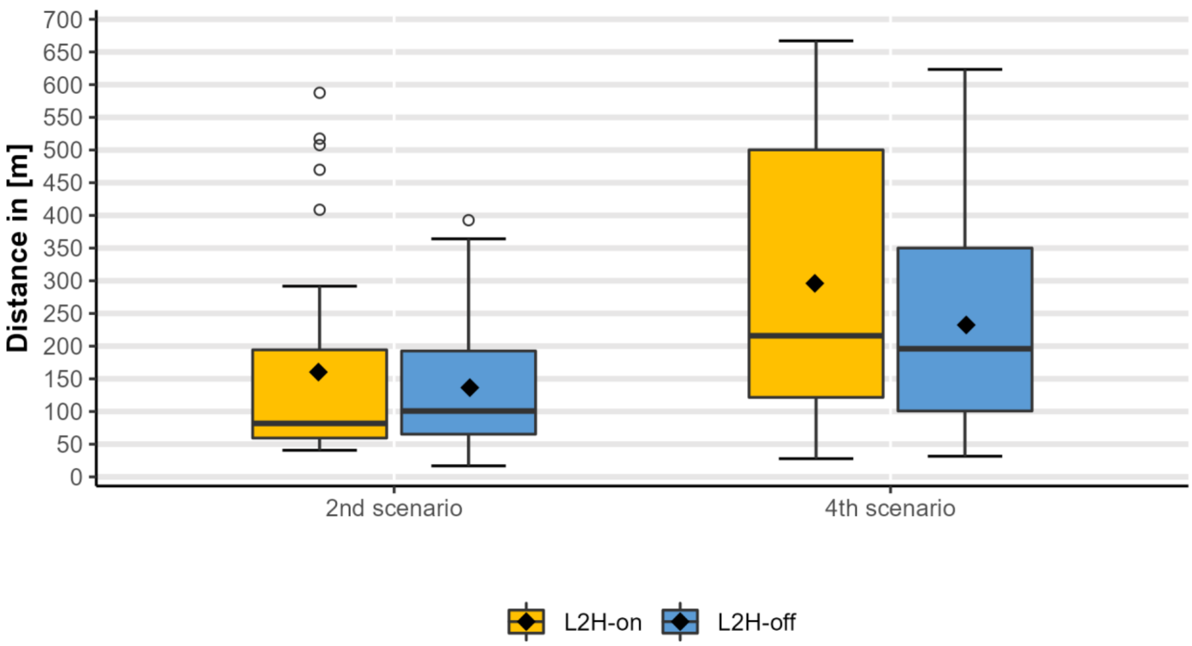


Figure 5-13: Distance to the ODD limit (in m) where the center of the vehicle's bounding box crossed the marking between lanes for the first (2nd scenario) and second encounter (4th scenario), split into the L2H-on (*n* = 37) and L2H-off group (*n* = 37).

In addition to the timing of first interventions, the distance to the lane end at lane change (Figure 5-13) was analyzed, providing insights into the handling of the situation. In line with first direct driver input, no difference between functions was found, *F*(1,72) = 1.74, *p* = .191 ($M_{H\text{-}on}$ = 228.19 m, *SD* = 127.68, *n* = 37; $M_{H\text{-}off}$ = 184.46 m, *SD* = 139.98 m, *n* = 37). Lane changes were conducted significantly earlier in the second encounter (4th scenario) with knowledge on function behavior and scenario design, *F*(1,72) = 54.29, *p* < .001, η_p^2 = .12

(M_{first} = 148.42 m, *SD* = 191.21 m; M_{second} = 264.23 m, *SD* = 184.46 m). The interaction was not significant, $F(1,72) = 1.60$, $p = .209$.

Albeit an early visibility of the lane end (first sign at 600 m before ODD limit), some participants (Scenario 2: $n_{H\text{-}on}$ = 14; $n_{H\text{-}off}$ = 21; Scenario 4: $n_{H\text{-}on}$ = 7; $n_{H\text{-}off}$ = 11) only reacted shortly before or even after the function issued a request to take direct control (FDCR) at either 2 s or 4 s before lane end (order balanced over sample and scenarios). As can be seen from the reaction times to the request in Scenario 2 (Table 5-4), drivers intervened close to the onset of the DCR signal, speaking for a certain preparedness to intervene despite no anticipative driver-initiated deactivation. Minimum values suggest that some participant may have started steering before the FDCR, but did not overrule the function (see Table 5-1 for detection thresholds applied). Mean differences in intervention times for drivers of the L2H-off group with and without hold of steering control at FDCR onset were less than 200 ms, indicating little influence of hand posture on non-anticipative interventions in this scenario. The latest intervention to the FDCR is observed for a driver with hold of steering control at that time, indicating that hold of steering control is not a sufficient prerequisite for fast interventions. Differences caused by variations in FDCR timing are not analyzed due to the high number of anticipative actions by drivers.

Table 5-4: Reaction times (*M, SD, min, max*) to the function direct control request (FDCR) in the first encounter (Scenario 2) of the ODD limit (lane end) for all participants in both L2 groups who received an FDCR in this scenario.

Group	**No hold of steering control at FDCR**					**Hold of steering control at FDCR**				
	M	*SD*	*Min*	*Max*	*n*	*M*	*SD*	*Min*	*Max*	*N*
L2H-off	0.91	0.53	0.14	1.61	12	0.72	0.94	0.00	2.96	9
L2H-on	0.54	0.52	0.00	1.01	3	0.61	0.55	0.00	1.56	11

For the lane end scenario, higher lateral accelerations can be observed for L2H-off users in both encounters (compared to L2H-on users), with lesser variation in the second encounter (4^{th} scenario; see Figure 5-11). The trajectories of each driver within each of the two lane end scenarios are displayed in Figure 5-14 (L2H-on) and Figure 5-15 (L2H-off). They illustrate the findings discussed above, i.e., mainly an effect of sequence with earlier interventions in the fourth scenarios (second lane end) and slightly more participants with very early interventions in the L2H-on group than in the L2H-off group in the first lane end encounter (second scenario).

The anticipation of the upcoming system limit is further visible in the changes in hand position towards higher steering control ratings (based on video analysis, see procedure described for FOT) in relation to the system deactivation. L2H-off users initially have lesser contact to the steering wheel, but medium to high control at system deactivation. However, some users in the L2H-off group use the hands-free option only scarcely overall.

The slower intervention in the first encounter compared to the second encounter of the ODD limit might be attributed to different reasons. Firstly, participants might have initially expected the function to act on its own, i.e., to perform a lane change, albeit explicit mention in the pre-drive manual of function limits. Secondly, participants might have actively tested the function to see how it behaves at ODD limits. Similarly, participants might have waited, independently of their function knowledge, to perform the necessary lane change as there was no need for an earlier intervention, as discussed for the lane drift.

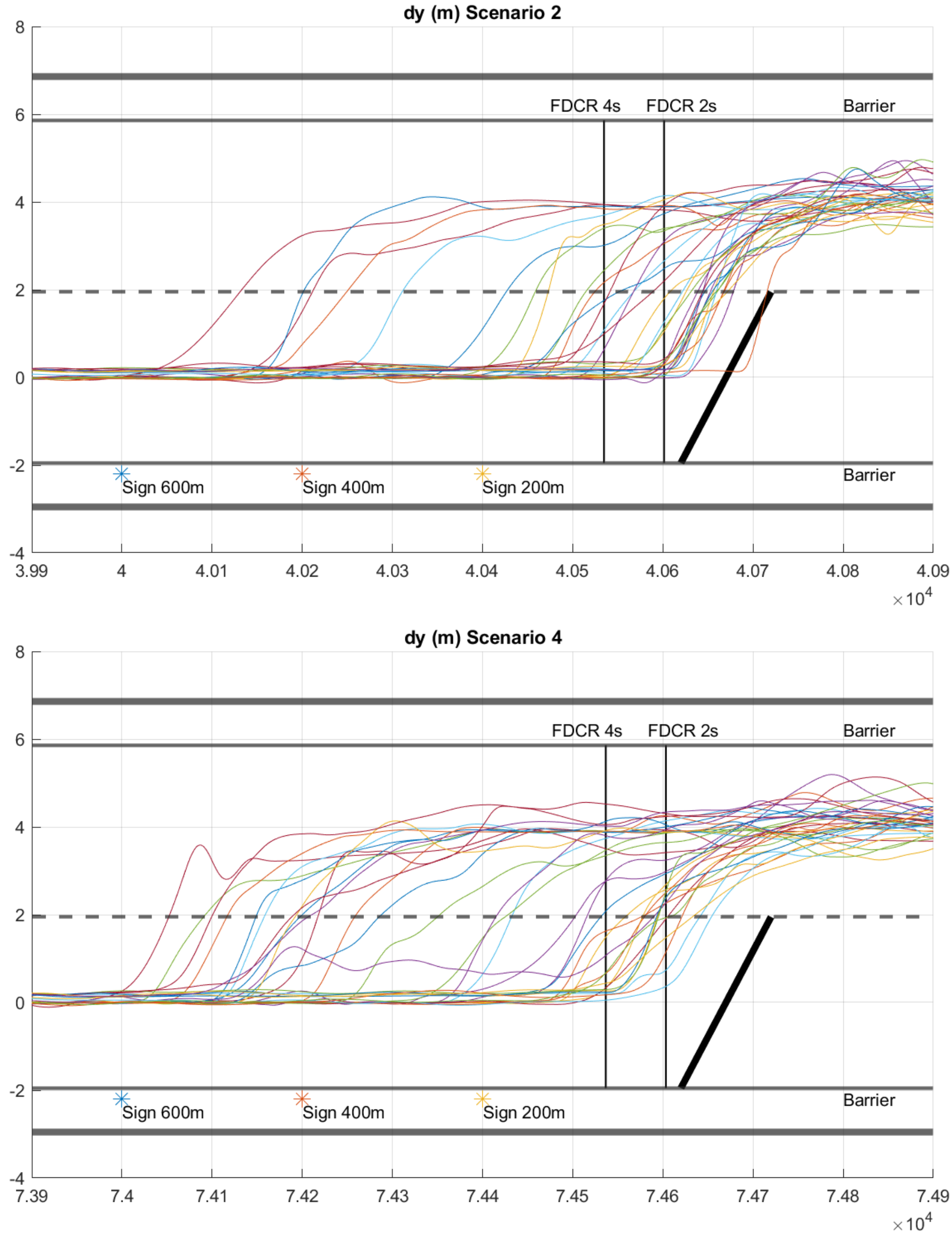


Figure 5-14: Trajectories of each participant in the lane end scenario in the L2H-on group (including one collision in Scenario 2) for the first encounter (Scenario 2; top) and second encounter (Scenario 4; bottom). Barrier not to scale.

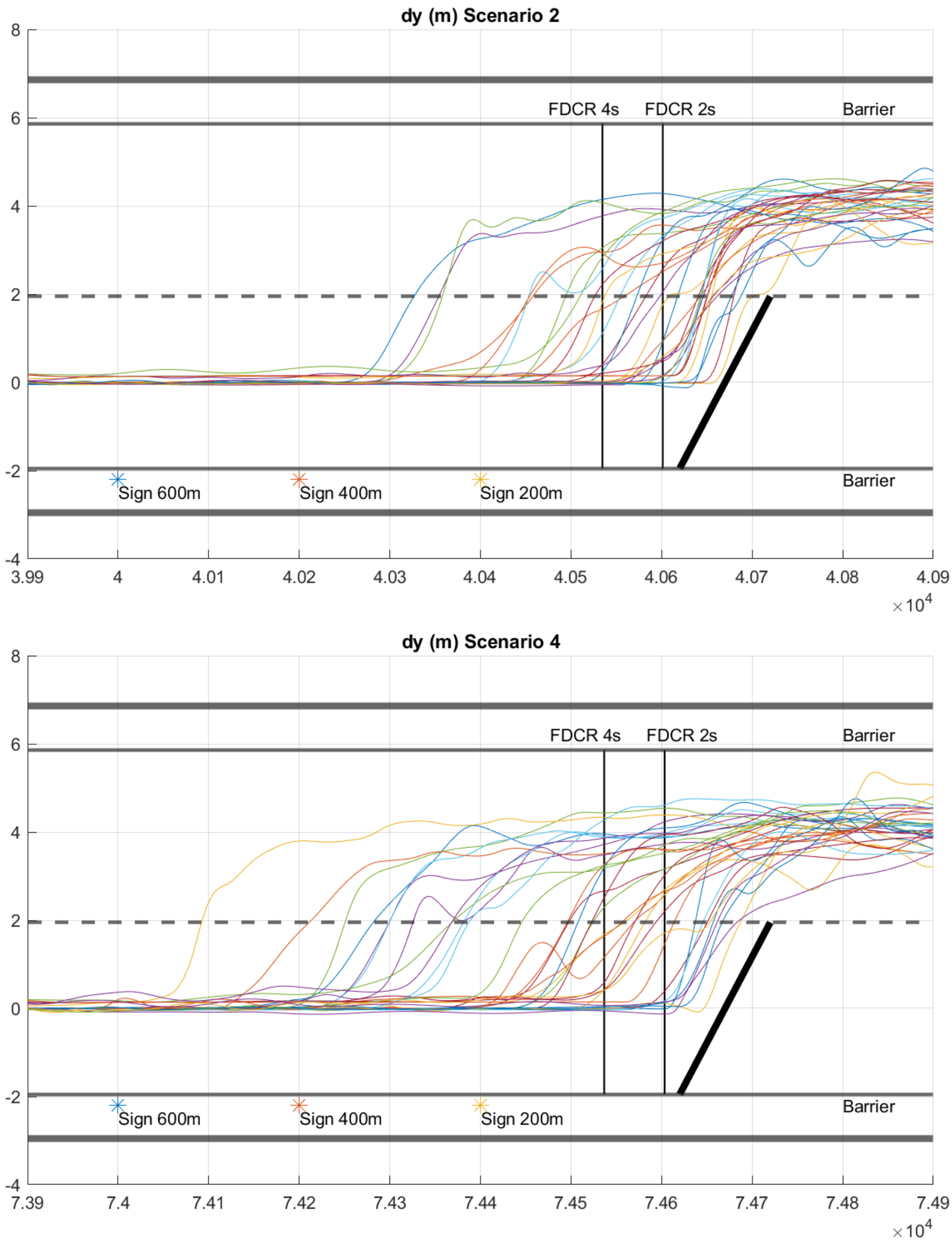


Figure 5-15: Trajectories of each participant in the lane end scenario in the L2H-off group for the first encounter (Scenario 2; top) and second encounter (Scenario 4; bottom). Barrier not to scale.

All three reasons are supported by the analysis of interview data. Firstly, interview data suggest some participants did not expect the function to issue a FDCR at the lane end, albeit clear instructions before the drive regarding function limits and the according functional behavior at such. Some also confused the FDCR with a DCR issued by the DMS, missing the connection between the DCR and the system limit. One participant was even convinced that the function had changed lanes on its own. Secondly, one participant (L2H-off experienced, using a L2H-off function in the study) stated in the post-drive interview:

> *The lane was going to the end and I knew the system was not going to change lanes on its own. In the first situation I waited until the lane was just about to merge […], just my normal reaction.*

In fact, this participant prepared for the ODD limit in time by moving the hands closer to the steering wheel early on and keeping contact to the steering wheel well before intervening. At FDCR, an intervention follows immediately.

5.5.3 Discussion

This fourth simulator study was designed to investigate potential differences in the detection and compensation of steering failures and the handling of ODD limits when interacting with different types of L2 functions (RQ1). Additionally, ratings and interaction behavior were investigated regarding their generalizability over samples, e.g. for different levels of ADAS experience and different cultural backgrounds (RQ2). No systematic disadvantage was observed for interactions with the L2H-off function with an EOD-based DMS in comparison to a L2H-on function with an HOD-based DMS. In the following, this observation will be discussed with regard to the assumed challenges in interaction with L2 functions.

The primary research focus targeted differences in driver interventions between L2 functions for two different types of scenarios (silent failures and ODD limits), of which none could be found. Similarly to the little difference found between L2 function designs for behavior and ratings, prior ADAS experience did not result in systematic changes in the interaction with L2 functions apart from slightly earlier establishment of hold of steering control by experienced drivers. This effect was however without consequences regarding intervention times. Although experienced drivers reported a higher involvement in the driving task during L2 use and a better role and system understanding before as well after use, no objective difference in gaze attribution to relevant AOI was found between experience groups. This lack in effect might be due to the difficult transfer of day-to-day experience into the experimental setting where an unknown function had to be used on an unknown track. Only qualitative results suggest a difference in expectations towards the function, e.g. regarding the two L2 novices who were confused by the lack of change in the HMI, indicating normal operation, during the lane drift. Similar as for ADAS experience, cultural differences did not impact the conclusions derived for L2H-off functions. No systematic effects relating to the difference between L2 functions were found by separate descriptive analysis of the US and DE sample. However, descriptive differences between US and DE samples became apparent, although not in primary focus of this study, e.g. regarding the intervention at ODD limit, where US drivers changed lanes later than DE drivers. Regarding, e.g., the gaze proportion during engagement in NDRT, only minor differences between the US and DE sample was observed in this study.

5.5.3.1 CQ1 (hands-off = mind-off) and CQ3 (Misuse)

Regarding the monitoring behavior analyzed as part of CQ1 (hands-off = mind-off), no differences between the L2 functions were observed in the amount of visual attention attributed to the road, although trust was significantly higher towards the L2H-off function. Thus, higher trust levels did not yield negative effects on monitoring performance in combination with the prototypical DMS. Differences in the gaze pattern when engaging in a visual-verbal NDRT revealed no significant differences between functions, but indicated that an EOD-based DMS interrupts prolonged visual distraction resulting in more and shorter fixations to the NDRT. No difference in the number of questions answered was found between the functions (CQ3: misuse). Furthermore, the EOD-based DMS issued a higher number of requests, indicating the need for the monitoring of visual attention. The EOD-DMS support did however not result in a significant increase in eyes-on road proportions in this study. Overall, eyes-on road ratios were likely reduced by the frequent visual-verbal NDRT offer as well as the low traffic density and little visual complexity of scenarios implemented. The higher frequency of requests was not perceived as annoying by L2H-off users. Overall, DMS were regarded as positive in terms of subjective driving safety.

