Compilation of suitable safety indicators for the evaluation of Human-Machine Interaction of level 3 systems

Reports of the Federal Highway Research Institute
Automotive Engineering F 153b

bast
Federal Highway Research Institute

Compilation of suitable safety indicators for the evaluation of Human-Machine Interaction of level 3 systems

by

Fei Yan, Jürgen Pichen, Sven Schmitz
Kai Sklorz, Martin Baumann

Institut für Psychologie und Pädagogik
Universität Ulm

Reports of the Federal Highway
Research Institute
Automotive Engineering F 153b

bast
Federal Highway
Research Institute

The Federal Highway Research Institute publishes the findings from its work and research in the Reports of the Federal Highway Research Institute publication series. The series consists of the following sub-categories:

A - General topics
B - Bridges and Structural Technology
F - Automotive Engineering
M - People and Safety
S - Highway Construction Technology
V - Traffic Engineering

It should be noted that the reports published under the names of the authors do not necessarily reflect the views of the publisher.

The brochures of the Reports of the Federal Highway Research Institute's publication series can be obtained directly from Fachverlag NW in der Carl Ed. Schünemann KG, Zweite Schlachtpforte 7, D-28195 Bremen, telephone: +49 421 36903-53.

Summaries of research findings and their publications are reported in brief in the Forschung kompakt information service. This publication is issued free of charge; please contact the Federal Highway Research Institute, Press and Communication office, for a copy.

The reports of the Federal Highway Research Institute (BASt) are partly available as a free download in the electronic BASt archive ELBA.
https://bast.opus.hbz-nrw.de

Imprint

Report of research project 82.0768
Zusammenstellung geeigneter Sicherheitsindikatoren für die Bewertung der MMI von Level 3 Systemen

Specialist supervisor:
André Wiggerich

Section:
Automated Driving

Publisher:
Bundesanstalt für Straßenwesen
Brüderstraße 53 | D-51427 Bergisch Gladbach
Telephone: +49 2204 430
www.bast.de | info@bast.de

Concept, editing:
Press and Communication office

Design:
MedienMélange:Kommunikation!

Printed and published by:
Fachverlag NW in der Carl Ed. Schünemann KG
Zweite Schlachtpforte 7 | D-28195 Bremen
Telephone: +49 421 36903-53 | Fax: +49 421 36903-48
www.nw-verlag.de

ISSN 0943-9307 | ISBN 978-3-95606-771-6

Bergisch Gladbach, January 2024

This report has also been published in German (report number: F 153).
It is accessible at the electronic archive ELBA under https://bast.opus.hbz-nrw.de

Abstract - Kurzfassung

Compilation of suitable safety indicators for the evaluation of Human-Machine Interaction of level 3 systems

With the aim of identifying suitable indicators and criteria for evaluating the safe human-machine interaction for SAE level 3 systems up to 60 km/h in the context of automated driving, this research project has started with a focus group interview to identify relevant publication channels and list of keywords regarding indicators for the evaluation of human- machine interaction at SAE Level 3. Based on the identified list of keywords, literature reviews have been conducted to extract relevant publications from the identified publication channels. According to the defined inclusion and exclusion criterion, 38 papers have then been selected and used for meta-analysis to study the influence of different takeover situations on takeover performances. The results of meta-analysis have indicated that drivers' takeover performances measured by the categories of takeover time, takeover quality and subjective workload are different in static and dynamic situations. After that, expert interviews have been conducted with six international experts to help interpret the results of meta-analysis and develop checklist items. In the end, 16 checklist items assigned in six categories of system requirements have been developed and can be used by international experts to evaluate the safety of the human-machine interaction of SAE Level 3 systems up to 60 km/h in production vehicles. This checklist has been further developed to an online application, which can be used as an easy-to-implement and efficient evaluation procedure in relation to the traffic safety relevant interaction quality of the system.

Zusammenstellung geeigneter Sicherheitsindikatoren für die Bewertung der Mensch-Maschine-Interaktion von Level 3 Systemen

Mit dem Ziel, geeignete Indikatoren und Kriterien für die Bewertung der sicheren Mensch-Maschine-Interaktion für SAE Level 3 Systeme bis 60 km/h im Kontext des automatisierten Fahrens zu identifizieren, wurde dieses Forschungsprojekt mit einem Fokusgruppeninterview begonnen, um relevante Publikationskanäle und eine Liste von Schlüsselwörtern bezüglich Indikatoren für die Bewertung der Mensch-Maschine- Interaktion auf SAE Level 3 zu identifizieren. Basierend auf der identifizierten Liste von Schlüsselwörtern wurde eine Literaturrecherche durchgeführt, um relevante Publikationen aus den identifizierten Publikationskanälen zu extrahieren. Anhand der definierten Ein- und Ausschlusskriterien wurden 38 Arbeiten ausgewählt und für eine Meta-Analyse verwendet, um den Einfluss verschiedener Übernahmesituationen auf die Übernahmeleistung zu untersuchen. Die Ergebnisse der Meta-Analyse haben gezeigt, dass die Übernahmeleistungen der Fahrer, gemessen an den Kategorien Übernahmezeit, Übernahmequalität und subjektive Arbeitsbeanspruchung, in statischen und dynamischen Situationen unterschiedlich sind. Anschließend wurden Experteninterviews mit sechs internationalen Experten durchgeführt, um die Ergebnisse der Metaanalyse zu interpretieren und Checklistenelemente zu entwickeln. Am Ende wurden 16 Checklistenpunkte entwickelt, die sechs Kategorien von Systemanforderungen zugeordnet sind und von internationalen Experten zur Bewertung der Sicherheit der Mensch-Maschine-Interaktion von SAE Level 3 Systemen bis zu 60 km/h in Serienfahrzeugen verwendet werden können. Diese Checkliste wurde zu einer Online-Anwendung weiterentwickelt, die als einfach zu implementierendes und effizientes Bewertungsverfahren in Bezug auf die verkehrssicherheitsrelevante Interaktionsqualität der Systeme genutzt werden kann.

Contents

1 Background

With the development of automation technology in the automobile domain in the last decades, automated or autonomous driving gradually comes into people's daily life, which is expected to increase traffic safety by reducing human errors and release drivers by executing different driving tasks (National Highway Traffic Safety Administration, 2017). According to SAE International (2021), there are six levels of driving automation, from no driving automation (level 0) to full driving automation (level 5), as shown in Figure 1. From SAE Level 0 to SAE Level 2, drivers need to drive and supervise the support features at the same time, while some driver support features are engaged. For example, at SAE Level 2, drivers still need to monitor the road and also take over the driving task facing critical events. Currently, many car manufacturers have made SAE Level 2 (partial automation) possible with the introducing of assisted features, such as Tesla's Autopilot, Audi's Traffic Jam Assist, BMW's Driving Assistant Plus (TEOH, 2020). From SAE Level 3, drivers don't need to drive and monitor the traffic situation and only intervene when the automated system requests. It can be said that at SAE Level 3, automated driving is first truly achieved. In this project, we will focus on the indicators that describe a safe human-machine interaction at SAE Level 3 (see Figure 1, red marked).