5.5.3.2 CQ2 (prolonged transition times) and CQ5 (safety level)

No differences in intervention times (CQ2: prolonged transition times) were observed by comparison of the two L2 functions in the lane drift and lane end scenario. Users of the L2H-on function intervened earlier by tendency in approach to the widely visible lane end, especially German users in the second encounter (4^{th} scenario). This behavior did however not translate into earlier lane changes in comparison to the L2H-off group and does not have any safety implications as lane changes were on average still conducted early on in both groups. Consideration of intervention time might in this specific scenario and in difference to the more time-critical scenario in Study 1 and 3 (Section 5.2.3.2.3 and Section 5.4.2.5LINK not be a very insightful indicator apart from the occurrence of an intervention in good time. Overall, intervention times show a high variability depending on the affordances of the respective scenario, but also depending on driver expectations. The latter is especially visible at ODD limits, where drivers stated that the first encounter influenced their behavior in the second encounter, which is also visible in the handling behavior of the ODD limit. Furthermore, some participants explicitly tested the function's behavior in the first encounter with according attention attributed to the road and function, visible, amongst others, in the movement of the hands towards the steering wheel. In conclusion, late interventions do not necessarily need to result in worse intervention quality and late interventions do not necessarily indicate a lacking detection of intervention needs. Reasons for non-successful interventions rather seem to lie elsewhere.

Two non-successful interventions were observed, one for the fast lane drift (0.6 m/s; L2H-off function) and one for the first encounter with the lane end (L2H-on function). A systematic disadvantage for L2H-off functions in terms of safety can thus not be derived (CQ5: safety level). In both cases, the failure to intervene did not result from a reduced amount of control on the steering wheel, as in both cases drivers established hold of steering control beforehand.

The driver who was using the L2H-off function established hold of steering control after receiving a DMS request shortly before onset of the lane drift. After analysis of these cases, changes to the DMS and HMI design seem to be sensible. Firstly, EOD requests should be terminated only after visual attention has been established, not after an arbitrarily quick glance towards relevant AOIs has been detected. Rather, the processing of relevant information and an overview on the current traffic situation need to be ensured. Secondly, the HMI design of DMS requests and a potentially more urgent FDCR should be clearly distinguishable to enforce a correct, timely response by the user.

Whereas conclusions regarding the effect of functional design are admissible, the study design does not allow for conclusions regarding the drift velocity on handling performance. Albeit one incident occurred in the fast lane drift but none in the slower drift scenario, this effect cannot be attributed to drift velocity, for reasons discussed above and also based on the unbalanced sequence of scenarios. Likewise, no conclusions towards FDCR timing can be drawn, as many participants either intervened before a FDCR was issued or deliberately waited to see how the system would behave. Only in case of the one incident observed in the lane end scenario, an earlier FDCR might have enabled a successful handling of the situation, as the driver only looked at the traffic scene after the FDCR had been terminated in the HMI. However, this late intervention would not have occurred, had the FDCR been understood as such.

5.5.3.3 CQ4 (Mode confusion)

Driver role understanding as well as function understanding was (very) good, but not flawless (CQ4: mode confusion). Interestingly, the average agreement to the necessity to keep at least one hand on the steering wheel after activation of the L2 system decreased after use in the L2H-on group, an effect that might be due to the tolerance of the L2H-on function’s DMS towards hands-free episodes (i.e., 15 s hands-free driving until first DMS warning) or to the relative simplicity of the track with little need for lateral control. Positive changes in driver role understanding were observed in such a way that higher agreement to those seating positions (L2H-on) and visual attention (L2H-off) adequate during manual driving was perceived necessary after the drive for L2 use as well. Overall, the meaning of DMS requests was univocally understood subjectively, although case examples as the ones resulting in collisions with the barrier reduce this finding slightly.

5.5.4 References

Damböck, D., & Bengler, K. (2012). Übernahmezeiten beim hochautomatisierten Fahren. In *5. Tagung Fahrerassistenz*.

Körber, M. (2019). Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Associa-*

tion (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics (Vol. 823, pp. 13–30). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-96074-6_2

Metz, B., Landau, A., & Just, M. (2014*).* Frequency of secondary tasks in driving–Results from naturalistic driving data*. Safety science, 68,* 195-203. doi:10.1016/j.ssci.2014.04.002

Mok, B. K. J., Johns, M., Lee, K. J., Ive, H. P., Miller, D., & Ju, W. (2015, June). Timing of unstructured transitions of control in automated driving. In *2015 IEEE intelligent vehicles symposium (IV)* (pp. 1167-1172). IEEE.

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car. In A. L. Kun (Ed.), *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '12* (51–58). New York, New York, USA: ACM Press. https://doi.org/10.1145/2390256.2390264

Schneider, N., Ahrens, L., & Pruksch, A. (2022). Controllability of lateral drift failures while driving with SAE Level 2 Advanced Driver Assistance Systems. In *14. Workshop Fahrerassistenz und Automatisiertes Fahren*, pp. 183-194.

Transportation Research Board (2000). *Highway capacity manual.* Washington, DC, 2(1).

5.5.5 Appendix

5.5.5.1 Screener for recruitment in Germany

***Welche Erfahrungen haben Sie gesammelt mit:**

	unbekannt	bekannt, aber nie genutzt	selten genutzt	regelmäßig genutzt
Tempomat (CC): Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit.	○	○	○	○
Abstandsregeltempomat (ACC): Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Mindestabstand zum Vorderfahrzeug ein. *Weitere Bezeichnungen sind z.B. aktive Geschwindigkeitsregelung bei BMW, adaptiver Geschwindigkeitsregler bei Opel oder Distronic (Plus) bei Mercedes Benz*	○	○	○	○
Teilautomation (ACC + Spurhalteassistent; L2): Dieses System regelt die Geschwindigkeit des Fahrzeugs auf eine eingestellte Geschwindigkeit und hält dabei immer einen festgelegten Mindestabstand zum Vorderfahrzeug ein. Außerdem erkennt dieses System die Fahrstreifenbegrenzungen und hält das Fahrzeug in den Begrenzungen. *Das System heißt z.B. Travel Assist bei VW, Intelligent Drive Assist bei Ford oder Autopilot bei Tesla.*	○	○	○	○

***Ich habe Erfahrung mit einem Assistenzsystem, das…**

Bitte wählen Sie alle zutreffenden Optionen aus. (Mehrfachnennungen möglich)

	Habe ich noch nie benutzt	Habe ich bereits mehrfach genutzt	Nutze ich regelmäßig (mehrmals im Monat)
…einen von mir ausgewählten Mindestabstand zum Vorderfahrzeug herstellt (d.h., mein Fahrzeug passt eigenständig die Geschwindigkeit an das vorausfahrende Fahrzeug an).	○	○	○
… die Geschwindigkeit des Fahrzeuges (entsprechend der eingestellten Geschwindigkeit oder den Geschwindigkeitsbegrenzungen) kontrolliert (d.h., mein Fahrzeug übernimmt eigenständig Beschleunigung und Bremsen)	○	○	○
… das Fahrzeug im eigenen Fahrstreifen hält (z.B. lenkt mein Fahrzeug eigenständig gegen, damit mein Fahrzeug nicht vom Fahrstreifen abkommt).	○	○	○

5.5.5.2 System understanding (before drive)

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute zu nutzende **Fahrerassistenzsystem**.

	Nicht zutreffend	Zutreffend	Unsicher
Das Level 2 System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer.	○	○	○
Das Level 2 System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden.	○	○	○
Wenn das Level 2 System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit.	○	○	○
Das Level 2 System passt die Geschwindigkeit an langsamer vorausfahrende Fahrzeuge an, außer diese stehen oder die Differenzgeschwindigkeit ist zu hoch.	○	○	○
Das Level 2 System erkennt immer, wenn es eine Situation nicht meistern kann.	○	○	○
Das Level 2 System lenkt selbstständig innerhalb des eigenen Fahrstreifens.	○	○	○
Wenn das Level 2 System aktiv ist, hält es einen voreingestellten Mindestabstand zum vorausfahrenden Fahrzeug ein.	○	○	○
Wenn das Level 2 System aktiv ist, steuert es die Geschwindigkeit des Fahrzeugs entsprechend der eingestellten Geschwindigkeit.	○	○	○
Wenn das Level 2 System aktiv ist, führt es Fahrstreifenwechsel durch, ohne dass der Fahrende dabei selbst lenken muss.	○	○	○

5.5.5.3 System understanding (after drive)

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte Fahrerassistenzsystem.

	Nicht zutreffend	Zutreffend	Unsicher
Das Level 2 System erfordert nach Aktivierung zu jeder Zeit mindestens eine Hand des Fahrers am Steuer.	○	○	○
Das Level 2 System kann jederzeit vom Fahrer durch Bremsen, Beschleunigen oder Lenken übersteuert werden.	○	○	○
Wenn das Level 2 System aktiviert ist, ist das System verantwortlich für die Fahrsicherheit.	○	○	○
Das Level 2 System passt die Geschwindigkeit an langsamer vorausfahrende Fahrzeuge an, außer diese stehen oder die Differenzgeschwindigkeit ist zu hoch.	○	○	○
Das Level 2 System erkennt immer, wenn es eine Situation nicht meistern kann.	○	○	○
Das Level 2 System lenkt selbstständig innerhalb des eigenen Fahrstreifens.	○	○	○
Wenn das Level 2 System aktiv ist, hält es einen voreingestellten Mindestabstand zum vorausfahrenden Fahrzeug ein.	○	○	○
Wenn das Level 2 System aktiv ist, steuert es die Geschwindigkeit des Fahrzeugs entsprechend der eingestellten Geschwindigkeit.	○	○	○
Wenn das Level 2 System aktiv ist, führt es Fahrstreifenwechsel durch, ohne dass der Fahrende dabei selbst lenken muss.	○	○	○
Wenn ich, während das Level 2 System aktiviert ist, bremse, fallen alle Assistenzfunktionen weg und ich muss selbstständig lenken sowie Geschwindigkeit und Abstand regeln, bis ich diese wieder aktiviere.	○	○	○

Es gibt einen Unterstützungsmodus, bei dem das Fahrzeug die Beschleunigung regelt aber der Fahrende selbstständig lenkt.	○	○	○
Es gibt einen Unterstützungsmodus, bei dem das Fahrzeug die Lenkung übernimmt aber der Fahrende selbstständig die Beschleunigung regelt.	○	○	○

5.5.5.4 Role understanding (before drive)

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute zu nutzende Fahrerassistenzsystem.

	Nicht zutreffend	Zutreffend	Unsicher
Ich muss das Level 2 System stets überwachen, wenn das System aktiviert ist und mich vergewissern, dass das System die Fahraufgabe entsprechend den aktuellen Bedingungen (z. B. Geschwindigkeitsbegrenzungen, Abstand zu anderen Verkehrsteilnehmern, Straßenbedingungen) durchführt.	○	○	○
Ich darf mich mit anderen Tätigkeiten, wie z.B. E-Mails schreiben, beschäftigen, wenn das Level 2 System aktiviert ist.	○	○	○
Ich muss zu jeder Zeit bereit sein die Fahraufgabe übernehmen zu können.	○	○	○
Ich kann das L2 System jederzeit übersteuern.	○	○	○
Ich darf mich von der Überwachung des Verkehrsraums abwenden, wenn das Level 2 System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind.	○	○	○
Ich muss dem Straßenverkehr die gleiche Aufmerksamkeit widmen wie beim Fahren ohne Fahrerassistenzsysteme (z. B. Straßenzustand, Interaktion mit anderen Fahrern).	○	○	○
Ich muss die gleiche oder eine ähnliche Sitzposition wie beim Fahren ohne Fahrerassistenzsysteme einnehmen.	○	○	○

5.5.5.5 Role understanding (after drive)

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte Fahrerassistenzsystem.

	Nicht zutreffend	Zutreffend	Unsicher
Ich muss das Level 2 System stets überwachen, wenn das System aktiviert ist und mich vergewissern, dass das System die Fahraufgabe entsprechend den aktuellen Bedingungen (z. B. Geschwindigkeitsbegrenzungen, Abstand zu anderen Verkehrsteilnehmern, Straßenbedingungen) durchführt.	○	○	○
Ich darf mich mit anderen Tätigkeiten, wie z.B. E-Mails schreiben, beschäftigen, wenn das Level 2 System aktiviert ist.	○	○	○
Ich muss zu jeder Zeit bereit sein die Fahraufgabe übernehmen zu können.	○	○	○
Ich kann das Level 2 System jederzeit übersteuern	○	○	○
Ich darf mich von der Überwachung des Verkehrsraums abwenden, wenn das Level 2 System aktiviert ist und keine anderen Fahrzeuge in meiner Nähe sind.	○	○	○
Ich muss dem Straßenverkehr die gleiche Aufmerksamkeit widmen wie beim Fahren ohne Fahrerassistenzsysteme (z. B. Straßenzustand, Interaktion mit anderen Fahrern).	○	○	○
Ich muss die gleiche oder eine ähnliche Sitzposition wie beim Fahren ohne Fahrerassistenzsysteme einnehmen.	○	○	○

5.5.5.6 Subjective monitoring performance

*Wie aufmerksam haben Sie das Systemverhalten und den Umgebungsverkehr überwacht, wenn Sie das Level 2 System aktiviert hatten?

	Sehr unauf-merksam	2	3	4	5	6	Sehr auf-merksam
	○	○	○	○	○	○	○

5.5.5.7 NDRT engagement

***Stellen Sie sich vor, Sie wären in Ihrem privaten Fahrzeug, das über eine Level 2 Assistenzsystem (wie Sie es heute erlebt haben) verfügt. Bitte geben Sie an, wie oft Sie sich mit folgenden Tätigkeiten beschäftigen würden, wenn Sie auf der Autobahn fahren und das Level 2 Assistenzsystem aktiviert ist.**

	Nie	Sehr selten	Selten	Gelegentlich	Oft	Sehr oft
Handy oder ähnliches Gerät (Laptop, externes Navi, Tablet, ...) in der Hand - **Bedienung** z.B. SMS/WhatsApp Nachrichten verfassen oder lesen	○	○	○	○	○	○
Handy oder ähnliches Gerät (Tablet, ...) in der Hand - **Sprechen** z.B. Telefonieren ohne Freisprechanlage; Aufnehmen von Sprachnachrichten	○	○	○	○	○	○
Handy oder ähnliches Gerät (Tablet, ...) fest installiert bzw. mit Freisprechanlage verbunden - **Sprechen** z.B. Telefonieren mit Freisprechanlage; aufnehmen von Sprachnachrichten über Sprachbefehle	○	○	○	○	○	○
Bedienung von Systemen im Fahrzeug (nicht direkt relevant für die Fahraufgabe) z.B. Bedienen der integrierten Navigation; Einstellungen im Infotainmentsystem vornehmen; Verstellen des Sitzes; Einstellen der Klimaanlage	○	○	○	○	○	○
Essen/Trinken/Rauchen z.B. Öffnen einer Dose; Essen eines Apfels; Anzünden einer Zigarette	○	○	○	○	○	○
Körperpflege/Make-Up z.B. Frisieren; Make-Up; Nagelpflege	○	○	○	○	○	○
Interaktionen mit Beifahrern z.B. Unterhalten mit Beifahrer; Gestikulieren vor Beifahrer; Blicke zum Beifahrer	○	○	○	○	○	○
Suchen; Greifen; Kramen z.B. in einer Tasche	○	○	○	○	○	○
Sonstiges	○	○	○	○	○	○

5.5.5.8 DMS evaluation

*Beachten Sie bitte, dass die folgenden Fragen sich nicht auf das Gesamtfahrzeug oder die Simulation beziehen, sondern auf das heute erlebte Fahrerbeobachtungssystem (Warnhinweise aufgrund von Fahrerverhalten, **nicht** Übernahmeaufforderungen).

Bewerten Sie hier nur die Umsetzung des soeben erlebten Fahrerbeobachtungssystems.

	Stimme überhaupt nicht zu 1	2	3	Neutral 4	5	6	Stimme voll zu 7	Ich habe keine Warnhin-weise be-kommen
Die Warnungen kamen zu häufig.								
Bei einer Level 2 automatisierten Fahrt würde ich mich mit dem Warnsystem sicherer fühlen als ohne das Warn-system.								
Ohne die Warnungen würde ich mich mehr mit fahrfrem-den Tätigkeiten (z.B. am Tablet spielen, Handy bedie-nen, trinken, etc.) beschäftigen.								
Ich empfand die Warnungen als lästig.								

	Nie 1	2	3	4	5	6	Immer 7	Ich habe keine Warnhin-weise be-kommen
Ich konnte nachvollziehen, warum eine Warnung ertönt ist.								
Die richtige Reaktion nach Auftreten einer Warnung war mir klar.								
Ich habe die Warnungen bewusst ignoriert.								
Die Warnungen lenkten meine Aufmerksamkeit wieder auf das Fahrgeschehen.								

6 Requirements for L2H-off

The goal of this project was to generate a reliable set of data, information and knowledge by application of different methods to assess potential challenges related to L2 hands-free driving as well as to derive recommendations on how these shall, should or may be compensated. Two transformation steps were central to achieving the project's goals: Firstly, the results and review of all analyses have to be aggregated and interpreted with regard to the five potential challenges for hands-free monitoring motivating the project (for an overview, see Chapter 1). In a second step, the conclusions drawn on driver behavior in interaction with L2 functions and the design options discussed shall be transformed into technology-independent guidance on the design of L2H-off functions. The final SP 5 (Figure 6-1) summarizes the efforts described above by providing an overall assessment for each of the five challenges based on the state of the art, discussions lead within this project and, first and foremost, the results of all data collections (first transformation; Chapter 6.1). The second transformation of knowledge generated within this project aims at providing guidance on the design of L2H-off functions (Chapter 6.2). The goal has been to create a stand-alone document comprising basic principles for the design of L2H-off functions within the context addressed by this project.