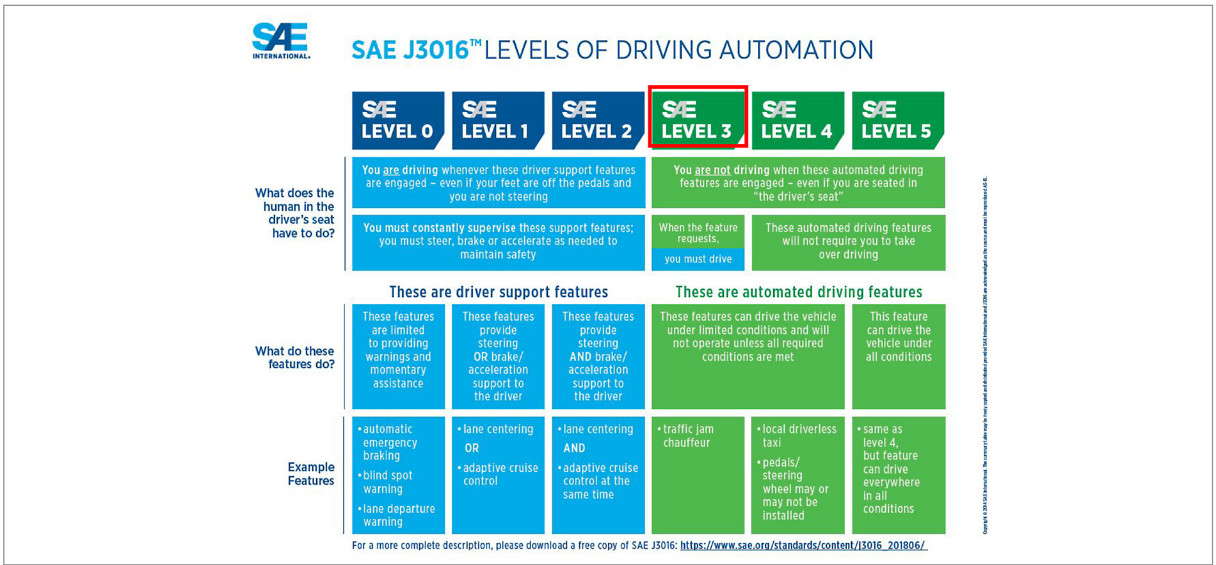


Fig. 1: Defined 6 levels of driving automation by SAE

1.1 SAE Level 3: Conditional Driving Automation

According to the taxonomy of driving automation defined by SAE, human's role changes with the increasing automation levels. At SAE Level 3, driving automation systems need to perform the entire dynamic driving task, so the user does not have to do so and can engage in non-driving related activities. However, the user is expected to take over the driving task if a system failure occurs or when the driving automation system is about to leave its operational design domain. Then the user is expected to be able to resume the driving task when alerted (SAE, 2021). Although many car manufacturers have announced their plans to release higher levels of driving automation (SAE Level 3 or above) recently, only Mercedes-Benz and Honda have released the first vehicles equipped with an approved level 3 feature, namely "Drive Pilot" and "Traffic Jam Pilot" in 2021. The system automatically executes lateral and longitudinal control while monitoring the vehicle's surroundings, using data from high-definition mapping and external sensors.

1.2 Problems of SAE Level 3 in Europe

As SAE Level 2 systems have been available in many series-production vehicles, there are already relevant criteria for assessing the safe human-machine interaction, which were investigated and reported in the BASt project FE 82.0708/2017 "Development and simulation of scenarios and continuous automation functions for an assessment of human-machine interaction safety" and FE 82.0709/2017 "Development of a checklist procedure for the evaluation of the safety of human-machine- Interaction of Continuous Automation Functions". However, explicit criteria for higher automation levels (SAE Levels 3 and 4) have not been developed due to a lack of vehicle technology regulations at UN ECE level.

The draft regulation UN-R 157 was finally adopted in June 2020 at UN ECE level. With this, it is possible for the first vehicle systems for automated driving at speeds up to 60 km/h (Automated Lane Keeping System, ALKS) to have necessary international regulatory framework for its approval. Although the design possibilities of the human-machine interaction will be standardized to a high degree because of detailed vehicle technology requirements, it is expected that various car manufacturers will define their specific design suggestions of human-machine interaction with different concepts within the regulatory framework. In order to promote potentially better human-machine interaction concepts in the context of consumer protection, the appropriate indicators and criteria for assessing the safe human-machine interaction for these systems in extension of the existing criteria for SAE Level 2 systems are needed.

2 Overview of the project

2.1 Goal and Research Question

Now with clear regulations for operating SAE Level 3 automated vehicles at UN ECE level, it is possible to develop safety criteria for the human machine interaction. So the research question of this project is to identify appropriate indicators and criteria for evaluating the safety of human-machine interaction for first SAE Level 3 systems up to 60 km/h based on a literature review.

The goal of the project is to develop a checklist which is able to evaluate the safety of human machine interaction at SAE Level 3 automated vehicles. Also, there will be a concept for the digital solution of the checklist to evaluate the safety of SAE Level 3 vehicles efficiently in practical application.

2.1.1 Approach

In order to answer the research question in this project, a six-step approach will be applied (see Figure 2). At the beginning publication channels will be identified with regard to safety indicators for the evaluation of the Human-Machine-Interaction (HMI) of SAE Level 3 systems. Within the identified publication channels, a first sample of about 30 publications will be generated based on an initial list of keywords. These samples can then be used to compile a list of criteria for the classification of publications. This will mainly be done within Work package 1.

In Work package 2, publication channels will be systematically queried and relevant publications will be extracted based on the list of keywords. A sample of 10-20% of all identified publications will be randomly selected. If necessary, further aspects will be added to the keyword list and the list of criteria.

In Work package 3, a meta-analysis will be then carried out based on the relevant publications. The results of each publication will be summarized statistically and the effect size for each safety indicator will be calculated.

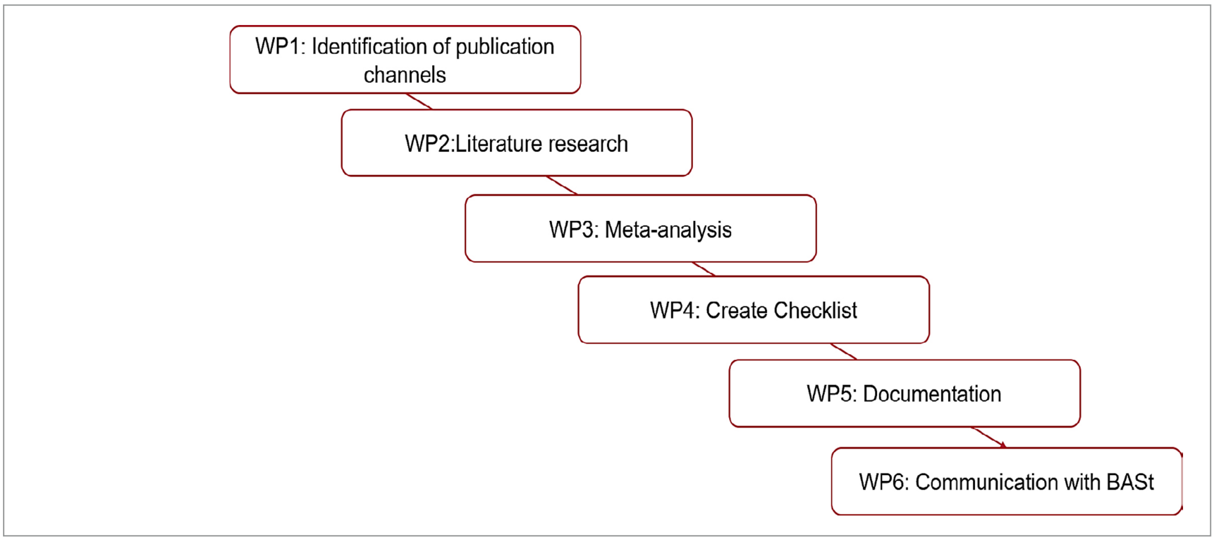


Fig. 2: Overview of the approach of the BASt project

In Work package 4, a checklist to evaluate the safety of the HMI of level 3 Systems will be created based on the results of the literature research and the meta-analysis. The results will be further evaluated in the form of expert interviews.

During the whole project, regular meetings will be arranged between UULM and BASt to present and discuss the results (Work package 5). Finally, a scientific report will be prepared, where the results and the methodological approach will be summarized and integrated (Work package 6).

3 Project Progress of Work Package 1: Identification of Publication Channels

3.1 Goal of work package 1

In work package 1, it aims to provide a first version of identified publication channels and keywords related to the evaluation of safe human-machine interaction at SAE Level 3 in the context of automated driving based on the focus group interview with domain experts.

3.2 Focus Group

3.2.1 Participants

Two moderators conducted the focus group with experts on the field. There were 12 participants in total. The educational level of the 12 experts on traffic research was at least a Master Degree and four of them had a doctoral degree. Eight females and four males participated, where one of the experts was an extern and 11 were from the human factors department of the Ulm University. The overall age distribution was around 30 years old.

3.2.2 Questions for Interview

Three main questions were asked during the focus group, with the possibility of adding additional comments on the process and approach towards the project goals. First, the project goal and the process achieving them was presented. Afterwards the questions were asked and there was a time frame of around 10 minutes to answer each of them. The three questions that were asked subsequently were as follows:

- Q1: Which publication channel could be relevant for the safe interaction between the driver and the system at SAE Level 3?
- Q2: At SAE Level 3, which scenarios or use cases are safety critical for HMI?
- Q3: Based on the safety critical scenarios, which key words are relevant for safe interaction between driver and system at SAE Level 3?

During the answering phase, participants could write down their ideas and estimation of the publication channels and situations anonymously in paper form and in a word cloud through an online service on the third question (Q3). The suggested answers by experts for each question were collected by the moderator and then discussed in the focus group.

3.2.3 Results

Regarding Q1: “Which publication channel could be relevant for the safe interaction between the driver and the system at SAE Level 3?”, experts mentioned first the international and national Standards, such as ISO TC22 and UN regulation No. 157 “Automated Lane Keeping Systems”. The second publication channel suggested by experts are literature database (e.g. ebsco), Journals in the field of traffic psychology, engineering psychology, human factors (e.g. Accident Analysis and Prevention, Transportation Research Part F: Traffic Psychology and Behaviour, Human Factors) and also the FESTA handbook for automated driving. In addition, experts also mentioned the relevant deliverable from the EU project SHAPE-IT (Supporting the interaction of Humans and Automated vehicles: Preparing for the Environment of Tomorrow) and ACM or IEEE conference (e.g. CHI, AUTOMOTIVEUI, ITSC, IV).

For Q2: ”At SAE Level 3, which scenarios or use cases are safety critical for HMI?”, the mentioned critical scenarios at SAE Level 3 by experts are listed in Table 1, which includes the scenarios where weather condition is bad, or sensor is defect etc. Especially, the scenario 6 is suggested by most experts, namely the unexpected or unpredicted event about traffic situations or other road users, in which automation is not able to solve the new problem facing unexpected situations and needs drivers to take over the driving task.

Critical Scenarios at SAE Level 3
1. Passing through a narrow passage/end of road type/road worts
2. Accident/accident ahead, construction side ahead/stopped vehicle/other vehicle bralkes down/ Traffic jam
3. Weather condition/bad weather take over/weather condition/bad weather/weather changes
4. No traffic rules apply(illegaled road)
5. Defect sensors/seonsor malfunction_affected by weather/system breaks down ()TOTB< 10 s and driver is distracted)
6. Unexpected changes/unpredictable behavior of other road users/unplanned events/unforseen event(missing lane markings)/Unexpected event are not communicated to the vehicles
7. Ambulance/police/firefighters
8. Interaction with vunarable users, e.g. cyclists
9. Missing traffic signs, or damaged/unreadable by AV/missing lane markings
10. Situations where driver takes over suddenly
11. Driver state/situation where take over hy human is not possible, e.g. sleeping
12. Rental car, in which driver has no ideas about the automation level of car

Tab. 1: Suggested critical scenarios at SAE Level 3 in Focus Group

Concerning Q3: ”Based on the safety critical scenario, which key words are relevant for safe interaction between driver and system at SAE Level 3?”, each expert was asked to type the relevant keywords in an online tool (https://www.menti.com/wzohe885w7). It is an interactive process, which means after typing each keyword, each expert can see the output from others at the same time. Besides, they were also allowed to type the same key word from others, if they found it also relevant. Below is an overview of suggested keywords by experts (see Figure 3), and the bigger the font size of the keyword is, the

Fig. 3: Suggested key words related to evaluation of safe human-machine interaction at SAE L3

more frequently it is mentioned. It can be seen that “situation awareness” is mostly suggested by experts, and “distraction” followed. Besides, “take over maneuver” or “take over time” are also strongly suggested.

3.3 Sample Papers

The selection of the sample papers is mainly based on two references. On the one hand, suggested publication channels and keywords by experts in the focus group were taken into account. On the other hand, two review papers (ZHANG et. al., 2019; WEAVER et. al., 2020) regarding take over scenarios were considered, which provide a good range and reference for selecting sample papers. Totally, a sample with 33 sample papers in relation to take over scenarios in SAE Level 2 and SAE Level 3 driving was selected by experts who have well-founded knowledge of the research topic in the department.

Based on the collected sample papers, the list of keywords was revised and supplemented as necessary. Both this list of keywords and the list of relevant publication channels will again be presented to the expert panel for discussion and additions.

3.4 First Version of the List of Keywords and identified Publication Channels

The list of Keywords that were extracted from the focus group and also throughout sample papers were merged into a list, which is the starting point for the extensive literature research in the next step. The list includes the keywords with its frequency and also the corresponding publication channels. It can be seen that except the frequent mentioned key words related to the research area, such as “automated driving” (18), “human automation interaction” (5),” driver behavior” (6) and “driving simulator” (6), the most relevant and frequent keywords in content are “(driver) take over” (12), “situation awareness” (4), “reaction time” (3), “distraction” (3), “non-driving related tasks” (3) (marked in yellow). Coincidentally, they are almost consistent with the suggested keywords by experts in the focus group.

In addition, the identified publication channels have also been listed in the previous progress report of work package 1 next to the keywords, and it can be summarized that Elsevier publisher including the journals “Accident Analysis and Prevention”, “Transportation Research Part F: Traffic Psychology and Behaviour” and the journal “Human Factors” are the two main publication channels relevant to our project topic.

3.5 Summary

A focus group interview with human factors experts in the automobile domain has been conducted to explore the publication channels and key words related to the evaluation of safe human-machine interaction at SAE Level 3 in the context of automated driving. In addition, a sample of 33 relevant papers have been selected and used to update the list of keywords and publication channels from the focus group. In the end of the work package 1, a first version of keywords and publication channels has been provided.

4 Project Progress of Work Package 2: Literature Search and Selection

4.1 Goal of the Work Package 2

In the work package 1, the first version of identified publication channels and keywords have been provided. Based on this, work package 2 aims to systematically search and select relevant publications about indicators that influence the safe interaction between driver and system at SAE Level 3.

4.2 Literature Source and Search Strategy

Firstly, the suggested keywords in the work package 1 have been used for literature search, such as “automated driving”, “take over”, “human automation interaction”, “driver behavior” and “driving simulator”. The publication channels have been updated in this project period and it mainly includes databases such as PsycINFO (APA), PsycArticles (APA), PSYNDEX, Web of Science), and the targeted journals (Accident Analysis and Prevention, Transportation Research Part F: Traffic Psychology and Behaviour, Human Factors) and conference proceedings (Conference on Human Factors in Computing Systems, ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, IEEE Intelligent Vehicles Symposium, IEEE International Conference on Intelligent Transportation Systems).