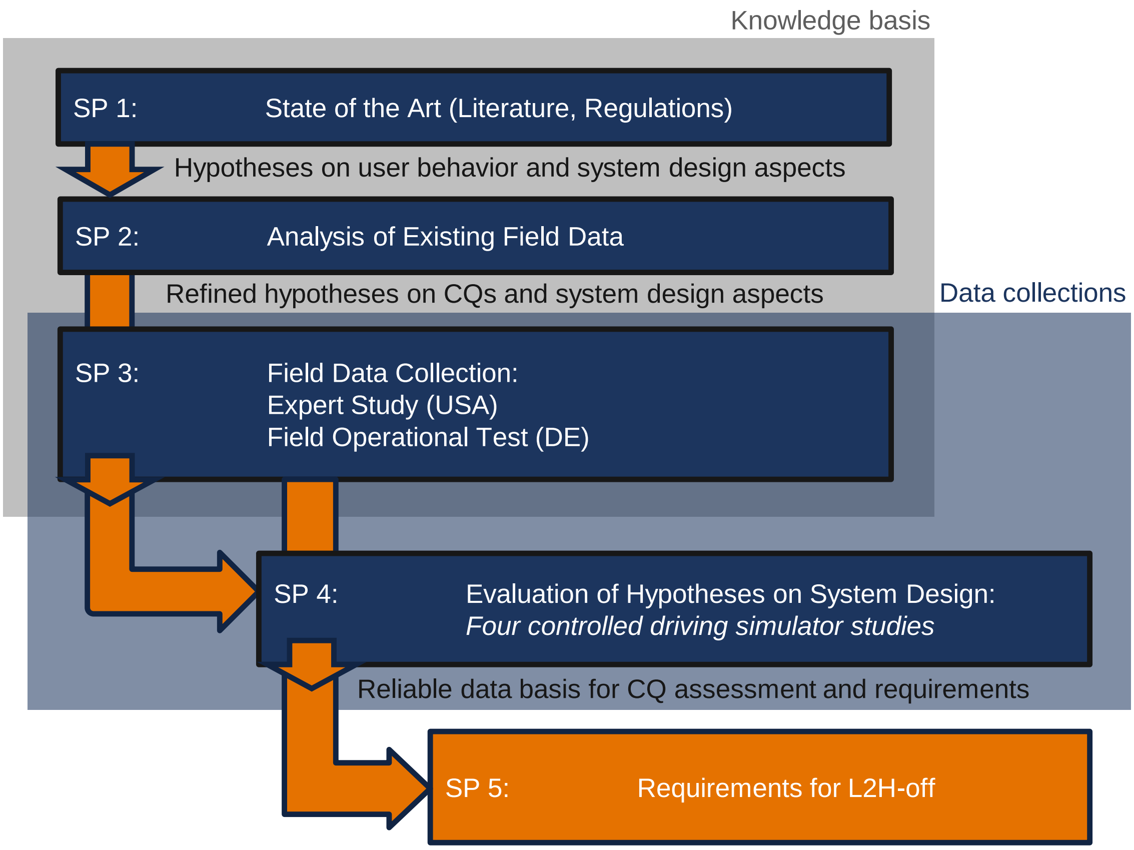


Figure 6-1: Overview on the five subprojects and the role of SP 5 within the project.

6.1 Aggregation of Analysis Results

Documentation by N. Grabbe, K. Bengler (Lehrstuhl für Ergonomie, TU München)

In the following, the results of all project studies (for an overview, see Chapter 1) are aggregated and compared with literature findings in order to derive conclusions for each of the five challenges and questions (CQs) (cf. Section 2.1). In this regard, aggregation means that the detailed results per study are transferred to a more abstract level in order to provide compact and concise findings for each CQ on a more general level. The five CQs were the starting point of our study. In Sections 6.1 and 6.2 the aggregated results are transferred as conclusions into system design recommendations and guidelines in. In total, eight different studies with specific contributions regarding the CQs were conducted (see Table 6-1). The studies were based on a literature review in the beginning (see Table 6-1) and their results will be mirrored with the conclusions drawn from prior studies in the following.

Table 6-1: Overview of project studies and their respective contributions to the CQs. Green = studies based on field data, yellow = study based on survey data, blue = studies based on driving simulator data

Study	CQ1	CQ2	CQ3	CQ4	CQ5
Literature	x	x	x	x	(x)
Field data	x				
Expert study US				x	x
FOT DE	x	x	x	x	x
User survey US			x	x	
Simulator study 1	x	x	x	x	x
Simulator study 2	x	x	x	x	x
Simulator study 3	x	x	x	x	x
Anchor study (study 4)	x	x	x	x	x

The structure regarding each CQ is as follows: First, the definition and the considered constructs and metrics are provided and described. Second, conclusions on every construct level are derived leading to main conclusions regarding each CQ in the end.

6.1.1 CQ1: Hands-off = Mind-off?

Documentation by N. Grabbe, K. Bengler (Lehrstuhl für Ergonomie, TU München)

6.1.1.1 Definition

There are concerns that a lesser driver involvement in the driving task (exacerbated by the lack of contact with the steering wheel during L2H-off) will reduce the driver's attention to the driving task probably leading to a state of mind-off.

In order to answer the concern of mind-off, we should take a closer look at the driving task. According to Geiser (1985), the primary, secondary, and tertiary driving tasks can be distinguished. Here, the primary driving task is the focus which requires the driver to keep the vehicle on course and adapt the speed if required, e.g. to maintain a safe distance to other traffic participants. In particular, three levels of the driving task can be differentiated: navigation, guidance, and stabilization. If we combine the three information processing level of perception, cognition, and action with the three driving levels, then we can recognize that for navigation and guidance, mainly visual perception in form of gazes to the road scene is required whereas for the stabilization, action (motor activity) in form of hand position in relation to the steering control and foot position in relation to the pedals is required, for which the hand position is decisively when comparing L2H-on and L2H-off. Moreover, also cognitive processes need to be considered since mind-off directly refers to human information processing. The focus in case of driving should be directed to visual perception and motor activity. These are observable parameters indicating mind-on/-off and highly relevant for driving. Nevertheless, Glaser et al. (2016) have shown that the duration of glancing at the road did not influence the detection rate or the reaction time in case of imminent danger which indicates that not only perception, but also cognition needs to be considered as well. For this, we refer to the conclusions of CQ3 and 4 in Sections 6.1.3 and 6.1.4.

In total, five constructs and eight metrics (see Table 6-2) were used to draw conclusions regarding CQ1.

Table 6-2: Assignment of constructs and metrics for CQ1

Information process level	Construct	Metric
Perception	Visual attention	• Number of eyes-off road glances >2 s • Attention ratios (eyes-on road, instrument cluster/steering wheel, other)
	Monitoring	• Number of hands-off/eyes-off warnings • Subjective rating of monitoring performance
	Perceptual readiness at transitions	• Number of eyes-off road glances >2 s and visual attention ratio 30 s before and 10 s after the transition
Action/motor activity	Motoric ability for safe vehicle guidance	• Hands-on/-off proportion • Classified hand positions
	Motoric readiness at transitions	• Level of motoric control rating-based on hand position 30 s before and 10 s after the transition

In addition, we conducted interviews and noted remarks in an experimental protocol that are used when making a contribution to a respective construct. For the metrics' definitions, we refer to the respective study chapters

6.1.1.2 Conclusions

In the following, the conclusions are presented and compared to findings in the literature.

Visual attention

We observed a tendency for higher eyes-on road proportion for L2H-off subjects than L2H-on subjects. Furthermore, L2H-on participants showed a higher dispersion with outliers towards lower eyes-on road proportions. In particular, simulator study 1 (Section 5.2) showed significantly higher eyes-on road proportion for L2H-off driving compared to L2H-on. If a visual NDRT is offered motivationally, more eyes-off road glances above 2 s could be observed for L2H-off participants. Without a motivational offer of NDRT, no difference could be found between functions. The 3 s-DMS implemented in Study 3 (Section 5.4) showed significantly fewer eyes-off road glances above 2 s than the 5 s-DMS but no difference could be found in the eyes-on road ratio. On the one hand, a slight improvement of visual attention by a more conservative DMS-approach can be found. On the other hand, DMS-request after 3 s visual inattention were rated rather as too short compared to 5 s which might increase the probability of disuse in daily life. Most importantly, no objective difference in terms of safety could be found between the 3 s-DMS and the 5 s-DMS. In general, over all studies, the 5 s-DMS was evaluated well in terms of timing (good trade-off between safety and comfort) and the warnings were rated as clear in the meaning of perceiving the warning, understanding its reason, and comprehending what the task driver should do. In addition, in simulator study 1 (Section 5.2), no L2H-off participant asked for earlier warnings than 5 s. Effects with regards to first contact use (experts vs. novices) could not be found.

According to literature, L2H-off without DMS leads to an increased visual distraction compared to L2H-on, ACC, and manual driving (Boos et al., 2020; Josten, 2021; Llaneras et al., 2013; Noble et al., 2021; Othersen, 2016). All L2H-off experiments in this project have been conducted with a DMS and don´t show the effects reported in literature. Comparing the experiments and their results leads to the conclusion that an adapted DMS eliminates the negative effects regarding visual attention during hands-free monitoring that were identified in literature and even converts them into positive effects.

Moreover, literature shows that with 3-step monitoring requests, L2H-off gaze ratio to the road is better than without (Blanco et al., 2015; Kurpiers et al., 2019; Llaneras et al., 2017; Victor et al., 2018). Furthermore, 3-step monitoring requests have been shown to prevent very long eyes-off-road times (e.g., 4 s eyes-off road) (Victor et al., 2018). Our results confirm these findings in the way that L2H-on drivers show higher maximum of eyes-off road times than L2H-off drivers. Thus, we can conclude that a 3-step DMS improves the eyes-on road ratio and leads to fewer long eyes-off road times. Overall, the implemented DMS fulfills its purpose.

Monitoring

The warning cascade (3-steps) is predominantly terminated in stage one. If a visual NDRT is offered motivationally, many eyes-off warnings have occurred. Without a motivational offer of NDRT, a rather small number could be observed. No differences in the number of warnings could be found between 3 s-DMS and 5 s-DMS which could be an indicator that subjects adapted their gaze behavior to the DMS cascade. Furthermore, in the FOT, no differences in hypothetical eyes-off warnings (5 s, 8 s, 13 s) between L2H-on users and L2H-off users could be observed but high outliers exist for L2H-on subjects. The lesser effectiveness of hands-on wheel DMS during NDRT engagement may be further illustrated by one L2H-off driver in the anchor study, who, during the lane drift, reacted primarily with hands-on wheel to an EOR but only very briefly attributed attention to the driving scene, hence not noticing the imminent collision with the barrier. This provides further evidence that a well-designed EOD might be beneficial. In addition, the objective findings on monitoring coincide with the subjective assessment of monitoring. The warnings were evaluated by the subjects as clear in the meaning of perception, understanding their reason, and comprehending what action should be taken by the driver. However, single cases of misunderstandings regarding the correct behavior upon FDCR in difference to DMS requests were observed (see Section 5.5.2). In addition, a lower annoyance by eyes-off warnings than by hands-off warnings could be found. Effects with regards to first contact use (experts vs. novices) could not be observed.

Literature reports that monitoring behavior is better when using 3-step monitoring requests than without (Blanco et al., 2015; Kurpiers et al., 2019; Llaneras et al., 2017; Victor et al., 2018). Our results confirm these findings, hence the monitoring behavior is improved by 3-step monitoring requests.

Perceptual readiness at transitions

According to the FOT results, driver-initiated activations tend to be somewhat more distracting for L2H-off participants. This is also confirmed by higher eyes-on instrument cluster attention ratio and some subjects reported that L2H-off is more complex than L2H-on. These findings indicate that a clear-cut function should be preferred over a multi-step function. Prior to system-initiated deactivations (with & without FDCR), L2H-off users tended to be more alert than L2H-on users indicating a higher potential to anticipate function limits. Otherwise, no differences or abnormalities in transitions regarding mental readiness could be found.

Motoric ability for safe vehicle guidance

L2H-off users had their hands on the steering wheel on average for approximately 45 % of the time during the FOT but a large inter-individual spread in hands-on wheel times exists. It can thus be argued that a certain awareness of the need for hands on the steering wheel exists which may be interpreted as a balanced trust. If L2H-off drivers had a hold of steering control, then they rather had both hands on the steering wheel instead of just one hand. In contrast, L2H-on driver mainly kept contact to the steering wheel but some outliers exist with relatively high hands-off proportions, indicating misuse. Even when within the admissible hands-off tolerance of the DMS, hands-off wheel is considered misuse during L2H-on use. If L2H-on drivers had hold of the steering wheel, then they had both hands on the steering wheel for 60% of the time, and just one hand on the steering wheel during 40% of the time. The FOT also showed that during L2H-off use, the participants most frequently (30 % of the time) put their hands on their laps and were therefore ready to quickly intervene. The second and third most frequently observed hand positions are both hands (12 % of the time) or one hand (7 % of the time) on the upper half of the steering wheel, respectively. Therefore, the hands are mainly placed close to or on the steering wheel. Placing the hands behind the head, far away from the steering wheel (e.g., passenger seat, grabbing for objects), or occupying the hands with objects was not observed during the FOT with the sole exception of infotainment use for a few seconds. No effects regarding first contact use (experts vs. novices) could be observed.

According to literature, most drivers deliberately leave their hands on the steering wheel when using L2H-off functions (Naujoks et al., 2015). This high percentage of hands-on steering wheel cannot be confirmed but a significant and appropriate proportion of hands-on times exists, which shows a balanced trust as well as an awareness of the need for hands on the steering wheel when it is necessary.

Motoric readiness at transitions

In the FOT, it could be seen that the closer the transition, the closer the L2H-off subjects get to the ready-to-drive hand position (motoric control). The anchor study (study 4) confirms this finding for upcoming, visible direct control requests and silent function failures. For the latter, changes in hand position occur proactively when observing the failure until a decision is made on the need for intervention. Thus, an awareness of when it is necessary to put the hands on the steering wheel exists, given sufficient (visual) involvement in the driving task. Effects with regard to first contact use (experts vs. novices) could not be found.

According to literature, drivers adapt their engagement to the traffic state, engaging more in processing the tertiary task while driving in low-velocity ranges compared to driving at higher speeds (Naujoks et al., 2016). We can confirm this finding by concluding that drivers adapt their hand position to the traffic state in such a way that they are aware of when it is necessary to put their hands on the steering wheel.

Main conclusions

Overall, evidence suggests that hands-off does not lead to mind-off when using a proper glance-based DMS because visual attention is improved. The monitoring behavior in interaction with the glance-based DMS is good, since drivers establish perceptual and motoric readiness and actively prepare for transitions (anticipation). In general, sufficient motoric ability for safe vehicle guidance could be observed. The glance-based DMS clearly showed positive effects resulting in better visual attention and good monitoring behavior. In order to completely assess the proposed challenge of mind-off (CQ1), we refer to the cognitive component analysis in CQ 3 and 4 in Sections 4.1.3 and 4.1.4.

6.1.2 CQ2: Prolonged transition times

Documentation by J. Josten (fka GmbH)

6.1.2.1 Definition

The necessity for the driver to move at least one hand towards the steering wheel to establish physical readiness for direct control and to provide steering input is undisputed. However, the impact of this motoric process as an additional component of driver interventions is unclear. There are concerns that the process of returning the hand(s) to the steering wheel (i.e., the time needed until establishing hold of steering control) as well as longer reaction times in general may lead to an increased risk of accidents when interacting with L2H-off functions in comparison to a reference that does not allow hands to be removed from the steering wheel or provides a different level of assistance. The following chapter is concerned with this second challenge/question motivating the current project (CQ2; prolonged transition times).

As discussed for single studies of this project (e.g., Study 4, Section 5.5), the timing of driver interventions is highly dependent on the affordances and (perceived) urgency of the scenario at hand. Next to the analysis of the timing and type of driver interventions, the quality of transitions from function to drivers taking direct control will be analyzed in the following chapter (see Table 6-3), to give perspective on the relevance of any identified differences regarding the timing of actions.

Table 6-3: Assignment of constructs and metrics to CQ2

Construct	Metrics
Timing of driver actions	• (Mean / Max.) Reaction time to DMS requests [s] • Hands-on (reaction) time [s] • Intervention time [s]
Type of driver interventions	• Number of driver- and system-initiated deactivations (for FOT) • Intervention type
Outcome of transitions in direct control	• Number of incidents • Distance-based metrics (e.g. TTC_{min}, distance to lane end at lane change) • Vehicle-dynamics based metrics (lateral acceleration) • TOC-rating of transitions
Other	• Interview data • Video-based analyses

Intervention times can only be compared between L2 functions due to the necessity to clearly detect the onset of lateral or longitudinal input, whereas the outcome of driver actions within the driving situation under test can also be derived for participants driving without any driver assistance (L0; only for results of Study 1, see Section 5.2.3). Analyses thus differentiated between metrics indicating the first reaction or direct control input of the driver during L2 use and metrics indicating the overall outcome of driver interventions for L2 and L0 (see subchapter on outcome of transitions in direct control).

6.1.2.2 Conclusions

In the following, the conclusions based on the project's studies are presented and compared to findings in the literature. All data collections apart from SP2 and the online user survey provided input to the conclusions on prolonged transition times.

Timing of driver actions

As reported in the literature overview (see Chapter 1.1), prior studies investigating the effects of hands-free driving in direct comparison to L2H-on functions have often been conducted without an adapted driver monitoring system (DMS; Cahour et al., 2021; Garbacik et al., 2021; Gold et al., 2013; Josten, 2021; Othersen, 2016). These studies found at least slightly increased mean transition times in case of system-initiated deactivations or a reduced controllability of transitions. A delay in transition times of approximately 0.3 seconds (Gold et al., 2013; Josten, 2021) has been attributed to the motoric intervention of taking the hands to the steering wheel. Study designs without an adapted DMS however do not allow for the investigation of the potential compensatory effect of a higher visual involvement in the driving task caused by a DMS based on eyes-on detection (EOD) (see Chapter 6.1.1) on the anticipation of limits or the timing of interventions.