Based on these suggested keywords, the identified publication channels were systematically queried and the relevant literature were extracted. A sample of 10-20% of all identified publications has been randomly selected, which was used to update the list of keywords. Besides, the selected papers used for the meta-analysis from two review papers by ZHANG et. al. (2019) and WEAVER (2020) are also taken into account, which covers almost all relevant publications addressing take over request at SAE Level 2 or 3 from 2012 to 2020. Furthermore, reference papers which have cited the review papers from ZHANG et. al. (2019) and WEAVER (2020) have been searched from 2019 until 2021.

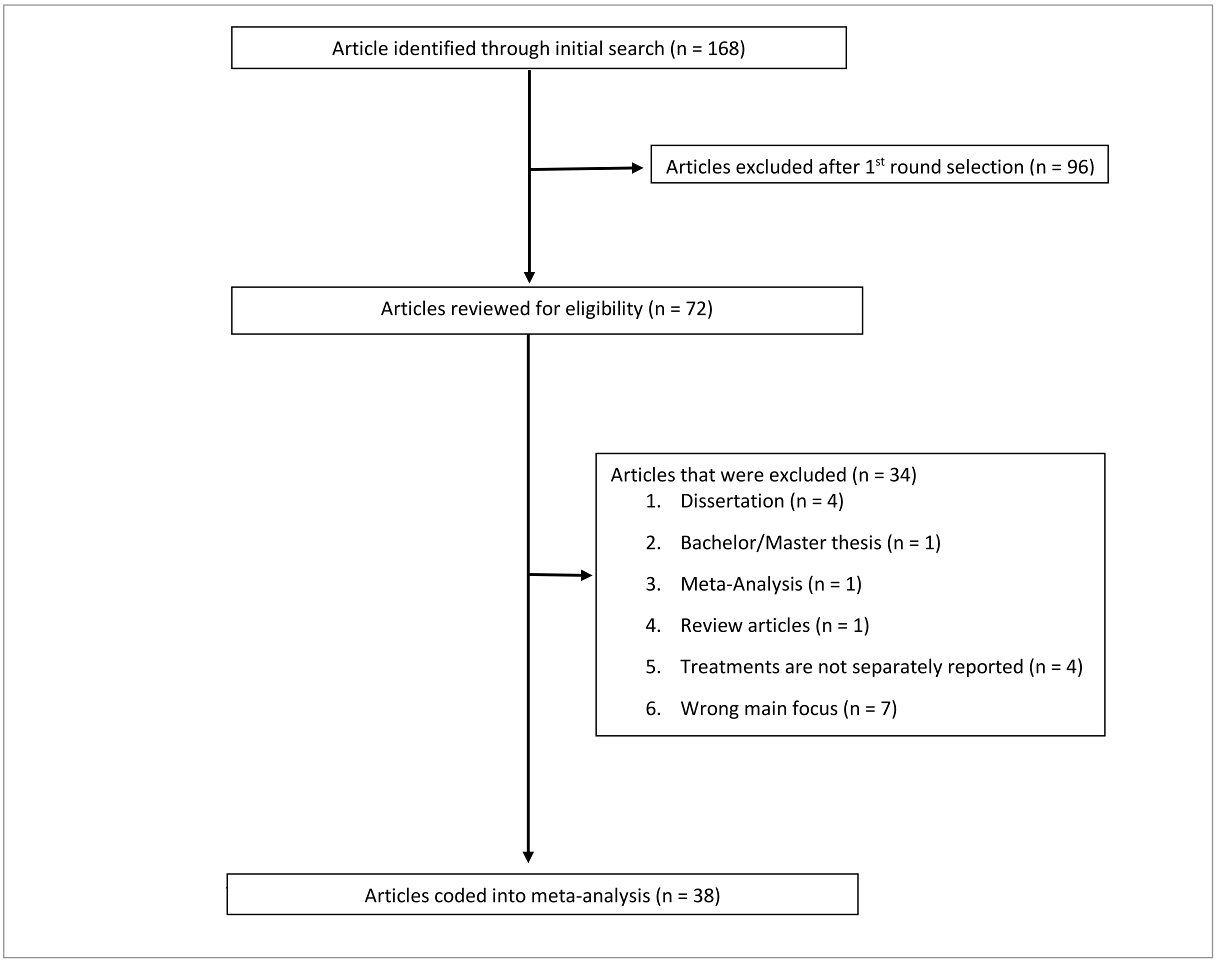
4.3 Inclusion Criterion

An inclusion criterion has been defined, in order to help selecting appropriate publications:

1. The study had to involve a transition from conditionally, or highly automated driving (i.e. SAE Level 3 and above) to manual driving.
2. The study had to involve an automation-to-manual take-over performed by a human (e.g. braking, steering, button pressing).
3. The study had to involve a transition in response to a TOR or a critical event in the environment.
4. The study had to report a mean take-over time (TOT) or variables reflecting take over quality, such as time to collision (TTC), and maximal acceleration. At least, it should be possible to calculate these values based on the reported information.

4.4 Literature Selection and Exclusion Criterion

After initial searching and filtering, totally 168 relevant full-text publications have been retrieved and documented with their titles, authors and years, experimental design, measurement, and the non-driving-related task (NDRT). In the first round, the range of retrieved publications has been narrowed down by selecting representative sample papers for the meta-analysis. For this purpose, 38 common papers both chosen by ZHANG et al., (2019) and WEAVER et al., (2020) in their review papers regarding take over request have been first selected. Besides, reference papers (43 papers) since 2019 have also been chosen, which makes supplements to the selected common papers by ZHANG and WEAVER. Hence, 72 records are left in this round. In the second round, the pre-selected papers have been reviewed for eligibility. 34 papers were excluded based on the exclusion criterion (see Table 2). In the end, 38 relevant publications have been considered for the meta-analysis in the work package 3. For each selected article, the title, author and year, sample characters, experiment design and independent variables, measurement, modality of NDRT, modality of takeover request (TOR), the classification of the situation triggering TOR, and fidelity of the driving simulator are documented in the previous progress report of work package 2 and work package 3.



Tab. 2: Flowchart of the inclusion and exclusion of articles

4.5 Summary

In work package 2, relevant publications with regard to indicators that influence the safe interaction between driver and assistance systems at SAE Level 3 has been extracted based on the identified publication channels and list of keywords. After literature selection, 38 relevant publications have been selected for the meta-analysis in the work package 3.

5 Project Progress of Work Package 3: Meta-Analysis

5.1 Goal of the Work Package 3

The selected literature in work package 2 will be then used to conduct the meta-analysis for work package 3. The results of each publication will be summarized statistically and the effect size for each safety indicator will be calculated.

5.2 Previous Reviews and Meta-Analyses

There are some narrative reviews on takeover request in the context of automated driving from 2014 to 2017, in which the findings are summarized by the domain experts (ERIKSSON & STANTON, 2017; KÖRBER & BENGLER,2014; RADLMAYR & BENGLER, 2015). However, narrow reviews have many limitations, such as the resolution of the conflicting results and the decisions are too subjective. In order to overcome this, meta-analysis have been taken into account, which have quantitative procedures and focus on effect size instead of statistical significance (ROSENTHAL & DIMATTEO, 2001). The two representative reviews for takeover scenario using meta-analysis are ZHANG et al. (2019) and WEAVER et al. (2020). ZHANG only focused on the measurement of take over time and found that driver took over more quickly when given a short time budget, using a hands-free device, performing a non-visual NDRT, having experience of takeover before the experiment, and receiving an auditory or vibrotactile takeover request. WEAVER made supplements to ZHANG's review, e.g. the measurement of takeover quality was also considered and they examined the effects of time budget, NDRT, and information support on takeover timing and quality measures. They found that engagement of NDRT degraded the takeover performance, especially there was an overlap of the resource demands between NDRT and the driving task.