EOD-based systems were found to issue a higher number of requests compared to HOD-based (hands-on detection) DMS. Mostly, drivers complied quickly with DMS requests regardless of the L2 function and frequency of requests. No indication was found for earlier compliance with L2H-on function requests, but half of the studies revealed advantageous behavior in the L2H-off group either regarding mean or maximum reaction times to requests. This effect in itself might be due to different reasons, e.g., gaze on road might be established faster than hold of steering control, or the higher tolerance in the timing of the L2H-on DMS cascade might result in a less prompt response, similar to the timing of interventions in less urgent scenarios. Most DMS requests were terminated after the first stage. However, in some occasions – the majority of these in interaction with L2H-off functions – the request cascade was not terminated until the function deactivated and began braking (Stage 3 request). Still, even when L2H-off users terminate DMS requests in later stages, this will result in less time during which the driver is unaccounted for compared to a 15 s onset-delay in hands-on requests (HOR).

Regarding the analysis of hand postures during the use of L2H-off functions, strategic adaptations were observed to establish higher levels of control before system-initiated transitions (see Study 4 in Chapter 5.5 and the FOT in Chapter 4.4). On the other hand, the difference between establishing hold of the steering wheel (hands-on time) and intervention time (first direct control input by the driver) was quite small in the field-operational test (FOT), an effect that might be due to analyzing only interventions after function direct control requests (FDCR) and thus ruling out scenarios with a higher potential for anticipative actions or higher monitoring requirements. The driving simulator studies show that L2H-off users use the available time to take hold of steering control in case of anticipatable system limits such as road works. This is not possible in sudden events such as cut-outs (see Study 1, Chapter 5.2).

Overall, no disadvantage could be observed for L2H-off functions due to different intervention times. In the majority of scenarios, no difference between L2 functions was found at all. It could thus be argued that the adaptations to the L2 design (i.e., the DMS) compensated potential adverse effects of hands-free monitoring. The similarity of interventions between L2 functions differed depending on the scenario, e.g., depending on the possibility of anticipation by the driver. In some scenarios, a FDCR was issued as a last resort, when drivers did not intervene earlier on their own account, i.e., did not or could not anticipate the upcoming functional limit. Where implemented, the timing of FDCR issued before reaching functional or ODD limits ranged between 2 s to about 4 s (Study 2, Chapter 5.3; Study 3, Chapter 5.4; Study 4, Chapter 5.5) over the different studies. The type of L2 function did not systematically influence the time with which drivers intervened after a FDCR, indicating no disadvantage for L2H-off users in case of time-critical requests for direct control. L2H-off users showed no benefit in cases where they did not react to upcoming system limits before the FDCR was issued, i.e., the EOD-based DMS did not increase driver's preparation to the limit for those drivers who did not show anticipative actions before the FDCR.

A seemingly diametrical effect was observed in responses to an upcoming lane change when comparing different simulator studies: A higher degree in anticipative behavior based on lane change behavior was observed for L2H-off users compared to L2H-on users in Study 1 (Chapter 5.2), albeit with no significant effect on the mean timing of first direct input (albeit a significant difference in minimum TTC values). For the same scenario setting, a tendency for earlier interventions in the 5-s-EOR L2H-off group compared to L2H-on users was observed in Study 3. However, only this specific L2H-off group in Study 3 showed indications for anticipation in comparison to L2H-on users and all interventions in this scenario were carried out in time regardless of the experimental group. By contrast, L2H-on users intervened earlier (by tendency) than L2H-off users in face of the upcoming lane change in Study 4. This difference in effect in similar scenarios might be explained with seemingly small differences in the scenario design. Respectively, a higher complexity of the roadworks scenario in Study 1 and 3 (i.e., more surrounding vehicles; Chapters 5.2 and 5.4) and the thereby limited anticipation potential might make it more appealing to conduct a lane change in advance, given the opportunity by anticipation. By comparison, early lane changes in Study 4 (Chapter 5.5) provided no advantage over late, but well-prepared lane changes due to the unobstructed visibility of the upcoming track and lack of interfering traffic. The conclusion that later interventions might as well be results of a more prepared direct input has already been postulated by Gold et al. (2013). Whereas both L2H-off and L2H-on users in Study 4 have plenty of remaining time to conduct the lane change albeit absolute differences in timing, the anticipative effects observed in Study 1 and 3 indicate a higher visual attention to the road that enables users to act based on single cues such as road signs albeit no full visibility of the ODD limit. The advantages observed for L2H-off groups in Study 1 and 3 are thus considered more relevant than the slightly later lane changes with L2H-off functions in Study 4.

Experienced users established contact with the steering wheel descriptively earlier than inexperienced users, an effect most prominently visible for foreseeable ODD limits (see Study 4).

The difference in intervention times between experienced and novice ADAS users in the foreseeable scenario was smaller, but indicated earlier interventions for experienced users as well.

As already reported above, less tolerance regarding inattention (i.e., earlier timings of the first DMS request) did not positively influence the response to EOR requests or interventions in case of system limits. The complexity of the L2 function design (multi-step versus clear-cut) did not influence the timing of actions in transitions. However, users adapt the timing of interventions to the level of assistance used as fallback in case of L2 limits (i.e., L1 or L0).

Type of driver interventions

No difference in the propensity to overrule L2 functions in case of failed adaptations to speed limit reductions was found between L2H-off and L2H-on functions (see Chapters 5.2 and 5.4). The frequently observed steering inventions as first direct driver input could be explained with the driving scenarios being designed to impose lateral control needs, e.g., lane changes, evasive maneuvers, or curves. However, this result shows that although hands-free driving is more common with L2H-off functions, drivers can establish hold of steering control in time when necessary. This effect is in line with the ratings of physical control as discussed for CQ1 (see 6.1.1). Looking at the FOT results (Chapter 4.4), no difference between functions, but between deactivation types became apparent. For system-initiated deactivations, regardless of the function used, the same pattern as in the driving simulator studies was observed, albeit a hypothetically more diverse encounter of scenarios: L2 users intervened nearly always by steering regardless of admissible hand posture, although most FDCR in the FOT were issued as a mere information on system deactivation, thereby providing no time to potentially adapt hand positions based only on a pre-warning of the system. For driver-initiated deactivations, functions were overruled by braking or deactivated by pressing the respective button. Again, no difference between functions was obvious. Whether the difference between system- and driver-initiated deactivations was due to the characteristics of those driving situations where driver- or system-initiated deactivations occurred was not in the focus of this project.

The number of driver- and system-initiated deactivations was only compared for field data, i.e. in the FOT. Data show no indication that drivers might be less willing to overrule L2H-off functions due to lesser contact with the steering wheel. On the contrary, a descriptively higher number of driver-initiated deactivations was observed in the L2H-off drives, whereas a similar number of system-initiated deactivations was observed. However, the results on driver-initiated deactivations need to be weighed against influences of the FOT design (e.g., presence of experimenter only for L2H-off drives; no control over traffic conditions).

Outcome of transitions in direct control

Hands-free monitoring can result in detrimental effects on steering control and situation handling (e.g., Garbacik et al., 2021). Other reports argue that incident rates, given an adapted DMS, are not dependent on hand posture, but rather on other factors such as trust (e.g., Gustavsson et al., 2018; Pipkorn et al., 2021; Victor et al., 2018). Higher maximum steering wheel angles after hands-off monitoring have been observed in prior studies without an EOD-based

DMS when drivers reacted to a FDCR without the possibility to anticipate the intervention need (Josten, 2021). The same effect, with lesser difference between functions, was observed in a test track study (Josten, 2021).

In line with intervention times, no systematic disadvantage in the number of incidents was observed for L2H-off functions. However, not all drivers intervened in case of system limits and not all drivers who did intervene did so successfully, but this effect was not dependent on the L2 function used. The number of incidents observed in some scenarios suggests that these might have been designed slightly too challenging, especially in combination with a visually challenging NDRT, as can be seen, for example, in the number of drivers triggering the risk mitigation function (RMF) – with as well as without L2 assistance (L0) – and the number of participants who subsequently overruled the RMF (see Study 1, Chapter 5.2). Reasons for incidents therefore lie in the affordances of the scenario itself, but also in misunderstandings of DMS requests and FDCR (see Study 4, Chapter 5.5) and in the way the intervention was carried out (i.e., too little steering or braking input with regard to the required input; Study 3, Chapter 5.4).

The differences found in timing are put into perspective when considering the overall handling of the situation for the lane change in Study 4 (Chapter 5.5), as discussed above. Contrastingly, tendencies found in transition times for the lane change in Study 1 (Chapter 5.2) are intensified in terms of metrics including the overall handling of the situation, e.g., the resulting minimum TTC and handling quality. For most comparisons, however, no differences between experimental groups (including the manual driving group were included) were found regarding the metrics used to evaluate the outcome of driver interventions, e.g., minimal TTC to obstacles, the distance to obstacles or ODD limits or TOC ratings of candidates in the field data.

Maximum lateral acceleration, interpreted here as a (relative) indicator for the quality of interventions, varied with the timing and urgency of interventions. A tendency for higher lateral accelerations was observed in groups with descriptively later interventions in predictable scenarios (Study 4, Chapter 5.5) as well as in descriptive comparison of predictable and urgent scenarios (roadworks versus cut-out; Study 1, Chapter 5.2). If differences between L2 functions were observed (Study 1, Chapter 5.2), these differences support the assumption that a higher likelihood for anticipation leads to more controlled interventions, favoring L2H-off functions. This finding does not support the above stated hypothesis that later interventions might primarily be an indicator of better preparation.

No systematic effects were observed when comparing different user samples (i.e., ADAS expertise, cultural differences), exempt the above described dependency of intervention time and resulting lateral accelerations. The complexity of systems as well as the timing of the first EOR had no systematic effect on the intervention success.

Main conclusions

Over all studies, no systematic disadvantage of L2H-off functions could be observed regarding the timing and success of driver interventions. This conclusion is valid across different samples and functional designs. Differences in timing solely due to a physical disadvantage of hands-off wheel should be small in theory as well as based on literature findings (e.g., Josten, 2021), making it likely that other influences, such as physical (i.e., reducing the distance to the steering wheel), visual (see 6.1.1) or mental preparedness to intervene, compensate any influence of reduced physical control. Such compensation would allow L2H-off users to successfully intervene in most situations.

All studies within this project combined the effect of an adapted DMS (with proposed effects on visual attention) with the variation of accepted hands-free episodes during L2 use (with proposed effects on establishing physical readiness). Neither visual attention nor hand posture in L2H-off use or L2H-on use was controlled for, except regarding the boundaries set by the respective DMS. The expected range in effect regarding hands-on or intervention times when hands are not only off the wheel but, e.g., used for interacting with different items or devices, was not targeted in the conducted studies. Any physical disadvantage from hands-off wheel can, however, be expected to be highly dependent on the possibility to anticipate the upcoming system deactivation or limit.

As can be seen from percentages of hold of steering control during the drives, drivers interacting with L2H-on functions show a tendency to make use of the possibility to remove their hands from the steering wheel, although Study 1 did not find differences between L2H-on and L0 regarding hands-off proportions (Chapter 5.2). Thus, the physical disadvantage to establish readiness for direct control might be more likely when interacting with L2H-off functions, but is certainly not exclusive to them. As could further be shown by case analysis (Study 4, Chapter 5.5), the number of incidents observed in the studies in interaction with different L2 functions and in literature (e.g., Victor et al., 2018), hold of steering control is not enough to guarantee a timely and, most importantly, successful intervention.

Timing and quality might have been influenced negatively as the visual NDRTs offered in most studies might have systematically reduced visual attention towards the traffic, which likely reduced anticipation and its associated behavior such as anticipatory hand posture changes. Where anticipation is not possible, FDCR have proven effective in eliciting direct control regardless of hand posture (see also Josten, 2021). No incident in all studies conducted in this project could be clearly linked to hand posture. Therefore, enabling a high visual and potentially cognitive involvement in the driving task seems more relevant than the physical component for EOR-controlled hands-free monitoring.

6.1.3 CQ3: Foreseeable misuse

Documentation by D. Schwarze (fka GmbH)

6.1.3.1 Definition

There are concerns that the use of L2H-off functions will lead to increased foreseeable misuse or disuse, particularly with respect to an increased initiation of non-driving related tasks. In order to address these concerns, the different types of misuse, defined in ISO 21448:2022 should be considered. Depending on the causal relationship to the hazardous behavior, two types of misuse need to be distinguished: direct and indirect misuse. Direct misuse could be a cause for the occurrence of hazardous situations of use. For example, activating the L2 function outside of its ODD is considered as direct misuse which can be viewed as a potential triggering condition leading to unforeseen, potentially hazardous situations of use. Indirect misuse, on the other hand, may result in a reduced interaction quality or controllability of the function's limits (e.g., reduced monitoring behavior, although the function requires continuous monitoring by the driver to ensure that the DDT is performed correctly).

To answer the question of whether L2H-off functions will lead drivers to increased misuse or disuse, three constructs mainly considering indirect misuse were considered (see Table 6-4). They were divided into six metrics and collected in all project studies and the user survey US. In addition, interviews were used when appropriate. For the metrics' definitions, we refer to the respective study chapters (see Chapters 4 and 5).

Table 6-4: Assignment of constructs and metrics for CQ3

Construct	Metric
Distraction	• Objective NDRT (1-9) proportion • Subjective NDRT (1-9) engagement • Number of tasks worked on
Misuse	• Trust (TiA) • Acceptance (CTAM) • Time H-off (L2H-on)
Disuse	• Trust (TiA) • Acceptance (CTAM)

6.1.3.2 Conclusions

In the following, the conclusions are presented and compared to findings in the literature.

Distraction

There was no significant difference between L2H-on and L2H-off functions in terms of involvement in non-driving related tasks. Simulator Study 1 (Chapter 5.2, results on objective NDRT engagement) showed a greater amount of engagement for the L2H-on group while the other studies did not show a clear difference. This effect might thus be a consequence of the different types of NDRTs (reading-typing vs. reading-speaking). The different NDRTs did not reflect any advantage of the hands-off group in terms of hand posture. According to literature the involvement in NDRTs is more common while using L2 than while driving manually (Solís-Marcos et al., 2018; Noble et al., 2021; Llaneras et al., 2013). An increased initiation of NDRTs while using L2 functions compared to manual driving could not be observed in the data (Chapter 5). Independent of the non-driving related tasks motivated in the simulator studies, the FOT and the online user survey (Chapters 4.4 and 4.3) showed that NDRTs in real-traffic tend to be rather “non-critical” activities (e.g., interaction with passengers, vehicle related inputs). Moreover, the FOT (Chapter 4.4) and the simulator studies (Chapter 5) indicated that alerts issued by the function (DMS) seem to interrupt NDRT interactions but also seem to induce more frequent changes in attention (number of fixations) between NDRT and street (results of Chapter 4.3 on NDRTs and Chapter 5.5 for the evaluation of interaction behavior during continuous use). Similar findings of Llaneras et al. (2017) confirm the finding that DMS alerts are an effective countermeasure to interrupt NDRT interactions. Overall, a DMS can lead to a reduction of NDRTs and therefore avoidance of misuse.

Misuse

When considering the metric hands-off time for the construct misuse when driving with a L2H-on function, users of these functions apparently use the possibility to temporarily remove their hands from the steering wheel. Further, some L2H-on users are sometimes not even aware that they are required to keep their hands on the steering wheel (Chapter 4.3; driver role during L2 use). The FOT and on-road studies of Banks et al. (2018) and Morando et al. (2021) showed similar outcomes. The authors observed that some L2H-on drivers appeared to use the function in a hands-free fashion. On the other hand, L2H-off drivers, even with experience, do not always make use of the opportunity to take their hands off the steering wheel in all situations (Chapter 4.3; Chapter 4.4). It might thus be beneficial to be aware of the function’s limitations and the ODD. Overall, it seems that at least some drivers adapt their hand posture to the current situation of use and that a DMS based on hands-off wheel detection does not result in continuous contact with the steering wheel.

Disuse

According to Kim et al. (2021) or Feldhütter et al. (2019) the likelihood of engaging in non-driving related tasks increases with a better attitude towards automated driving features. All our studies (trust scale; Chapter 4.4 and Chapter 5) showed trends for higher trust in L2H-off functions but not overly high compared to L2H-on functions. The slightly higher trust could be explained due to more usage situations (e.g., more frequent DMS alerts and, therefore, more frequent interaction with the function) with the L2H-off functions than with the L2H-on function during the test drives. However, the differences in the level of trust are not so high that an increase in misuse or disuse can be expected. The authors’ findings (Kim et al., 2021 and Feldhütter et al., 2019) can, therefore, not be confirmed by the studies’ results. Furthermore, the studies with experienced L2 drivers also showed that gaining experience over time with L2 functions seems to increase the level of trust and acceptance but does (subjectively) not lead to more misuse or disuse (e.g., Chapter 4.3).

Main conclusions

In conclusion, the studies showed no evidence that the use of L2H-off functions leads to an increase in the engagement in non-driving related tasks compared to L2H-on functions. Instead, the outcomes hint at intentional abuse/misuse of L2H-on functions for some L2H-on users who seem to use the opportunity to temporarily remove their hands from the steering wheel while the function is active. The trust and acceptance evaluations revealed no significant differences between the levels of trust of the L2 functions and thus do not indicate increased levels of disuse or misuse through the use of L2H-off functions.

6.1.4 CQ4: Mode confusion

Documentation by P. Dautzenberg (Institut für Kraftfahrzeuge, RWTH Aachen University)

6.1.4.1 Definition

There are concerns that with the introduction of L2H-off functions, drivers are no longer aware of their tasks and roles as drivers, the functional design and logic of the function used as well as its ODD, which also makes it difficult to anticipate function limitations. These concerns arise from the additional physical degree of freedom that L2H-off functions provide, but also going along with a less direct connection to the vehicle. As described beforehand (see e.g., Chapter 4.3), mode confusion can arise if a driver experiences two or more different functions. The risk for mode confusion is increased when multiple functions are available within one vehicle and the alternating system modes appear similar to the user (Boos et al., 2020; Kurpiers et al., 2020).

Mode confusion is one possible reason for deficient mode awareness. Mode awareness again combines two major aspects (Boos et al., 2020; Kurpiers et al., 2020): Knowledge-based confusion and behavior-based confusion. The first aspect, knowledge-based confusion, implies the knowledge about which mode is currently active and the knowledge about the respective function’s abilities and limits, as well as the tasks and roles as driver. Understanding the system and its limitations as well as understanding one's own tasks when interacting with the function appear to be essential prerequisites for mode awareness. The second aspect of mode awareness, behaviour-based confusion, implies mode compliant behaviour. In distinction to misuse, mode errors (behavior not compliant with the currently active mode) are described as actions that are appropriate for an erroneously assumed automation mode but inappropriate for the actual, currently active mode. Thereby, mode errors correspond to unintentional errors and can in turn be seen as an indicator of poor mode awareness. Therefore, when assessing mode confusion, it is important to gather objective as well as subjective data to validly interpret potentially occurring behavior that is not compliant with the respective mode.

As described in Chapter 5.1, there is only little to no research on the occurrence of mode confusion when switching between different L2 functions (H-on and H-off). However, there is some literature that provides guidance on what design principles can be helpful in reducing the risk of mode confusion. Systems with clear-cut modes, for example, of either on or off should increase mode awareness and decrease mode confusion as there are fewer transitions and function characteristics the driver has to go through and differentiate (Consumer Reports, 2020). Furthermore, there are indications that systems providing gaze-based attentiveness requests should increase mode awareness or decrease mode confusion (Kurpiers et al., 2019). These assumptions were examined by means of the studies within the current project.

In total, we focused on two main constructs and six metrics (see Table 6-5). In addition, we conducted interviews and noted remarks in an experimental protocol that are used when making a contribution to a respective construct. For the metrics’ definitions, we refer to the respective study chapters.

Table 6-5: Assignment of constructs and metrics for CQ4

Construct	Metric
Knowledge-based confusion	• System understanding • Role understanding • Reported behavior related to warnings
Behavior-based confusion	• Time H-off, although L0/L1 (for L2H-off users only) • Time H-off, although L2H-on (for L2H-off users only) • Number of attempted activations of L2, although not available

6.1.4.2 Conclusions

In the following, the conclusions are presented and compared to findings in the literature.

Knowledge-based confusion

Over all studies, no significant differences between L2 functions (L2H-on and L2H-off) could be found when analyzing the questionnaire results regarding system and driver role understanding. In fact, over all studies and both L2 functions (H-on and H-off) there was a good to very good understanding of system functionality, functional limits and driver responsibilities, which is an essential prerequisite for mode awareness. However, it should be considered, that this finding might be supported by the usage of standardized manuals introducing the systems to participants in a compact, efficient and study-focused way (see Chapters 4.4 and 5). The user survey, by comparison, revealed that a majority of drivers either did not consult the function's manual at all, stressing the need to inform the driver during use, or was not aware of any situations not recommended for use of the function (see Chapter 4.3). Experienced ADAS users showed a slightly better system and role understanding when rating the correctness of statements than novice ADAS users, but without consequences for user behavior (see Study 4, Chapter 5.5). In sum, it appears that L2H-off systems are neither unclear nor more confusing than L2H-on systems.

When comparing clear-cut and multi-step L2 systems, no significant differences could be found regarding the questionnaire-based assessment of system and role understanding (see Study 2, Chapter 5.3). However, qualitative interviews revealed slight uncertainties when using multi-step systems, which include L1 functions (ACC; Study 2) and/or combine L2 system designs incorporating L2H-on and L2H-off. These descriptive results go in line with the assumption made by the Consumer Reports (2020), which states that systems with clear-cut modes, of either on or off, might increase mode awareness/decrease mode confusion. To reduce the risk of mode confusion with multi-step systems, participants actively suggested that mode changes (L2H-on to L2H-off, off to stand by, ACC to L2) and references to (necessary) driver actions should be displayed saliently and distinguishably in the HMI when several modes are combined in one system. Accordingly, it might be useful to not only prominently display the currently active mode and mode changes but also explicitly indicate which user behavior is required when changing between modes. These findings should be kept in mind when designing L2 systems.

Behavior-based confusion

Over all studies, there were little behavior-based indications for mode confusion when looking at inappropriate H-off times (H-off times when driving manually, in L1 or L2 H-on mode) and attempted activations of Level 2. Over all studies and both L2 functions (H-on and H-off), it appears that most participants were aware when to activate the L2 function and when hands-free driving was applicable. In general, no significant differences between the L2 functions (H-on and H-off) could be found, again indicating that L2H-off functions seem not more confusing than L2H-on functions.

As with subjective data, no significant differences could be found between clear-cut und multi-step L2 functions when looking at inappropriate activation attempts and hands-free driving. However, FOT data on hands-free driving, when L2H-off mode was not activated and hands-free driving was therefore not applicable, indicate that mode confusion might have occurred for some participants, especially when using L2 functions which combine L2H-on and L2H-off functions. This finding goes in line with the qualitative interviews, indicating that multi-step functions, i.e. functions integrating multiple assistance modes, might be confusing for some users.

Main conclusions

The results of the studies conducted within the current project indicate that there are no significant differences regarding mode confusion between L2H-off and L2H-on functions. For both L2 functions (H-on and H-off) a rather good to very good understanding of system functionality, functional limits and driver responsibilities was observed both regarding objective and subjective data. This is an essential prerequisite for mode awareness. The good role and system understanding might, however, been enforced by the manuals used for instructional purposes in the beginning of each study. Based on our data, the initial concern that with the introduction of L2H-off functions, drivers are no longer aware of their tasks and roles as drivers, ODD understanding, and system functional logic, can be denied.

Nevertheless, mode confusion might occur with some users when the function offers different but rather similar modes (e.g., L2H-on and L2H-off), which goes in line with literature-based findings (see Boos et al., 2020; Kurpiers et al., 2020). However, salient and distinguishable indications of currently active modes and mode changes may help to distinguish between the modes L2H-on and L2H-off, if offered in the same vehicle. This should be considered when designing L2 functions.

6.1.5 CQ5: Safety level

Documentation by N. Grabbe, K. Bengler (Lehrstuhl für Ergonomie, TU München)

6.1.5.1 Definition

There is uncertainty as to what level of safety (in terms of contributing to road safety) can be achieved by introducing L2H-off. We distinguish two major aspects: objective safety and perceived safety. Objective safety can be seen from two different perspectives: safety-I which is outcome-oriented and safety-II which is process-oriented. According to Hollnagel (2014), the safety-I perspective is based on a reductionist approach focusing on a few negative events (e.g., failures, near-crashes, accidents) and only on outcomes in the way of numbers like criticality metrics which can be seen as a narrowed and isolated view (i.e., siloed-thinking (Hollnagel, 2020)) on what is happening or how something works. Instead, the safety-II perspective follows a complexity-oriented holistic approach focusing on the frequent and positive events or the system performance as a whole process of different interacting elements/agents (driver, vehicle, environment) in a socio-technical system leading to a system outcome. Or in other words, safety-I follows an analytical approach breaking something apart and looking at the pieces individually, whereas safety-II is about synthesis putting information together to see an overall pattern of how things come together. Overall, these two different but also complementary perspectives ensure a holistic picture.

In total, we considered three constructs and sixteen metrics (see Table 6-6). In addition, we conducted interviews and noted remarks in experimental protocol that are used when making a contribution to a respective construct. For the metrics' definitions, we refer to the respective study chapters (see Chapters 4 and 5).

Table 6-6: Assignment of constructs and metrics for CQ5

Construct	Metric
Objective safety (safety-I)	• Number of accidents/incidents • Number of incident classes • Number of safety-criticality levels • Number of controllability levels (TOC-rating) <u>Criticality metrics at transitions & for incident candidates:</u> ○ TTC ○ THW ○ Longitudinal & lateral distance ○ Longitudinal & lateral acceleration
Objective safety (safety-II)	• Synthesis of study results over all five CQs
Perceived safety	• Hands-on/-off proportion (L2H-off mode) • Trust (TiA) • Acceptance (CTAM) • Preferred L2 system & L2 intention to use

6.1.5.2 Conclusions

In the following, the conclusions are presented and compared to findings in the literature.

Objective safety (Safety-I)

We observed a few accidents as a mere fraction of all those interactions in the simulator studies (Chapter 5) which were designed to test the driver performance close to controllability limits (excluding, e.g., speed limit changes in simulator study 2, see Chapter 5.3). However, there was no difference between L2H-on and L2H-off regarding accident proneness, and no systematic, group or system-related effects could be found. These accidents can rather be seen as individual coping problems. In particular, as implemented in the simulator studies (Chapter 5), the following controllability conclusions adapted to the ISO 26262:2018 with regards to the function limits as roadworks/end of lane and cut-out and the function failure as lane drift can be drawn:

Roadworks/end of lane:

- L2H-on: normally controllable (5 accidents out of 116 interactions → 95.69% successful)
- L2H-off: normally controllable (2 accidents out of 133 interactions → 98.50% successful)

Cut-out:

- L2H-on: normally controllable (1 accident out of 60 interactions → 98.33% successful)
- L2H-off: normally controllable (5 accidents out of 97 interactions → 94.85% successful)

Lane drift (contact to barrier as relevant criterion):

- L2H-on: simply controllable (0 accidents out of 77 interactions → 100% successful)
- L2H-off: normally controllable (1 accident out of 74 interactions → 98.65% successful; for 1st lane drift only: 1 accident out of 37 interactions → 97.30% successful)

No collisions occurred in any of the scenarios which were unique to the second simulator study (i.e., curve, traffic jam). In simulator study 1 (Chapter 5.2), an emergency braking maneuver (EB) was initiated in case the driver did not provide direct input within five seconds after the situation became visible. Even if the EB was not overruled by the driver, no accident could occur. Cases, in which drivers did not overrule the EB, were counted as collisions (no overruling of EB by $n_{L2H\text{-}on} = 4$ of $N = 19$ and $n_{L2H\text{-}off} = 2$ of $N = 19$ in roadworks scenario). In simulator study 3, using the same types of scenario, a function direct control request (FDCR) was issued in the roadworks and cut-out scenario, but no EB was implemented.

We can see that, for L2H-off, both function limits and function failures were normally controllable. The same can be observed for L2H-on except that the lane drift is simply controllable.

All transition types in the FOT (Chapter 4.4; driver-initiated activation/deactivation/lane change, system-initiated deactivation) were predominantly uncritical (see candidate thresholds in the

FOT, Chapter 4.4) in terms of criticality metrics (TTC, THW, longitudinal and lateral distance/acceleration) and no differences between L2H-on and L2H-off could be found. Especially, no differences in the controllability level (TOC-rating) in the case of direct control or direct control requests could be found between both systems where the driver showed good controllability in general (see FOT; Chapter 4.4). Only two driver-initiated deactivations (L2H-off) and six system-initiated deactivations (L2H-on: 2, L2H-off: 4) were rated dangerous based on objective data (non-acceptable risk). It can be seen that more dangerous events occurred for L2H-off subjects. However, we need to be aware that the sample size of L2H-on participants is considerably smaller than L2H-off participants due to limited data which is why a final comparison is difficult. Hence, on basis of a relative evaluation, there is rather no difference between both systems concerning dangerous events in the FOT.

Regarding the eight events rated dangerous for all users, the main issues have been roadworks affecting lane detection, bumps, or little lateral distance to the crash barrier. In addition for L2H-off drivers, one dangerous driver-initiated deactivation has been due to a closely cutting-in truck and one dangerous system-initiated deactivation is due to mode confusion. In most events, the driver reacted or intervened well but objectively the situations have been determined as crash relevant in terms of safety criticality.

Further, no differences between the 3-s-DMS and 5-s-DMS (Study 3, Chapter 5.4) could be observed in terms of objective criticality metrics. Further, effects with regards to first contact (experts vs. novices) could not be found. Especially in the FOT (Chapter 4.4), no differences in the number of incident candidates, safety criticality level, and controllability level (TOC-rating) between L2H-on and L2H-off could be found. Most of the incident candidates are normal driving and thus not safety-critical. Only 14 incidents exist as increased risk or crash-relevant whereas the proportion between L2H-on and L2H-off is balanced (see FOT, Chapter 4.4). Here, it is noticeable that vehicle dynamics-based longitudinal incidents predominate compared to distance-based incidents but vehicle dynamics-based lateral incidents did not occur at all. The incidents at increased risk were mainly unpleasant (acceptable risk) in terms of controllability (TOC-rating) whereas one incident was perfect. The incidents which are crash-relevant were mainly dangerous (non-acceptable risk) in terms of controllability (TOC-rating) whereas one incident was unpleasant. In particular, major issues for L2H-on are cut-in scenarios by other road users (especially in traffic jams) and major issues for L2H-off are roadworks (lane detection, bumps). Minor issues are a lane change in the traffic jam and fall short of distance in roadwork for L2H-on drivers, passing on the right, unnecessary intervention or inappropriate braking by the driver, and overtrust during a lane change for L2H-off drivers. That coincide with the fact that participants reported a lack of trust, especially in road works, and closely approaching cars.

According to the literature, L2H-off leads to poorer direct control quality compared to L2H-on (Cahour et al. 2021; Garbacik et al. 2021; Gold et al. 2013; Ishida & Itoh, 2017; Josten 2021; Josten et al. 2016; Othersen, 2016), including, amongst others, the number of incidents and the quality of direct input after transitions. Our study results contradict this statement, at least regarding the metrics analyzed within this project. In contrast, we can conclude that L2H-off shows a mainly good and safe and similar direct control performance compared to L2H-on which can be attributed to the implemented DMS as the DMS is the main difference between the L2H-on and L2H-off users as well as between our project studies and the studies in the literature. In addition, literature shows that in critical function failure scenarios in the case of L2H-off, visual attention is not sufficient for an appropriate driver response because neither reminders nor explicit instructions on function limitations and supervision responsibilities prevented 28% of drivers from crashing with their eyes on the conflict object indicating an automation expectation mismatch (Victor et al. 2018; Gustavsson et al. 2018). However, we could not find clear indications of automation expectation mismatch for L2H-off participants.

Perceived safety

We could observe a balanced trust and acceptance for both L2H-on and L2H-off indicating no over-/undertrust. The participants rated the L2H-on & L2H-off systems with a good overall intention to use. In particular, L2H-off was preferred over L2H-on as it is more comfortable and participants had a clear desire to use L2H-off but also large dispersion could be observed. Moreover, there was a tendency that EOR (L2H-off) to produce a higher feeling of safety than HOR (L2H-on). Especially, in the FOT, participants reported that L2H-off drives smoother and more stable than L2H-on.

In rain and spray, L2H-off drivers sometimes experienced problems/frequent function drops, which was perceived as annoying and could lead to mode confusion and disuse. Besides, many subjects recommended using L2H-off meaning the hands-off option only at speeds up to 130 km/h. Also, many participants in the FOT (Chapter 4.4) reported that at interchanges the vehicle sometimes does not recognize speed limits and braked strongly which may cause danger for following vehicles or at least a surprise. Effects regarding first contact use (experts vs. novices) could not be found.

Objective safety (Safety-II)

The safety-II analysis can be seen as an overall conclusion over all five CQs. In particular, we aimed to synthesize and abstract all findings with regards to the five CQs, following a general scheme by Bubb (2021) and Plavsic (2010) for structure. L2H-off participants showed a good performance in every construct analyzed within this project. The same applies to L2H-on subjects with marginal reductions in some studies in visual attention (e.g., eyes-on road proportion in Study 1), perceptual readiness and misuse (e.g., hands-off percentage and NDRT engagement in the FOT), which are nonetheless on acceptable performance levels. Overall, the performance differences between both systems are small. The main differences lie in visual perception and misuse indicating the advantage of the adapted DMS in L2H-off. It can be argued that L2H-off users show no disadvantages compared to L2H-on users but even show slight advantages. When looking at the performance from the perspective of the five CQs, we can state that:

1) The DMS fulfills its purpose and mind-off did not occur.
2) The physical disadvantage of hands-free driving is compensated by taking the hands on the steering wheel and making a decision to act simultaneously (and not sequentially; cf. hands-on time versus intervention time), transition times are not prolonged systematically in the scenarios investigated within the studies, and driver interventions are successful.
3) Misuse and disuse levels are low and not more prominent than for L2H-on (although long-term effects were not evaluable within the current setting).
4) Mode confusion was not observed overall or in comparison to L2H-on.
5) The safety level is high and similar to L2H-on. (A comparison of L2 driving to manual driving (L0) was not in the focus of the current project apart from Simulator Study 1, where no systematic disadvantage in case of driver-detected system limits was found.)

It should be emphasized that for the human machine interaction (HMI) both functions are rated as good overall. In the FOT (Chapter 4.4), the L2H-off function is however rated to drive more smoothly and more stable. It is considered to be more comfortable but also more complex to use than L2H-on functions. Some issues could be further identified in the FOT (Chapter 4.4) regarding external conditions of use. In rain and spray, many mode switches happened in the case of L2H-off use which increases the probability of disuse and mode confusion. This could not be observed for L2H-on functions. In terms of the infrastructure, the main issue are roadworks due to failed lane detection or bumps, and false speed limit detection at interchanges. The interaction with other road users is challenging mainly in form of closely cutting-in vehicles and in the way that both functions rather impair the traffic flow in heavy traffic due to unnatural driving behavior (slow accelerating, abrupt braking). Potential speed effects cannot be completely assessed since they were not systematically evaluated but seem to play a rather minor role in the performance as no obvious performance differences between transitions at high and low speeds could be found. In terms of functional limits and failures, for both functions, end of lane and lane drift as implemented in the simulator studies were simply or normally controllable, whereas cut-outs as implemented in the studies were difficult to control for L2H-off users. In contrast, for L2H-on as well as manual driver, the cut-outs as implemented were normally controllable but tending to be difficult to control as the difference with respect to L2H-off is close. Finally, no relevant differences between first contact users and experienced ADAS users could be identified for prototypical systems in the simulator studies.

Ultimately, we have to bear in mind that the mentioned results are only valid under the given driver, vehicle/function, and environmental characteristics. For example, the results are sensitive to the specification of the warning cascades for the HOD and DMS as well as the design of HMIs. The driver population is generally representative but has some limitations regarding the technology affinity, familiarity with L2 automation, and duration of use. This means that the results are valid in terms of drivers who are open-minded to new technologies, have no experience with L2, and use the system on a short-term base. In contrast, effects through older drivers tending to struggle with new technology, drivers with extensive L2 experience, and long-term use cannot be assessed.