5.3 Research Question

Different to the mentioned reviews by ZHANG and WEAVER, this meta-analysis focus on the influence of the situation or scenario that triggers takeover request on takeover performance, which has not been investigated until now. The research question can be formulated as: "Do the situations/scenarios that trigger the takeover requests influence drivers' takeover performance?"

5.4 Classification of Scenarios triggering TOR

In the selected 38 papers, there are many types of situations that trigger the takeover request. An overview of the types of takeover scenarios and their corresponding papers can be seen in the Table 3. One category of takeover scenario (A) is the road block caused by road construction, broken or stopped vehicle, or faded lane markings in front. The second category of takeover scenario (B) is the braking or slow movement of the lead vehicle or the suddenly approaching vehicle from behind. Besides, system failure (C) can also trigger the takeover request. In the last category (E), no scenario has been described, but only a takeover request is given.

Takeover Scenarios	Paper ID
A. Road block: stationary vehicle; crashed vehicles; stopped or broke vehicle, obstacle	2, 6, 8, 10, 11, 12, 13, 18, 19, 20, 21, 23, 25, 26, 27, 28, 32, 33, 34, 35, 37
B. Lead vehicle suddenly brakes or moves slowly; approaching vehicle from behind	3 ,4 ,7 ,9, 15, 31, 32, 38
C. Automation reaches limits or system malfunction	7, 14, 24, 32, 36, 37, 38
D. Request to intervene	1, 5, 16, 17, 25

Tab. 3: Classification of scenarios triggering TOR

In order to conduct the meta-analysis, these types of takeover scenarios need to be further classified into two comparable categories. Therefore, two categories of takeover scenarios are defined: static takeover situations and dynamic takeover situations. Static takeover situations mainly refer to the category A, where the takeover request is usually caused by a static object in front. Different to this, dynamic takeover situation refers to the category B, in which the object that triggers takeover request moves in front or from behind. For category C and D, the takeover scenario has not been explicitly described and a further look of the relevant takeover situation in the paper has been conducted, which helps to classify the takeover scenarios into static and dynamic situations.

5.5 Category of Safety Indicators

To measure the takeover performance, safety indicators that can measure the safe interaction between drivers and automated systems have been taken into account. Totally three categories of safety indicators are considered in the meta-analysis. First category is takeover timing, which mainly includes hands-on time (GOLD et al., 2016), time to steer, time to brake, takeover time. The second category is takeover quality, such as Time to Collision (TTC) or maximal acceleration (GOLD et al., 2016). The last category is about the subjective measurement of takeover performance, such as workload. An overview of the safety indicators about takeover performance can be found in the Table 4.

Category of Safety indicators	Measure
Takeover time	Hands-on time, Time to steer, Time to brake, Takeover time
Takeover quality	Time to Collision, Maximal Acceleration
Subjective indicator	Workload

Tab. 4: An Overview of the safety indicators about takeover performance

5.6 Preparation of the Data Frame

In order to investigate the research question concerning the influence of situations triggering takeover requests on safety indictors, meta-analysis has been chosen. To prepare the data frame for conducting meta-analysis, relevant statistic measures (mean, standard deviation) were extracted from selected papers and were summarized in Excel, so that raw mean difference or effect size can be calculated. It was observed that some papers had directly reported the means or standard deviations. However, some papers didn’t mention these relevant values in texts and only provided figures about means and standard deviations. In this case, mean and standard deviation values were estimated from relevant diagrams using a tool named WebPlotDigitizer. With this, the data frames for each indicator have been separately prepared for meta-analysis, when at least four studies were available for each indicator. Moreover, there should be at least one study for the category of static situation trigger available and one for the dynamic situation trigger available.

5.7 Analysis and Results

R software (Version 4.1.1) was used for the meta-analysis using the package “meta”. As there were more than two studies per indicator category, a random-effects model was used to calculate effect sizes. Moreover, the relevant statistics (mean and standard deviation) were weighted based on the sample size for each selected paper.

Raw mean difference and Cohen’s d have been considered for the output of meta-analysis. Raw mean difference quantifies the effect size at the original units, which is appropriate when the dependent measure has same and meaningful units. However, Cohen’s d quantifies the effect size on a standardized scale, which is meaningful when measurements have different units. As all indicators per category have the same units in selected papers, both raw mean difference and Cohen’s d are reported. The results of meta-analysis for each safety indicator is reported below with the corresponding forest plots (LEWIS & CLARKE, 2001), which gives a graphical description of the meta-analysis data.

5.7.1 Takeover Timing

5.7.1.1 Takeover Time

Take-over time refers to the time until the actual driving manoeuvre begins, steering wheel angle > 2° or brake pedal pressure > 10% (KERSCHBAUM et al., 2015). This analyses studies whether these two dynamic and static situation triggers affect the overall takeover time. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen’s d). The result of the raw mean difference is illustrated in the Figure 4: the takeover time induced by the static trigger has an overall mean of 2.49 s (SD=0.78), while the dynamic trigger has in contrast an overall mean of 2.32 s (SD=0.68). Furthermore, it was examined if there is a significant difference between the dynamic trigger and the static trigger by using a z-test with an alpha level of 0.05. The raw mean difference (MD) between them is 0.17 s and Cohen’s d is 0.23 ($p < 0.01$), which indicates that the situation trigger has a significant effect on takeover time. Moreover, drivers in the static situations need more time to take over than drivers in the dynamic situations.

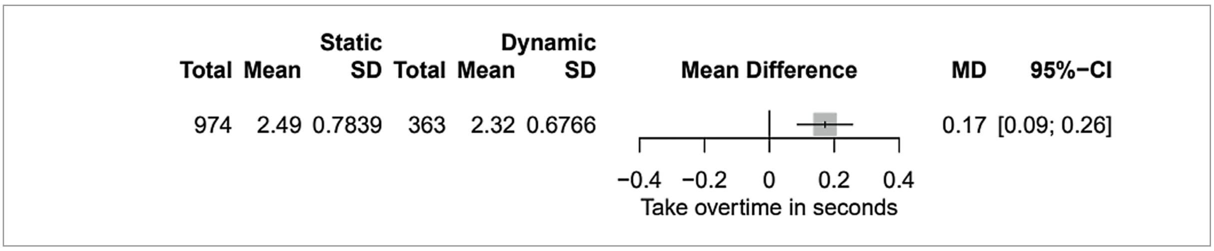


Fig. 4: Mean difference of takeover time

5.7.1.2 Hands on Time

Hands-on time refers to the time until the driver has left or right hand on the steering wheel (KERSCHBAUM et al., 2015). This analysis examines the effect of dynamic and static situation triggers on hands on time. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen’s d). The result of the raw mean difference is illustrated in the Figure 5: the hands on time induced by the static trigger has an overall mean of 1.8s (SD=0.72), while the dynamic trigger has in contrast an overall mean of 2.31s (SD=0.95). The raw mean difference (MD) between them is -0.51 and Cohen's d is -0.61 ($p < 0.01$), which indicates that the situation trigger has a significant and moderate effect on hands on time. Moreover, drivers in the dynamic situations need more time to put their hands on the steering wheel than drivers in the static situations.

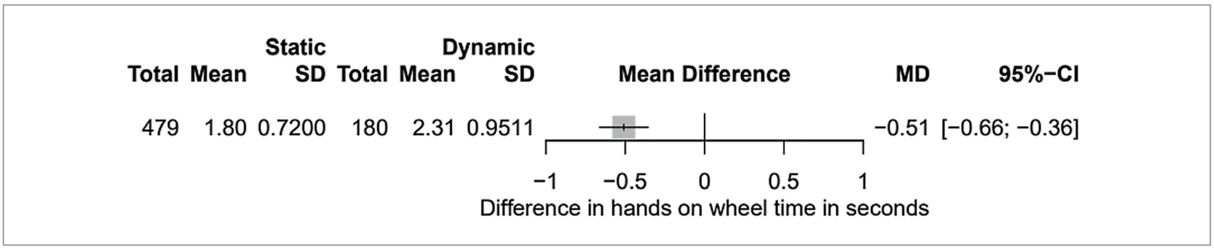


Fig. 5: **Mean difference of hands on time**

5.7.1.3 Time to break

Time to brake refers to the time between the driver has left or right hand on the steering wheel and the time driver brakes. This analysis examines the effect of dynamic and static situation triggers on time to brake. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen's d). The result of the raw mean difference is illustrated in the Figure 6: the time to brake induced by the static trigger has an overall mean of 3.7s (SD=1.01), while the dynamic trigger has in contrast an overall mean of 3.18s (SD=0.88). The raw mean difference (MD) between them is 0.51 and Cohen's d is 0.55 ($p < 0.01$), which indicates that the situation trigger has a significant and moderate effect on time to brake. Moreover, drivers in the dynamic situations need less time to start braking than drivers in the static situations.

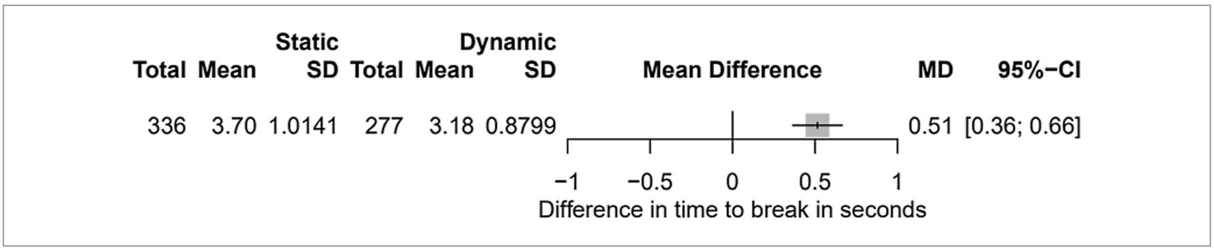


Fig. 6: **Mean difference of time to break**

5.7.1.4 Time to steer

Time to steer refers to the time between driver has left or right hand on the steering wheel and the time driver steers to left or right. This analysis examines the effect of dynamic and static situation triggers on time to steer. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen's d). The result of the raw mean difference is illustrated in the Figure 7: the time to steer induced by the static trigger has an overall mean of 3.24 s (SD=0.72), while the dynamic trigger has in contrast an overall mean of 4.59 m (SD=1.53). The raw mean difference (MD) between them is -1.35 and Cohen's d is -1.1 ($p < 0.01$), which indicates that the situation trigger has a significant and large effect on time to brake. Moreover, drivers in the dynamic situations need more time to start steering than drivers in the static situations.

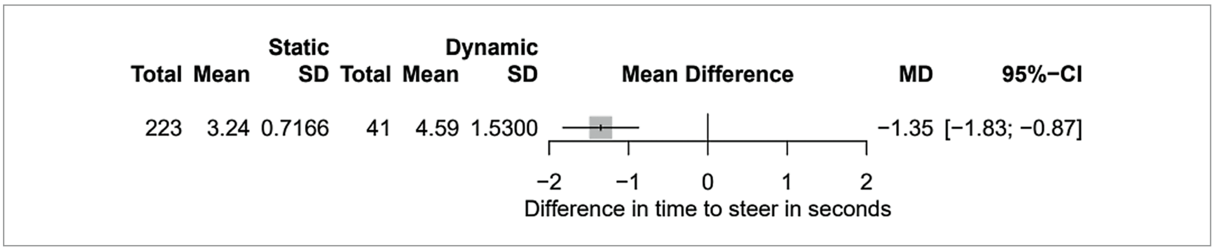


Fig. 7: **Mean difference of time to steer**

5.7.2 Takeover Quality

5.7.2.1 Maximal Acceleration

The maximal acceleration includes maximal lateral acceleration and maximal longitudinal acceleration. As the selected papers have only reported the longitudinal, so this analysis examines the effect of dynamic and static situation triggers on maximal longitudinal acceleration. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen's d). The result of the raw mean difference is illustrated in the Figure 8: the maximal acceleration induced by the static trigger has an overall mean of 4.4 m/s^2 (SD=2.6), while the dynamic trigger has in contrast an overall mean of 6.1 m/s^2 (SD=1.97). The raw mean difference (MD) between them is -1.7 and Cohen's d is 0.74 ($p < 0.01$), which indicates that the situation trigger has a significant and large effect on maximal longitudinal acceleration. Moreover, drivers in the dynamic situations brake more strongly than drivers in the static situations.

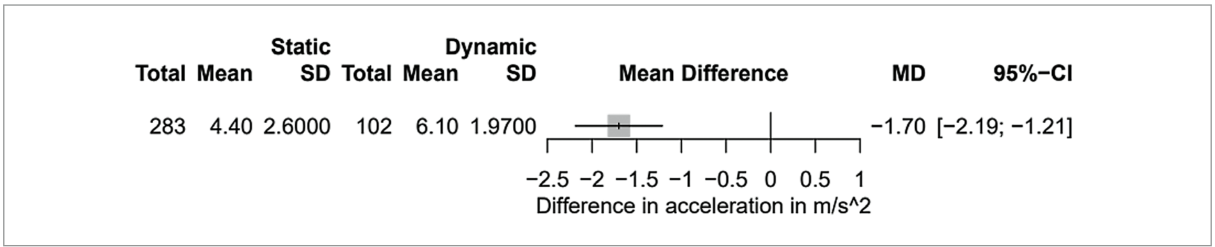


Fig. 8: **Mean difference of maximal acceleration**

5.7.2.2 Time to Collision

This analysis examines the effect of dynamic and static situation triggers on time to collision (TTC). Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen's d). The result of the raw mean difference is illustrated in Figure 9: the time to collision induced by the static trigger has an overall mean of 3.29s (SD=1.24), while the dynamic trigger has in contrast an overall mean of 2.39s (SD=0.88). The raw mean difference (MD) between them is 0.91 and Cohen's d is 0.84 ($p < 0.01$), which indicates that the situation trigger has a significant and large effect on time to collision. Moreover, drivers in the dynamic situations have a lower TTC than drivers in the static situations.