Main conclusions

There is a balanced perceived safety for L2H-off – referring to the continuum of over- and undertrust – that strengthens the cognitive component of the driving task, which in turn as a precondition (information processing) and promotes the ability to safely guide the vehicle. The safety level between L2H-off and L2H-on is similar with a tendency to improve the perceived (i.e. subjective) safety in favor of L2H-off. In particular, no safety level differences exist at transitions and especially in controllability in case of a direct control. On average, both systems showed good controllability. Potential issues identified in field tests are roadworks (lane detection, bumps) and cut-in scenarios by other road users. To summarize the safety-II results, it can be concluded that the DMS offers a great benefit and difference between both systems whereas hands-on the steering wheel does not play a significant role when using an attention-based DMS. It should be pointed out that this statement has to be treated with caution in case of sudden steering events as issues related to roadworks and cut-in scenarios indicate. Thus, the function design for L2H-off should be considered carefully for scenarios requiring sudden steering. The most important and decisive aspect when designing a L2H-off system is simply that the driver understands that she/he is responsible for the driving task and is supported (not substituted) in the driving task by a function not to be distracted, mode confused, etc., in the way that possible disadvantages are compensated by the DMS and HMI. General possible disadvantages of L2 are not exacerbated by L2H-off, rather these are even mitigated due to the DMS, and in order to compare effects between L2 and manual driving we refer to literature in general. Primarily, these conclusions are only valid for short-term evaluation of L2 use.

6.1.6 References

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied ergonomics*, 68, 138-145. doi:10.1016/j.apergo.2017.11.010

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., Fitch, G. M., Russell, S. M., Green, C. A., Cullinane, B., & Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (DOT HS 812 182). https://trid.trb.org/view/1367054

Boos, A., Feldhütter, A, Schwiebacher, J., & Bengler, K. (2020). Mode Errors and Intentional Violations in Visual Monitoring of Level 2. In *23rd IEEE International Conference on Intelligent Transportation Systems,* Virtual Conference.

Bubb, H. (2021). Inauguration. In *Automotive Ergonomics* (pp. 1-23). Springer, Wiesbaden.

Cahour, B., Jean-Francois, F., & Arnaud, K. (2021). Driving a Partially Automated Car with the Hands On or Off the Steering Wheel: Users' Subjective Experiences. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Springer eBook Collection: Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021): Volume III: Sector Based Ergonomics* 1st ed., Vol. 221, (pp.583–592). Springer International Publishing; Imprint Springer. doi:10.1007/978-3-030-74608-7_71

Consumer Reports (2020). *Active Driving Assistance Systems: Test Results and Design Recommendations.* Retrieved from https://data.consumerreports.org/wp-content/uploads/2020/11/consumer-reports-active-driving-assistance-systems-november-16-2020.pdf

Feldhütter, A., Härtwig, N., Kurpiers, C., Hernandez, J. M., & Bengler, K. (2019). Effect on mode awareness when changing from conditionally to partially automated driving. In Congress of the International Ergonomics Association (pp. 314-324). Springer, Cham.

Garbacik, N., Mastory, C., Nguyen, H., Yadav, S., Llaneras, R., & McCall, R. (2021). Lateral Controllability for Automated Driving (SAE Level 2 and Level 3 Automated Driving Systems). In *SAE Technical Paper Series, SAE Technical Paper Series* doi:10.4271/2021-01-0864

Geiser, G. (1985). Mensch-Maschine-Kommunikation im Kraftfahrzeug. *AUTOMOBILTECH Z*, 87(2).

Glaser, Y. G., Llaneras, R. E., Glaser, D. S., & Green, C. A. (2016). Relationship between driver eyes-off-road interval and hazard detection performance under automated driving (No. 2016-01-1424). *SAE Technical Paper*. doi:10.4271/2016-01-1424

Gold, C., Damböck, D., Bengler, K., & Lorenz, L. (2013). Partially automated driving as a fallback level of high automation. In 6. Tagung Fahrerassistenzsysteme.

Gustavsson, P., Victor, T. W., Johansson, J., Tivesten, E., Johansson, R., & Aust, L. (2018, October). What were they thinking? Subjective experiences associated with automation expectation mismatch. In *Proceedings of the 6th Driver Distraction and Inattention conference*, Gothenburg, Sweden (pp. 15-17).

Hollnagel, E. (2020). Synesis: The Unification of Productivity, Quality, Safety and Reliability. Routledge.

Hollnagel, E. (2014).Safety–I and safety–II: the past and future of safety management. CRC press.

Ishida, W., & Itoh, M. (2017, October). Effects of vehicle behavior for situation awareness when using a driving automation system. In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 888-893). IEEE.

International Organization for Standardization (2022) *Road vehicles— Safety of the intended functionality.* (ISO Standard 21448) Retrieved from https://www.iso.org/standard/77490.html.

International Organization for Standardization (2018a) *Road vehicles—functional safety—part 3: concept phase – annex B*. (ISO Standard 26262) Retrieved from https://www.iso.org/standard/68385.html.

Josten, J. (2021). *Nutzer und System im Nutzungskontext - Beherrschbarkeit von Teilautomation* [Dissertation, RWTH Aachen University]. Retrieved from https://publications.rwth-aachen.de/record/817292

Josten, J., Zlocki, A., & Eckstein, L. (2016). Untersuchung der Bewaeltigungsleistung des Fahrers von kurzfristig auftretenden Wiederuebernahmesituationen nach teilautomatischem freihaendigem Fahren. *FAT-Schriftenreihe*, (289).

Kim, H., Song, M., & Doerzaph, Z. (2022). Is driving automation used as intended? Real-world use of partially automated driving systems and their safety consequences. *Transportation research record*, 2676(1), 30-37.

Kurpiers, C., Biebl, B., Mejia Hernandez, J., & Raisch, F. (2020). Mode awareness and automated driving—What is it and how can it be measured? *Information*, 11(5), 277.

Llaneras, R. E., Cannon, B. R., & Green, C. A. (2017). Strategies to assist drivers in remaining attentive while under partially automated driving: Verification of human–machine interface concepts. *Transportation research record,* 2663(1), 20-26.

Llaneras, R. E., Salinger, J., & Green, C. A. (2013, June). Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway. In *Driving Assesment Conference* (Vol. 7, No. 2013). University of Iowa.

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). A model for naturalistic glance behavior around Tesla Autopilot disengagements. *Accident Analysis & Prevention*, 161, 106348.

Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment. *Transportation research part F: traffic psychology and behaviour*, *38*, 67-82.

Naujoks, F., Purucker, C., Neukum, A., Wolter, S., & Steiger, R. (2015). Controllability of partially automated driving systems–does it matter whether drivers are allowed to take their hands off the steering wheel?.*Transportation research part F: traffic psychology and behaviour*, *35*, 185-198.

Noble, A. M., Miles, M., Perez, M. A., Guo, F., & Klauer, S. G. (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. *Accident Analysis & Prevention*, *151*, 105959.

Othersen, I. (2016). *Vom Fahrer zum Denker und Teilzeitlenker: Einflussfaktoren und Gestaltungsmerkmale nutzerorientierter Interaktionskonzepte für die Überwachungsaufgabe des Fahrers im teilautomatisierten Modus* [Dissertation, Technische Universität Braunschweig]. GBV Gemeinsamer Bibliotheksverbund. Retrieved from https://link.springer.com/book/10.1007/978-3-658-15087-7

Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter? Transportation Research Part F: Traffic *Psychology and Behaviour*, 76, 14–25. doi:10.1016/j.trf.2020.10.001

Plavsic, M. (2010). *Analysis and modeling of driver behavior for assistance systems at road intersections* (Dissertation, Technische Universität München).

Solís-Marcos, I., Ahlström, C., & Kircher, K. (2018). Performance of an additional task during Level 2 automated driving: An on-road study comparing drivers with and without experience with partial automation. *Human factors*, 60(6), 778-792.

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human factors*, 60(8), 1095-1116.

6.2 L2H-off Guideline Document

Documentation by J. Hiller, L. Eckstein (Institut für Kraftfahrzeuge, RWTH Aachen University)

In the following, the second transformation of knowledge generated within this project aims at providing guidance on the design of L2H-off functions (L2H-off Guideline Document). The goal has been to create a stand-alone document comprising basic principles for the design of L2H-off functions within the context addressed by this project, similar to the European Statement of Principles on human machine interface for in-vehicle information and communication systems (ESoP).

1 Preamble to the L2H-off Guideline Document

The following document is a result of the “L2H-off” project. The project investigated five potential challenges and questions regarding adverse foreseeable driver behavior in combination with hands-free L2 driving. The project aimed at generating a reliable set of data, information, and knowledge on the investigated topics. The investigations conducted set the scope for requirements and recommendations described in this document. Amongst others, the project piloted and tested passenger cars on German and US motorways (multiple divided lanes, longitudinal traffic only). These on-road tests were accompanied by an expert study and complemented by simulator studies in similar infrastructural settings (cf. Figure 1). The primary aim of the project was the investigation of the vehicle-driver-interaction for L2H-off functions which comprise a glance-based Driver Monitoring System (DMS) and are implemented for dedicated ODDs. As L2H-off is a SAE level 2 function, the driver remains fully responsible. This holds true even if parts of the control execution of lateral and longitudinal dynamics as part of the driving task are handed over to the vehicle.

As further input to the project and to these guidelines, discussions with experts from the automotive industry and project external scientists from the U.S., Japan, Sweden, and Germany have been considered. Therefore, some of the following guidelines arise solely from the discussions and the requirements of these third parties and have not been finally validated.

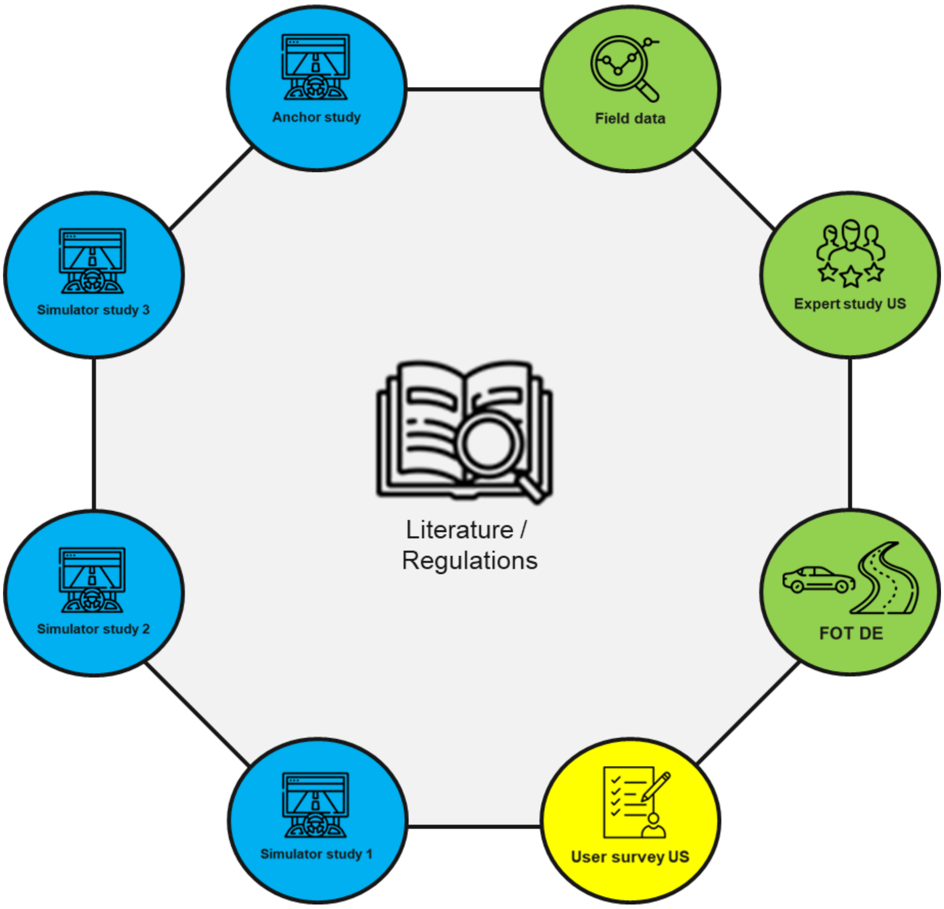


Figure 1: Overview of studies and data collections within the L2H-off project. (Icon source: Flaticon.com)

The goal of this document is to provide guidance on the design of L2H-off functions in order to achieve a sufficiently safe interaction of drivers with such functions. Similar to the European Statement of Principles on Human Machine Interface for in-vehicle information and communication systems (ESoP; Godthelp, 1998), the document proposes a structured set of guidelines, which should be followed when designing L2H-off driver assistance systems and its necessary components:

Apart from the **L2 function** itself, which controls longitudinal and lateral vehicle dynamics, a **Driver Monitoring System** (DMS) is required, which triggers **Driver Information and Warning** (DIW). If the driver is not following requests to take direct control, a **Risk Mitigation Function** (RMF) needs to be activated, which ultimately stops the vehicle. In order to avoid that the driver needs to take direct control immediately, the L2 system shall also comprise an **ODD Monitoring**, which anticipates that the L2 function is going to leave its Operational Design Domain (ODD) in the near future, allowing the DIW to provide timely information and warnings requesting the driver to take direct control. Moreover, a L2 function should always be accompanied by a **Collision Mitigation System** (CMS) in order to further reduce the probability of a collision.

The focus of this document is on the L2H-off function and the interaction between driver and L2H-off function as far as investigated within the project. Assumptions on other relevant parts of the L2H-off function or vehicle are described to the extent to which they are necessary to supplement and contextualize the project results (e.g. the ODD). Where necessary for the description of the L2H-off function, interactions between the driver and other and/or underlying functions are described. Unless specified otherwise, the fault-free operation of the L2H-off function and its elements is assumed.

Within the following document, each guideline (derived, e.g., from project results or the current state of the art) is always accompanied by an explanatory text with the aim of providing more context and meaning to the guideline itself. Where appropriate, references to literature, regulations or the project results themselves are given. The project results themselves can be found in the project’s final report and are linked here to this document via chapters and keywords, where feasible. Especially the project results should not always be seen as final confirmation. The empirical evidence may also serve as negative affirmation of a stated guideline and identify limits that should be avoided by proactive function design.

2 Terms and Definitions

Adaptive Cruise Control (ACC) means an enhancement to conventional cruise control systems, which allows the subject vehicle to follow a forward vehicle at an appropriate distance by controlling the engine and/or power train and potentially the brake.

Area of Interest (AOI) means a predefined area within the visual scene. (ISO 15007, 3.1.1)

Collision Mitigation System (CMS) aims at reducing the consequences of an accident by automatically reducing the speed and eventually other actions.

Direct control means that the driver executes the longitudinal and or lateral control.

Direct Control Request (DCR) means a request to the driver to take direct control for either only the lateral or both, the lateral and the longitudinal control of the vehicle.

DMS Direct Control Request (DDCR) means a DCR triggered by the DMS if the driver fails to follow a Hands-on-request.

Function Direct Control Request (FDCR) means a DCR triggered by the L2H-off function due to e.g. approaching ODD limitations.

Driver Information and Warning (DIW) means the central component of communication from the L2H-off function towards the driver.

Driver Monitoring System (DMS) means a technical solution which aims at monitoring the driver's behavior. This typically includes, among others, the driver's alertness as well as the monitoring of specific activities of the driver.

Dynamic Driving Task (DDT) means all of the real-time operational and tactical functions or tasks required to operate a vehicle in on-road traffic, excluding the strategic functions such as trip scheduling and selection of destinations and waypoints. The DDT may be executed either by a function or by the driver. (adapted from SAE J3016 2018, 3.13)

DDT Fallback means the response by the driver to either perform the DDT or achieve a minimal risk condition after occurrence of a DDT performance-relevant system failure(s) or upon operational design domain (ODD) exit. For a SAE L2 function, this always lies with the driver. (adapted from SAE J3016 2018, 3.14)

Eyes-On Request (EOR) means a request to the driver to return the eye focus (and thereby the visual attention) towards the areas relevant to perform the DDT.

Hands-On Request (HOR) means a request to take hold of the steering control.

Instrument Cluster (IC) means the cluster of displays and instruments usually located directly in front of the driver.

Monitoring means the activities and/or automated routines that accomplish real-time roadway environmental object and event detection, recognition, classification, and response preparation (excluding actual response), as needed to operate a vehicle. (SAE J3016 2018, 3.18.2)

Non-Driving-Related Task (NDRT) means all kinds of tasks or activities which are not directly related to the primary and secondary driving task (e.g. playing a game, receiving or sending text messages on a phone, etc.).

Object and Event Detection and Response (OEDR) means the subtasks of the DDT that include monitoring of the driving environment (detecting, recognizing, and classifying objects and events and preparing to respond as needed) and executing an appropriate response to such objects and events (i.e., as needed to complete the DDT and/or DDT fallback). For a SAE L2 function, the responsibility for the complete OEDR lies with the driver. (SAE J3016 2018, 3.20)

Operational Design Domain (ODD) means conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics. (SAE J3016 2018, 3.22)

Override of the Function by the driver means a situation when the driver provides an input to a control which has priority over the longitudinal or lateral control by the function, while the function is still active.

Overrule of the Function by the driver means the driver takes direct control of the lateral / longitudinal and thereby deactivates the lateral / longitudinal control.

Risk Mitigation Function (RMF) is activated, if driver information and warnings are not successful to motivate the driver to take direct control of the vehicle.

Transition means any change in direct control between the driver and the vehicle for the longitudinal and lateral control.

Visual Attention is used as a substitute for the current focus of the driver and means the area on which the gaze of the driver is on.

3 Guidelines for Level 2 Functions with Driver Monitoring System

3.1 General Assumptions on the Design of L2H-off Functions

3.1.1 The task of monitoring traffic and environment as well as identification and execution of appropriate actions remains with the driver.

- **Explanation:** Given that L2H-off constitutes a driver assistance system classified as SAE L2 function, the responsibility for the complete object and event detection and response (OEDR) remains with the driver.
- **References:** SAE J3016, 2018

3.1.2 Any vehicle equipped with the L2H-off function shall be able to detect imminent collisions with lead vehicles and perform appropriate measures to minimize risk.