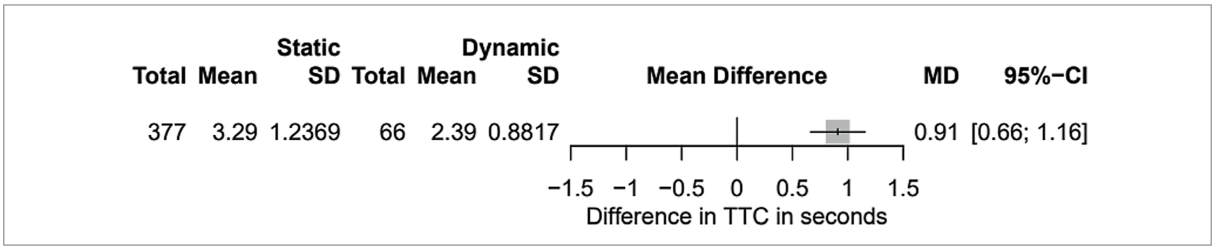


Fig. 9: **Mean difference of time to collision**

5.7.3 Subjective Measurement: Workload

To investigate the difference in perceived workload the overall scale from the NASA-TLX (HART & STAVENLAND, 1988) was used. The assessment measures six dimensions of the workload: mental demand, physical demand, temporal demand, performance, effort and frustration. This analysis examines the effect of dynamic and static situation triggers on subjective workload. Effect size was calculated as the static situation trigger minus the dynamic situation trigger (and divided by standard deviation for Cohen's d). The result of the raw mean difference is illustrated in the Figure 10: the score of subjective workload induced by the static trigger has an overall mean of 39.2 (SD=16.4), while the dynamic trigger has in contrast an overall mean of 29.65 (SD=14.4). The raw mean difference (MD) between them is 9.55 and Cohen's d is 0.62 ($p < 0.01$), which indicates that the situation trigger has a significant and large effect on subjective workload. Moreover, drivers in the static situations have higher workload than drivers in the dynamic situations.

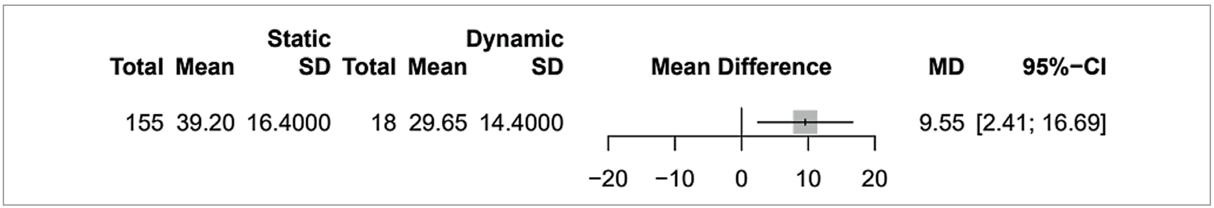


Fig. 10: Mean difference of subjective workload

5.8 Summary

Based on the results of meta-analysis above, it can be concluded that the classified static or dynamic situation triggering takeover request has a significant effect on takeover timing, takeover quality and subjective measurement. Especially, the effect size of Cohen's d is large for the indicators reflecting takeover quality.

Regarding takeover timing, it can be seen that although it takes drivers longer to regain the steering wheel in a dynamic situation, their takeover time overall is lower compared to a static situation. It may be explained that drivers feel more critical of dynamic situation trigger and therefore tend to take over more quickly. However, time budget also needs to be considered for further interpretation. Besides, drivers in a dynamic situation need more time to begin steering than to brake, while drivers in a static situation need more time to brake than to steer. It can be interpreted that drivers may brake strongly in dynamic situations and steer strongly in static situations. The maximal acceleration needs to be taken into account to help to understand it.

Concerning takeover quality, it can be summarized that drivers brake more heavily in dynamic situations than in static situations, which on the one hand indicates the dynamic situation is more critical than the static situation. On the other hand, it supports the conclusion regarding time to brake in dynamic situations, where drivers brake strongly and quickly in dynamic situations. In addition, the time to collision is shorter in a dynamic situation than in a static situation. As TTC is a measure that reflects the criticality of the traffic situation, it can be implied that the dynamic situation trigger is more critical than static situation trigger.

Regarding measurement of subjective workload, it can be seen that driver have higher workload in static situations than dynamic situations, which is opposite as expected. One possible reason could be the influence of the NDRT, which needs to be further studied.

6 Project Progress of Work Package 4: Development of Checklist

6.1 Goal of Work Package 4

Based on the results of the literature research in the work package 2 and meta-analysis in the work package 3, a summary view of safety indicators will be prepared, in order to develop a checklist for the evaluation of the safety of human-machine interaction at SAE Level 3. These results will be evaluated in comprehensive expert interviews to help to develop the checklist. The finalized checklist will be transferred into a digital application software that allows an efficient and reliable evaluation of human-machine interaction at SAE Level 3.

6.2 Expert Interview

In order to help to interpret the results of meta-analysis in the work package 3 and further develop the checklist for the evaluation of the safety of human-machine interaction at SAE Level 3, expert interviews involving 6 experts have been conducted.

6.2.1 Guidelines of expert interview

In order to achieve a comprehensive evaluation in the expert interviews, a corresponding expert interview guideline has been developed. First, a draft of the expert interview guidelines has been discussed with BASt in advance to collect initial feedbacks and suggestions. Based on this, the draft of expert interview guideline has been adjusted and tested with one expert. The results of the test interview were then communicated with BASt to collect the final feedback and finalize the guidelines of expert interviews.

Before opening the interview, some basic information with regard to expert's name and gender, expertise, date will be noted. At the beginning of the interview, a short introduction of the interviewer herself and also the background of the cooperation project between BASt and UULM will be given. After that, the main interview part will start with the following 6 interview questions:

Q1: If there is a system that support drivers to take over, do you think the assistance system should distinguish itself in different takeover situations?

Q2: If the answer is yes, in which kind of takeover situations do you expect this difference? Could you give an example?

Q3: We did meta-analysis to have a look at the influence of the type of takeover situations on takeover performance. We classified the takeover situations into static and dynamic situations: static takeover situations refer to the situations where the obstacle in front is static (e.g. there is a broken car in front); whereas in dynamic situations the surrounding obstacle has a dynamic movement (e.g. the lead car suddenly brakes). After meta-analysis, we found that:

a. In dynamic situations, driver´s takeover time overall is shorter and brake reaction is more quickly than in static situations.

b. Drivers in dynamic situations need more time to put their hands on the steering wheel and also more time to steer than in static situations.

c. Drivers brake more heavily in dynamic situations and TTC is shorter than in static situations.

d. Drivers’ workload is higher in a static situation then dynamic situation.

What do you think of the reasons for these results?

Q4: How can a system support drivers in different situations?

Q5: Please try to formulate a Checklist Item to evaluate the safe interaction between a driver and automated vehicles.

Q6: Do you have a specific system in mind that is already available on the market? Which items/functions do you propose to consider in the near future? Which items should be kept in mind for vehicle manufactures? Which items will not be considered for the production of series production vehicles? When or how far in future can it be implemented in series production vehicles?

It can be seen that the first two questions (Q1, Q2) aim to figure out whether experts find it necessary to provide different support in different takeover situations. The third question (Q3) is about the key findings of the meta-analysis, which need experts’ interpretations to help understand these results. Q4 aims to further collect experts’ opinions about specific support in different takeover situations. In the Q5, Experts are asked to suggest checklist items for evaluating safe interaction between drivers and systems. The last question (Q6) expects experts to give their considerations and visions on implementation of such assistance systems at SAE Level 3 in the future.

6.2.2 Selection of Experts

The selection of experts for interviews is based on two aspects. On the one hand, it is based on the broad network of relevant experts available in the department of Human Factors of Ulm University through numerous collaborations with universities, scientific institutions and industry partners. On the other hand, authors of relevant publications addressing the project topic are also considered.

In total, 6 nationally and internationally recognized experts have been selected, with the disciplines of psychology, computer science and engineering in the fields of automated driving, traffic safety, and human factors. They all have a PhD degree and one of them is a professor. Moreover, they all have expertise on driver-vehicle interaction, especially automated driving at SAE Level 3. One (Expert 1) of them is female and half of them come from industry, whereas the others come from universities in Europe.