- **Explanation:** Collisions are to be avoided by supporting the driver in performing appropriate measures. For example, a vehicle can be equipped with a collision mitigation system (CMS) (e.g. UN ECE R131 / R152, EU Regulation 2019/2144 (6.2)), which will brake the vehicle in case of an imminent collision.
- **References:** UN ECE R131 / R152, EU Regulation 2019/2144 (6.2)

3.1.3 The L2H-off function shall be able to

I. Monitor the operational design Domain (ODD) (cf. 3.2 Operational Design Domain, 3.6 ODD Monitoring),
II. Control the longitudinal and lateral dynamics (cf. 3.3 L2H-off Function),
III. Monitor the attention of the driver (cf. 3.5 Driver Monitoring) and
IV. Inform and warn the driver (cf. 3.7 Driver Information and Warning).

- **Explanation:** Within this document it is assumed that the above four components are part of a L2H-off function. To avoid the necessity for the driver to take direct control unexpectedly, the driver needs to be informed in advance to take direct control in an appropriate and safe manner. Therefore, the L2H-off function needs to monitor and evaluate whether the function's ODD is sufficient to cope with the upcoming driving task. For example, if the L2H-off function is not capable of driving through construction zones, the ODD Monitoring has to identify in advance that the ODD of the function is going to be exceeded when entering the construction zone. Since the driver is relieved from executing direct control by a SAE level 2 function, the driver may be inclined to direct the attention towards non-driving related tasks (NDRT). Since the driver is nevertheless responsible for driving, the driver's attentiveness needs to be monitored to initiate appropriate measures to direct the driver's attention by a driver monitoring system (DMS). If the DMS detects that the driver is not sufficiently attentive, suitable information and warnings are presented by the DIW in to direct the attention appropriately.

3.1.4 The L2H-off function should be equipped with a component to log the status of the function.

- **Explanation:** In cases or events, in which the safety of the driver or the surrounding traffic may be endangered, the status of the function needs to additionally be recorded (e.g. EU Regulation 2019/2144 (6.2), UN ECE R160) for further analysis and to facilitate a review of the requirements for L2 functions.
- **Reference:** EU Regulation 2019/2144 (6.2), UN ECE R160

3.1.5 The effectiveness of the L2H-off function should not be adversely affected by cyber-attacks, cyber threats, and vulnerabilities.

- **Explanation:** It is important to ensure that non-authorized third parties cannot tamper with the system. The effectiveness of the security measures can be demonstrated e.g. by compliance with UN Regulation No 155.
- **References:** UN ECE R155

3.1.6 If the L2H-off function performs software updates, the effectiveness of the software update procedures and processes should be demonstrated.

- **Explanation:** As there may be necessary changes to the system (e.g. update implemented regulations or maps), it is necessary to ensure that this process is without added risks e.g. by compliance with UN Regulation No 156.
- **References:** UN ECE R156

3.2 Operational Design Domain

3.2.1 The intended ODD shall be restricted to lanes that are well-defined (e.g. by sufficiently visible lane markings) and that are of sufficient width.

- **Explanation:** As the function needs to keep the vehicle within the lane, the lane needs to be defined for the function (e.g. UN ECE R79 (5.6.2.3.1.3)). Temporarily insufficient lane markings may be bridged by other inputs (e.g. road boundaries, infrastructural separation, surrounding traffic, map data) as long as the course of the lane can be reliably determined.
- **Reference:** UN ECE R79 (5.6.2.3.1.3), FOT DE (System deactivations)

3.2.2 The intended ODD of the L2H-off function shall be restricted to weather conditions that would allow the recognition of the environment and surrounding traffic participants in sufficient quality and detail by the function as well as by the driver.

- **Explanation:** Environmental conditions such as heavy rain, snow and sun glare are excluded from the ODD if they do not allow a sufficient detection of the surrounding traffic. Since the function needs to interact with the surrounding traffic in a safe way, the surrounding traffic and the environment need to be detected. If the driver needs to take direct control for whatever reason, he or she also needs to be able to perceive the environment and traffic in sufficient quality. Therefore, the availability of the L2H-off function needs to be restricted to those environmental conditions, which can be handled safely by the function as well the driver.

- **References:** Expert Study US

3.2.3 The L2H-off function shall include measures to verify that it operates within its ODD (cf. 3.6 ODD Monitoring).

- **Explanation:** Since the driver may want to activate the L2H-off function while operating the vehicle outside of the ODD of the L2H-off function the function always needs to versify, whether the vehicle is operated inside or outside the ODD. Based on suitable measures, the driver can be informed about potential OEDR limits of the function.

3.3 L2H-off Function

3.3.1 The activated L2H-off function shall detect the distance and the change in distance to a potential lead vehicle and adapt its speed to keep a sufficient distance.

- **Explanation:** A sufficient distance to the lead vehicle is defined by the minimum viable distance specified within e.g. ISO 15622 (6.2.3.1).
- **References:** ISO 15622 (6.2.3.1)

3.3.2 The activated L2H-off function shall aim to center the vehicle in the lane unless a different position is deemed reasonable due to the situation or resulting from driver input (e.g. when another vehicle is driving close beside).

- **Explanation:** For most situations, keeping the vehicle in the center of the lane is desirable (e.g. UN ECE R79 (5.6.2.1.1)). According to literature (cf. Gold et al., 2013, Josten 2021, Kerschbaum 2018) and in contrast to other lateral support functions, an additional 300 ms may be necessary for the driver in a hands-off usage situation to take hold of the steering control in time-critical situations when driving hands-free up to that point. If the vehicle is kept in the center of the lane, this gives the driver more time for the detection and correction of steering failures or for taking direct control in case of functional limits. In certain situations (e.g. close-by truck) it may be reasonable to deviate from the lane center to generate more distance between the vehicle and the other vehicle in lateral distance and thereby provide more time for the driver to intervene if necessary. Furthermore, if the vehicle's position within the lane is not controlled in an expectable manner by the L2H-off function, the driver may need more time to detect potential effects of functional failures like unintended lane drifts.
- **References:** UN ECE R79 (5.6.2.1.1), FOT DE (System-initiated deactivations), Anchor Study (Lane drift scenario), Gold et al., 2013, Kerschbaum 2018, Josten 2021

3.3.3 The activated L2H-off function should support the driver to comply with the traffic rules relating to the DDT in the respective country of operation (e.g. by providing speed limit information).

- **Explanation:** The driver remains ultimately responsible for the DDT and needs to monitor, whether the actual longitudinal and lateral control of the vehicle by the L2H-off function is in accordance with the locally relevant traffic rules (e.g. EU Regulation 2019/2144 (6.2), StVO), which may also change temporarily. A possible support by the function can be the display of changed speed limits.
- **References:** EU Regulation 2019/2144 (6.2), StVO, Study 2 (ACC usage)

3.3.4 The activated L2H-off function should monitor its ODD in advance, to adapt longitudinal and lateral control in response to known or detected ODD limitations and to eventually instruct the driver to take direct control within a sufficient timespan.

- **Explanation:** Since the L2H-off function typically has limitations in terms of perception, cognition and control documented by its ODD, the function needs to monitor in advance, whether one or more of these limitations are going to be reached. ODD monitoring recognizes potential limitations to adapt the behavior of the function accordingly (e.g. ISO 15622 (6.2.3) for reduced capabilities) and eventually to hand over the driving task to the driver. Even though the ODD is monitored, the OEDR and especially the DDT fallback primarily remains with the driver. As an example, the L2H-off function recognizes impaired perception by adverse weather conditions and consequently reduces the set speed while informing the driver. Another example is, that the L2H-off function recognizes an upcoming curve and reduces the set speed to safely guide the vehicle through the curve (e.g. ISO 21717 (6.4)).
- **References:** EU Regulation 2019/2144 (6.2), ISO 15622 (6.2.3), ISO 21717 (6.4)

3.3.5 The L2H-off function should include the option for the driver to adjust the set speed and may include the option to adjust the minimum longitudinal distance to the lead vehicle following the provisions by e.g. ISO 15622.

- **Explanation:** The behavior of the L2H-off function in longitudinal direction can build upon the provisions set by e.g. ISO 15622 (6.3). Drivers switching from ACC would be confused if the function would behave differently. Therefore, the options for setting the speed and the distance are adopted from there.
- **References:** ISO 15622 (6.3), Simulator Study 2 (Video-based interview)

3.3.6 The L2H-off function shall become active only if all of the following conditions are met:

I. The function has been activated by the driver by a deliberate action,
II. The vehicle is within the ODD of the function,
III. The function is available to the driver and
IV. The driver is classified by the driver monitoring as attentive.

- **Explanation:** The vehicle and the driver need to be in clearly defined states (e.g. UN ECE R79 (5.6.4.2.3)) which are suitable for a safe transition when the function becomes active. If no deliberate action of the driver was necessary for an activation of the L2H-off function, this may lead to mode confusion.
- **References:** UN ECE R79 (5.6.4.2.3), State of the Art Review (Activation principles), Expert Study US (Interview data)

3.3.7 The L2H-off function shall inform the driver of the driver's responsibilities and necessary actions in a timely and clearly perceptible fashion at least upon first activation after each ignition cycle (cf. 3.7 Driver Information and Warning).

- **Explanation:** With the availability of several driver assistance systems within the vehicle it may be confusing for the driver to distinguish between the L2H-off function and other functions. Therefore, the awareness for the responsibilities when operating individual functions or the vehicle might dwindle. If the driver is informed about responsibilities when using the function, the driver is more likely to stay attentive for the DDT and perform necessary actions.
- **Reference:** User Survey US (Reading of owner's manual), CQ4 (Knowledge-based confusion)

3.3.8 The L2H-off function shall no longer be active if the conditions under 3.3.6 are no longer valid for a sustained period (cf. 3.4 Transitions) and inform the driver upon such a deactivation.

- **Explanation:** As the conditions under 3.3.6 provide the necessary prerequisites for the operation of the L2H-off function, the function needs to deactivate, if one or more conditions are no longer fulfilled. Since some changes might not be directly obvious to the driver, the L2H-off needs to inform the driver about the change in function state.

3.4 Transitions

3.4.1 It shall be possible for the driver to deactivate the L2H-off function at any time.

- **Explanation:** Since the driver remains the responsible entity during L2H-off driving, the driver always needs to have the option to deactivate the function using a designated control element or primary vehicle controls (e.g. UN ECE R79 (5.6.4.2.4)).
- **Reference:** UN ECE R79 (5.6.4.2.4)

3.4.2 The driver shall be able to deactivate the L2H-off function by braking. Moreover, it may be to deactivate the L2H-off function by overruling steering or acceleration applied by the function itself.

- **Explanation:** As the responsibility stays with the driver, actions by the driver need to take precedence over the L2H-off function (e.g. ISO 15622 (6.3.1.2)). If the brakes are applied by the driver, the function deactivates and informs the driver accordingly. Deactivation due to a significant steering or acceleration action by the driver can be considered as a valid means to deactivate the function (US Expert Study).

- **References:** ISO 15622 (6.3.1.2), US Expert Study

3.4.3 The driver may be able to temporarily override the L2H-off function by accelerating or steering whereby the L2H-off function remains active.

- **Explanation:** Under some circumstances, the driver might want to perform small corrections of the actions / control provided by the function (e.g. UN ECE R79 (5.6.4.3), ISO 15622 (6.3.1.4)). These corrections need not necessarily lead to a deactivation of the function.
- **References:** ISO 15622 (6.3.1.4), UN ECE R79 (5.6.4.3)

3.4.4 A sustained override may lead to a deactivation of the L2H-off function.

- **Explanation:** If the driver continuously overrides the function's actions, the driver clearly wants to control the dynamics of the vehicle, directly. Therefore, the function may be deactivated in combination with an appropriate information of the driver.

3.4.5 The L2H-off function shall be deactivated if the monitoring of the driver (cf. 3.5 Driver Monitoring) or the ODD monitoring (cf. 3.6 ODD Monitoring) request a deactivation.

- **Explanation:** Both the attentiveness of the driver as well as the vehicle being within the limits of the ODD are necessary requirements for the operation of the L2H-off function. If any of those is no longer fulfilled, the respective monitoring system will request a transition to deactivate the function.

3.4.6 A risk mitigation function (cf. 3.8 Risk Mitigation Function) shall become active in the case the driver did not respond to a request to take direct control.

- **Explanation:** The L2H-off function is activated under the provision that the driver is attentive and ready to take direct control. If this is not the case, the function needs to transfer the vehicle into a safe state to avoid harm to the driver and the surrounding traffic.
- **Reference:** US Expert Study, Simulator Study 4 (Incident due to failure to act on FDCR)

3.5 Driver Monitoring

3.5.1 The Driver Monitoring System (DMS) shall monitor the driver's visual attention to ensure that the driver is sufficiently attentive.

- **Explanation:** Given that the OEDR remains a permanent task of the driver, the driver needs to sufficiently monitor the environment and the behavior of the L2H-off function and the vehicle. Visual attention is a necessary precondition for the assessment whether a response or action by the driver is needed. It is assumed, that the ability of the driver to perform the DDT fallback (e.g. ISO 21959 (7)) can best be assessed by monitoring the visual attention of the driver. To achieve this goal, the DMS triggers the driver information and warning (DIW).

- **References:** ISO 21959 (7)

3.5.2 The DMS may confirm that the driver has an appropriate driving posture to get in contact and operate the primary controls if necessary.

- **Explanation:** The driver needs to be able to take direct control of the vehicle to execute the DDT at any time. This is primarily to be assured by confirming the attentiveness of the driver. Appropriate driving posture and appropriate proximity of hands and feet to primary controls of the vehicle can be further criteria to confirm the readiness for execution of the DDT. Additionally, using the fastened safety belt can help to confirm that the driver is in the seat. However, moving the seat back (i.e. leaving an adequate driving position) and thereby compromising perception and execution of the DDT is considered abuse (e.g. ISO 21448 (7.3.4)).
- **References:** SAE J3016, 2018, ISO 21448 (7.3.4), Simulator Study 1 (Attention ratio), User Survey US (Impact of alerts)

3.5.3 The driver shall be deemed attentive if the driver's visual attention is directed towards the OEDR-relevant areas.

- **Explanation**: Since the OEDR is with the driver, the visual attention of the driver needs to be directed towards the areas relevant for the OEDR. If the driver directs the visual attention towards the road, safety-relevant information is sufficiently likely to be perceived by the driver.
- **References**: User Survey US (Misuse strategies)

3.5.4 The direction of the driver's visual attention shall be defined by measuring the driver's head posture and/or eye gaze.

- **Explanation**: To perform the OEDR the driver needs to direct the visual attention towards the relevant areas of the road scene by either turning the head towards that area or at least directing the eye gaze there. The visual attention of the driver towards the OEDR-relevant areas can be sufficiently measured using either the head posture and/or the eye gaze of the driver.

3.5.5 The DMS shall include the detection of the driver's eye biometrics to assess if the eyes of the driver are not closed, to ensure that the driver is awake.

Explanation: The DMS needs to ensure that the driver's visual attention is not only turned towards the road but that the driver is also awake / not drowsy. This can be sufficiently assured using a combined detection of head posture, eye lid closure and the driver's eye blinking (e.g. ISO 21959 (6.3)).

- **References**: ISO 21959 (6.3), User Survey US (Misuse strategies, new activities with L2)

3.5.6 The DMS shall detect whether the driver takes hold of the steering control. Additionally, the ability of the driver to perform the OEDR may be assessed by detecting the hand posture.

- **Explanation:** If the driver takes hold of the steering control, this can be deemed as physical readiness of the driver to take direct control. Additionally, if the driver's hands are not near the steering control or interacting with other elements of the vehicle, the time necessary to react to requests by the function is likely prolonged and the function may need to take this into account.
- **References:** FOT DE (Hand posture analysis)

3.5.7 The DMS should remind and warn the driver of the monitoring task via the driver information and warning (DIW) system if the driver is deemed inattentive (cf. 3.7 Driver Information and Warning).

- **Explanation:** The visual attention of the driver on OEDR-relevant AOIs is necessary. Therefore, the DIW notifies the driver if this condition is detected as not fulfilled (cf. Kurpiers et al., 2019). Depending on the degree of the inattentiveness, a warning further encourages the driver to return the attention to OEDR-relevant areas (cf. Kurpiers et al., 2019, Llaneras et al., 2017). Therefore, timely reminders to keep the visual attention towards OEDR-relevant areas lead to a higher overall attention of the driver.
- **References:** FOT DE (Number of warnings), Simulator Study 2 (Number of warnings), Simulator Study 3 (Number of warnings), Kurpiers et al., 2019, Llaneras et al., 2017

3.5.8 The driver shall be deemed attentive again after a detected aversion of visual attention from the road if visual attention is re-directed towards the road for at least 100 – 800 milliseconds depending on the situation. Gaze patterns should additionally be considered to define sufficient attentiveness.

- **Explanation:** The driver's visual attention must be directed to the road for at least 100 - 800 milliseconds (cf. ISO 15007 (3.1.4, minimum fixation duration), Schweigert, 2003) to allow the driver to perceive all relevant information necessary to perform the OEDR. The exact timing strongly depends on the complexity of the situation and the necessity of the driver to (re-)assess the situation. The gaze behavior of drivers can be divided into scanning and processing (Schweigert, 2003). Scanning can be seen as the refreshing of existing knowledge and processing as the generation of new knowledge. If the driver scans the OEDR-relevant areas in a regular interval, there is no need for re-processing the complete OEDR-relevant areas when returning visual attention towards the road. If the environment changes (e.g. new or more traffic participants, etc.), the driver needs additional time to process these new elements. Therefore, taking the gaze patterns including fixations and saccades of the driver into account and

differentiating between scanning and processing is an important part of the determination of the sufficient time span for deeming the driver attentive again.
- **References:** ISO 15007 (3.1.4), Schweigert, 2003, Anchor Study (Incident during lane drift)

3.5.9 The DMS shall provide suitable data necessary for the DIW to be able to inform and warn the driver in order to influence the drivers behavior (cf. 3.7 Driver Information and Warning).

- **Explanation:** The DMS has the task of monitoring and classifying the driver's behavior and attention whereas the DIW issues information and warnings to influence the attentiveness and behavior of the driver. The DIW cannot inform or warn the driver in case of inattentiveness, if the DMS does not provide data describing the state of the driver (e.g. attentiveness).
- **References:** FOT DE (System-initiated deactivations), Simulator Study 1 (System-initiated deactivations)

3.6 ODD Monitoring

3.6.1 The ODD monitoring shall enable the L2H-off function to adapt its behavior to changed ODD conditions if necessary.

- **Explanation:** If the ODD monitoring anticipates that the current parameters of the L2H-off function will lead to leaving the ODD in the near future, it may provide data to the function allowing an adaptation e.g. of the set speed in order not to exceed the maximum lateral acceleration defined by the ODD. If the driver attempts to activate the function, it may be sufficient for the ODD monitoring to simply inform the function whether the current driving state is within the ODD, so the function may become active. If the function is active and may not be capable to handle an upcoming driving task, the driver is informed well before the function hands back the execution of the DDT to the driver. Therefore, one solution is for the ODD monitoring to analyze data provided by vehicle sensors looking ahead in combination with map data, e.g. describing the curvature of the road ahead.

3.6.2 The ODD monitoring shall provide respective data of relevant and known disturbances to the L2H-off function and the DIW.

- **Explanation:** As the L2H-off function needs to either hand-over the execution of the driving task to the driver before leaving the ODD or to adapt its behavior (e.g. set speed) in order not to leave the ODD, the ODD monitoring needs to provide relevant data to the function. In general, the ODD monitoring compares the requirements from the detected DDT with the capabilities of the function defined by the ODD. If the ODD monitoring predicts a mismatch between requirements of the driving task and the ODD, the function needs to change its behavior or switch off. For example, using an underlying map, the ODD monitoring can anticipate upcoming changes in the driving task and check, whether the required functional performance might exceed the operational design domain. Also, known disturbances, such as construction sites, may be available to the ODD monitoring via external sources and can be used for the prediction of the availability of the function.

3.7 Driver Information and Warning

3.7.1 The DIW shall at least consist of

I. The component responsible for timing and displaying information and warnings,
II. The display of function status for example within the IC and
III. The component for emitting acoustic warnings.

- **Explanation:** To fully redundantly inform the driver about the responsibilities and necessary actions, at least three components are necessary. The first component is dedicated to the timing of information and warnings and displays these visually. The status of the function (i.e. off, stand-by and active plus parameters like set speed), is visually displayed by a second component in order to avoid that the system status might be temporarily overlayed by other information and warnings. Thirdly, a component to emit acoustic warnings is installed, since the multimodal presentation of warnings is more effective than a mere visual display.

3.7.2 Using the DIW, the L2H-off function shall inform the driver about

I. Its current status (off, stand-by and active plus parameters like set speed),
II. Driver-initiated changes of functional parameters (e.g. set speed),
III. The responsibilities and necessary actions by the driver and
IV. Any changes in the relevant states of the function.

- **Explanation:** As the driver is responsible for OEDR, the driver needs to be clearly informed about the function state (e.g. UN ECE R79 (5.6.2.2.1)) and any changes in relevant states, which may be driver initiated or function driven, e.g. automatic reengagement after overrides.
- **References:** UN ECE R79 (5.6.2.2.1), Expert Study US (Interview data), FOT DE (Driver-initiated activations)

3.7.3 Depending on the level of urgency of the information or warning, the L2H-off function shall use different modalities to inform and/or influence the driver.

- **Explanation:** The different levels of urgency need to be clearly distinguishable (cf. Kurpiers et al., 2019, e.g. UN ECE R79 (5.6.2.2.5)) by their design and implementation. As an example, an implementation may use the visual modality for a first request, add acoustic cues as part of an elevated request character and may employ short braking pulses to catch the driver’s attention at a final request stage. If the modalities of the warnings change with urgency level, it is easier to make the driver aware of the increasing urgency.
- **References:** UN ECE R79 (5.6.2.2.5), Kurpiers et al., 2019

3.7.4 Warnings to the driver should be displayed using optical displays within or close to the normal line of sight of the driver towards the road (e.g. IC or head-up-display). Since the driver is not looking towards displays, the visual presentation of warnings shall be accompanied by another modality, e.g. an acoustic warning.

- **Explanation:** The driver needs to be able to perceive the warning while focusing on the primary field of attention (e.g. UN ECE R79 (5.4.1.2)). This requires the use of an additional modality complementing the visually presented warning.
- **References:** UN ECE R79 (5.4.1.2)

3.7.5 If an acoustic signal is used as a warning modality and not as addition to a visually displayed warning, it shall be easily recognizable by the driver.

- **Explanation:** As the L2H-off function needs the driver as fallback for the DDT, the driver needs to be clearly aware of a potential need to react. If the warning uses an easily recognizable signal (e.g. UN ECE R79 (5.4.1.3)), the driver will be aware of the necessary reaction without the potential of confusion with other warnings or information.
- **References:** UN ECE R79 (5.4.1.3)

3.7.6 Requests for visual attention, i.e. if the driver is deemed inattentive as per 3.5.3, shall be invoked latest after five seconds (eyes-on request, EOR). This request shall be a visual warning in combination with at least one other modality unless it can be ensured that the driver has recognized the visual warning.

- **Explanation:** As the driver is responsible for OEDR, prolonged inattentiveness is not appropriate (cf. Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, e.g. UN ECE R79 (5.6.2.2.5)). Therefore, a request for visual attention, i.e. directing the eyes towards the driving scene (EOR) is displayed. In order to increase the likelihood of perception by the driver, a second modality is used complementing the visual presentation. The time span of five seconds has been shown to be as effective as three seconds. Not reacting to the EOR on purpose is deemed misuse of the function.
- **References:** UN ECE R79 (5.6.2.2.5), Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, State of the Art Review (System design approaches), Simulator Study 2 (Reaction time to EOR), Simulator Study 3 (Reaction time to EOR)

3.7.7 If the driver is deemed attentive by the DMS after the EOR was issued, the EOR shall be considered as confirmed.

- **Explanation:** In this case, the DIW successfully informed the driver about the necessary action and the driver is attentive again (cf. 3.5.8). Therefore, there is no need for further requests at this point (e.g. UN ECE R79 (5.6.2.2.5)).
- **References:** UN ECE R79 (5.6.2.2.5)

3.7.8 If the driver does not react to the EOR for at most three seconds after it was issued, the driver should be requested to return attention towards the driving task and to take hold of the steering control (hands-on request, HOR). This shall be an acoustic warning together with a visual warning.

- **Explanation:** Requesting the hands on the steering control is an amplified means of ensuring driver reaction and raising the awareness of the responsibility for the OEDR (cf. Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, e.g. UN ECE R79 (5.6.2.2.5)). The hands-on request can be used as further measure to enforce the shift of attention by the driver towards the driving task.
- **References:** UN ECE R79 (5.6.2.2.5), Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, State of the Art Review (System design approaches)

3.7.9 If the driver is deemed attentive again by the DMS and at least one hand of the driver is on the steering control, the HOR shall be considered confirmed.

- **Explanation:** With at least one hand on the steering control and the DMS stating that the driver is attentive, the driver is deemed to be within a state that allows for the continued operation of the L2H-off function. Since the continued operation requires the attentiveness of the driver, the DMS checks, whether the driver has looked back to the road scene. Requiring the driver to take hold of the steering control can be seen as further measure to educate the driver to avoid prolonged intervals of inattentiveness (e.g. UN ECE R79 (5.6.2.2.5)).
- **References:** UN ECE R79 (5.6.2.2.5)

3.7.10 If the driver stays inattentive even after the HOR, a DMS-initiated direct control request (DDCR) shall be emitted no later than five seconds after the HOR.

- **Explanation:** The DDCR constitutes the final DMS warning level. It asks the driver to take direct control of the driving task, as the driver stayed inattentive despite previous requests from the DMS via the DIW. As the responsibility for OEDR lies with the driver, a prolonged inattentiveness even after (multiple) request(s) needs to be sanctioned by the system (cf. Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, e.g. UN ECE R79 (5.6.2.2.5)).
- **References:** UN ECE R79 (5.6.2.2.5), Blanco et al., 2015, Llaneras et al., 2017, Kurpiers et al., 2019, State of the Art Review (System design approaches), Simulator Study 1 (L2H-on vs. L2H-off)

3.7.11 If the DDCR is answered by the driver by taking direct control, which requires an operation of any primary control element, the L2H-off function shall be deactivated.

- **Explanation:** The emitted DDCR needs to be terminated as soon as the driver takes appropriate actions (e.g. UN ECE R79 (5.6.2.2.5)).
- **References:** UN ECE R79 (5.6.2.2.5)

3.7.12 If multiple DDCRs have been issued over a sufficiently short period of time, the L2H-off function should be made unavailable to the driver for a sustained period of time.

- **Explanation:** Multiple DDCRs within a sufficiently short period of time clearly indicate that the driver is not vigilant enough to perform the OEDR. Therefore, activating the function is no longer possible to prevent misuse of the function.
- **Reference:** Expert Study US (State of the Art Analysis)

3.7.13 If the L2H-off function receives notice from ODD monitoring about an upcoming ODD limitation in a timely manner and the parameters of the L2H-off function cannot be adjusted accordingly to continue its operation, it should issue a request to the driver to take direct control (function direct control request, FDCR). If the function receives notice of the ODD limitation well in advance, the FDCR should be issued no later than five to six seconds in advance (cf. 3.6).

- **Explanation:** If the ODD monitoring detects that the upcoming DDT cannot be covered by the L2H-off function, the driver needs to take direct control of the driving task. This happens, while the function is still within the ODD. Depending on the timespan for this transition, the modalities used can adequately convey the urgency of the request. If the limitation is recognized early enough, the DIW may first issue a HOR to prepare the driver and raise awareness. According to literature (Pipkorn et al., 2021), successful driver interventions can be assumed even in higher automation levels within five to six seconds after an FDCR. Considering the lower limit, a timing of 2.7 seconds proved insufficient in some cases within our studies (cf. Simulator Study 3 (Reaction times to FDCR)). However, these considerations are only valid if an FDCR was issued, as five seconds were proven as too short for driver-detected interventions (i.e. without an FDCR) (cf. Simulator Study 1 (System limits: Roadworks)). This further highlights the fact that even a late warning still increases the number of successful interventions by drivers.
- **References:** Pipkorn et al., 2021, Simulator Study 1 (System limit: Roadworks), Simulator Study 3 (Reaction times to FDCR)

3.7.14 If the L2H-off function is deactivated without prior warning to the driver, this deactivation shall be clearly and well perceptibly communicated to the driver by issuing an FDCR.

- **Explanation:** While the ODD monitoring aims at handing over the execution of the DDT within several seconds, there may be sudden or unplanned events like a suddenly changing ODD, which require the driver to immediately take direct control. Therefore, this deactivation needs to be clearly and well perceivably communicated to the driver, e.g. by issuing a multimodal FDCR.
- **References:** Simulator Study 1 (System limit: Cut-in), Simulator Study 3 (Reaction times to FDCR)

3.7.15 An unanswered DCR must result in further measures to reduce the risk and consequences of potential accidents. These further measures should ideally also raise the driver's attention e.g. by applying the vehicle's brakes. As a last resort, the risk mitigation function (cf. 3.8 Risk Mitigation Function) shall be triggered.

- **Explanation:** Since a DCR (DDCR or FDCR) clearly asks the driver to take direct control, an unanswered DCR indicates that the driver is not able to monitor or execute the driving task. Therefore, the speed of the vehicle can be reduced to mitigate the risk and consequences of a potential accident. Reducing speed by braking may also serve as vestibular warning which ideally raises awareness of the driver for immediate action. If the driver appears to be unable to serve as fallback, the L2H-off function cannot continue operating and the risk mitigation function is triggered.

3.7.16 To ease the understanding of the L2H-off function and its limits by the driver, the DIW may display reasons for DCRs.

- **Explanation:** To raise awareness of the responsibilities of the driver but also for the limits of the L2H-off function, the DIW may display reasons for issued DCRs. This can also be done solely upon request by the driver. Either way, educating the driver on the reasons for DDCRs leads to a reduced number of DCRs in the future.
- **Reference:** Expert Study US (Mode confusion), Simulator Study 4 (ODD limit: lane end)

3.8 Risk Mitigation Function

3.8.1 The risk mitigation function (RMF) shall minimize the risks to the safety of vehicle occupants and other road users once the driver remains inattentive in spite of given warnings.

- **Explanation:** If the driver did not react to multiple warnings or a DCR, it has to be assumed that the driver may not be able to take direct control. Therefore, the RMF needs to take actions to minimize potential risks. The risk mitigation function includes an automatic deceleration of the vehicle in order to bring the vehicle to a safe stop. Moreover, the hazard warning lights are activated with the initiation of the risk mitigation function.

3.8.2 At the end of an RMF, the L2H-off function shall no longer be active.

- **Explanation:** If the RMF comes to an end without driver intervention, the driver is not available and therefore a necessary precondition of the function is not fulfilled (cf. 3.3.6). Consequently, the L2H-off function is no longer active at the end of an RMF. If the driver intervenes while the RMF is active, the continued attentiveness of the driver needs to be ensured by the DMS before the function can be activated again by the driver.

3.8.3 The driver shall be able to overrule the RMF and thereby end it by taking appropriate actions.

- **Explanation:** Since the driver is always in responsibility of the OEDR, several options need to be provided to allow for an immediate termination of the RMF. Appropriate driver actions include

I. Deactivation using a button,
II. Braking,
III. Accelerating or
IV. Steering.

4 Guideline Literature

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., Fitch, G. M., Russell, S. M., Green, C. A., Cullinane, B., & Morgan, J. F. (2015). *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts* (DOT HS 812 182). https://trid.trb.org/view/1367054

Godthelp, H., Haller, R., Harteman, F., Hallen, A., Pfafferott, I., & Stevens, A. (1998). European Statement of Principles on Human Machine Interface for In-vehicle Information and Communication Systems.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* doi:10.1177/1541931213571433

Josten, J. (2021). *User and system within the context of use – How users cope with partial automation* (PhD Thesis, RWTH Aachen University). Retrieved from https://publications.rwth-aachen.de/record/817292

Kerschbaum, P. J. (2018). *Design for Automation–The Steering Wheel in Highly Automated Cars* (Doctoral dissertation, Technische Universität München).

Kurpiers, C., Lechner, D., & Raisch, F. (2019). The influence of a gaze direction based attention request to maintain mode awareness. In *Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles*, Eindhoven.

Llaneras, R. E., Cannon, B. R., & Green, C. A. (2017). Strategies to Assist Drivers in Remaining Attentive While Under Partially Automated Driving: Verification of Human–Machine Interface Concepts. *Transportation Research Record: Journal of the Transportation Research Board*, 2663(1), 20–26. doi:10.3141/2663-03

Pipkorn, L., Victor, T., Dozza, M., & Tivesten, E. (2021). Automation Aftereffects: The Influence of Automation Duration, Test Track and Timings. *IEEE Transactions on Intelligent Transportation Systems*, *23*(5), 4746-4757.

Schweigert, M. (2003). *Fahrerblickverhalten und Nebenaufgaben* (Doctoral dissertation, Technische Universität München).

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, 60(8), 1095–1116. doi:10.1177/0018720818788164

7 Glossary

Table 7-1: Glossary for acronyms and established terms

Item name	Description
ACC	Adaptive Cruise Control
ACSF	Automatically Commanded Steering Function
AD	Automated driving
ADAS	Advanced Driver Assistance System
ADS	Automated driving system (SAE L3+)
ALKS	Automated Lane Keeping System
AoI	Area of Interest
AR	Attention Ratio / Attention Reminder
ASIL	Automotive Safety Integrity Level
AV	Automated Vehicle
CDA	Conditional Driving Automation = SAE Level 3
CID	Central Information Display
CS	Center Stack
CSF	Corrective Steering Function
CQs	Challenges and Questions
DA/DD/DL/SD	Driver initiated Activation/Deactivation/Lane change/System-initiated deactivation
DCR	Direct Control Request
DDCR	DCR after DMS warning cascade due to driver inattentiveness
DIW(S)	Driver Information and Warning (System)
DMC	Driver Monitoring Camera
DMS	Driver Monitoring System
DSM	Driver State Monitoring
EEG	Electroencephalography
EM	Emergency Maneuver
EOD	Eyes-on detection
EOG	Electrooculography
EOR	Eyes-on Request
ESF	Emergency Steering Function
ETJA	Extended Traffic Jam Assistant
ESoP	European Statement of Principles on HMI
FDCR	Function Direct Control Request
FOT	Field Operational Test
HARA	Hazard and Risk Analysis
HAZOP	Hazard and Operability Safety
HMI	Human-Machine-Interface
HOA	Hands-on Alert
HOD	Hands-on Detection
HOR	Hands-on Request
HOT	Hands-on time
HR	Heart rate
HUD	Head-up Display
IC	Instrument Cluster
ISO	International Organization for Standardization
LiDAR	Light imaging, detection and ranging
LCA / LKA	Lane Centering Assistant / Lane Keeping Assistant
LoA	Level of automation
L2H-off	See hands-free
L2H-on	See hands-on
MGOR	Mean off Road Glance Duration
MRM	Minimum Risk Maneuver
NDA	Non-disclosure agreement

NDRT / NDRA	Non-driving-related task
OEDR	Object and Event Detection and Response
ODD	Operational Design Domain
OEM	Original Equipment Manufacturer
PADS	Partially Automated in Lane Driving System
PDA	Partial Driving Automation = SAE Level 2
PEORT	Percentage eyes-off road time
PERCLOS	PERCentage eye CLOSure: percentage of time that the eyes were more than 80% closed
RAA	Richtlinie für die Anlage von Autobahnen, Guidelines for the design of German motorways.
RT	Reaction Time
RtIs	Request to Intervene
SLG	Single longest off road glance
TEORT	total eyes off road time
THW	Time Headway
TJA	Traffic Jam Assistant
TLC	Time to Line crossing
TOT	Take-over time
TTA	Time to activation
TTC	Time to Collision
TTR/I	Time To Respond (= Response Time) / Intervention Time
UX	User experience
VDA	Verband der Automobilindustrie
VRU	Vulnerable Road User

Verband der Automobilindustrie e.V. (VDA)
Behrenstraße 35, 10117 Berlin
www.vda.de
Twitter @VDA_online

Forschungsvereinigung Automobiltechnik e.V. (FAT)
Behrenstraße 35, 10117 Berlin
www.vda.de/fat

Verband der
Automobilindustrie

Forschungsvereinigung
Automobiltechnik