6.2.3 Conduction of Expert Interview

Each expert interview has been conducted between March 10th, 2022 and May 3th, 2022. The experts were interviewed online except of the first expert interview which was conducted offline. Before conducting interviews, an email invitation has been sent to the selected experts at the beginning of March, 2022 to find a possible date for each expert. In the email, the interviewer generally introduced the background of the project and the focused topic of the interview to the experts. However, the specific interview questions were not mentioned. After the negotiation and the confirmation of experts, an online Zoom link was sent to experts with the corresponding dates. There was an exception that two experts from automobile industry were not able to use Zoom software. As an alternative, Microsoft Teams was then used once for interviews. The duration of each interview was between 40 to 50 minutes and two of the interviews were conducted in German, whereas the other 4 interviews were conducted in English.

At the beginning of the interview, the interviewer greeted experts and introduced herself. Then a short introduction to the background and research question of the project was given by sharing the slides with experts at the same time. Experts were also told that their answers to the interview questions were anonymized and used to develop a checklist for evaluating the safety of the human-machine interaction at SAE Level 3. If they were clear and had no further questions, the main interview part started. Before asking the first interview question, the interviewer would ask experts if they agreed that the interviewer made a video recording from that moment, so that a detailed protocol can be made after the interview. The interview questions were shared with experts per slides, so that they could have a look at the questions at any time. The interviewer made also some notes during the interview and guided the expert to focus on the given interview question, when the deviation of expert focus appeared. After the expert answered the last interview question, the interview was finished and the video recording was stopped and saved.

6.2.4 Analysis and Results of Expert Interview

Totally, 6 expert interviews have been conducted either online or offline. The audios and videos of the online sessions were recorded and used for further analysis. For the face-to-face interviews, only the audio file was available for data analysis. First and foremost, the recordings of expert interviews were transcribed by employees of the university of Ulm without any transcription software and saved as a word document for further analysis. The results of expert interviews were summarized below:

With regard to the first question Q1 “If there is a system that support drivers to take over, do you think the assistance system should distinguish itself in different takeover situations?”, all experts agreed that there should be a distinction of system’s support in different takeover situations.

For the Q2: “If the answer is yes, in which kind of takeover situations do you expect this difference? Could you give an example?”, all experts answered that the biggest difference of system’s support in either critical or non-critical takeover situations should be the time to notification. Expert 2 emphasized that there should be a minimum time for the system requesting the driver to take over control. Besides, Expert 2 also mentioned that the transition from SAE Level 3 to SAE Level 2 assistance influenced the HMI (Human-Machine Interface).

One expert (Expert 1) mentioned that the system should provide a different kind of support regarding the level of urgency in takeover situations. This level of urgency could be communicated via a light display. With growing urgency, the car can communicate with different color, different presentation patterns, different phase of showing how urgent the situation is. For example, time for takeover in non-urgent situation is 10 seconds using yellow light display with a slow light movement; in urgent takeover situations where only 5 seconds was available, red light can be used for critical warning. Similar to Expert 1, Expert 3 also gave the example of corresponding support in critical situations using a big red indicator and multimodal warnings, such as auditory cues or tactical cues, whereas only visual cues with not so many colors can be used in non-critical situations.

Expert 4 and Expert 5 both mentioned the importance of providing early takeover notifications. Expert 5 thought although the regulations have defined the 10 seconds and 6 seconds for takeover request, it was suggested to give the notification much earlier than the regulated in non-critical situations. Expert 4 found that in non-critical situations the automated mode should be permanently displayed (for example, the steering wheel lights up permanently in blue). The illumination of the steering wheel can be changed in the case of a road works where you know that an obstacle is coming in the near future. In critical situations, it should make sure that a driver still has enough time to takeover and the notification should not frighten the driver. Hence, it is important to give more and clear information, such as a spatial warning (e.g. there is an obstacle in front that you have to react to). This can be supported by a red light in front of the driver drawing his attention in the right direction and combined with acoustic cues. Expert 6 stated that the support in planned takeover situations should be different compared to unexpected situations.

For the Q3 regarding the four key findings from the meta-analysis about the influence of different takeover situations on the safety indicators, all experts found that the first three points fitted together, but the fourth point was surprising.

Considering longer Hands-on time in dynamic situations, Expert 1, Expert 4 as well as Expert 5 explained that drivers needed more time to analyze the complex situation and to make a decision with the first motor action. Expert 3 thought the longer hands-on time was related to the fact that drivers needed to focus on braking. Expert 6 pointed out that difference of hands-on time in different takeover situations may be related to the non-driving related tasks or the perception of the situation.

For the shorter takeover time in dynamic situations, Expert 4 explained that drivers allowed themselves more time to prepare. Nevertheless, they braked and steered intensively in dynamic situations and therefore the takeover time was shorter. Expert 6 mentioned that drivers in dynamic situations could perceive relevant objects more quickly due to the dynamic movements and therefore react to them more quickly than to a static object.

With regard to the faster and stronger braking actions in dynamic situations, Expert 2, Expert 3 and Expert 5 had a similar interpretation. Expert 2 found that in dynamic situations especially when time for takeover was short or the system limit reached, drivers' needs to get more time were higher than in static situations, therefore they started faster with the braking action by decreasing speeds to save more time for final correct decisions. Expert 3 also mentioned that drivers needed more flexibility to react to the complex situation, therefore they needed to be quick, and by quick and strong braking they could have more freedom to react to uncertainties. Expert 1 and Expert 6 emphasized that braking action was intuitive, and it was a defensive mechanism and an unconscious reflection of difficult situations.

Regarding shorter TTC in dynamic situations, Expert 1 and Expert 4 explained that as being related to drivers needing more time to make decisions in dynamic situations. When time was running out, they still had to brake harder and hence had a short TTC.

For the results concerning workload, all experts stated that they expected the opposite results. An alternative explanation by Expert 1, Expert 2 and Expert 6 could be that some static objects created more uncertainty, and drivers needed additional effort to identify and assess the object, which led to higher workload. Besides, Expert 4 thought maybe that the driver didn't need more time to orientate himself in the dynamic situation.

For the Q4: "How can system support drivers in different situations?", experts gave their suggestions on the support in static and dynamic takeover situations. Expert 1 started with the suggestion of using light displays, which could first help to distinguish static and dynamic objects: static objects could be marked with flashing lights and dynamic objects with light with moving patterns. Furthermore, the system could show the position of the object (e. g. with light) and also the safe distance. In dynamic situations the system could calculate the trajectory movement of the surrounding obstacle. For static objects in front, the warning could be a looming sound. Similar to Expert 1’s opinion, but not restricted to the use of light displays, Expert 3 suggested to show the potential conflict points of collisions in real time for dynamic situations by using multimodal warnings, such as auditory cues or tactile cues, whereas it was sufficient to use visual cues to show the driver where the relevant obstacles were in static situations. Expert 4 also suggested to use a light display to support drivers in takeover situations and emphasized the importance of having a permanent light display, showing spatially where the vehicle was going within the next second, which needed to be highlighted for the dynamic situation. Expert 5 also expressed a similar suggestion that the system should support drivers in pointing out the traffic situation for them early, so that they could be aware of the traffic surrounding, for example using the birds eye view, and also be better prepared for urgent situations. Expert 6 suggested to use multi-modal warnings to provide assistance.

Expert 2 emphasized the aspect of timing when supporting drivers in different situations. Expert 2 said that in dynamic situations, it is important to know how immediate the action is needed and what is the capability of the system to prepare an action early enough. The typical times needed for L3 systems were at least 4 to 5 seconds in dynamic situations. In static situations, drivers had in total more time and there can be much more explainable aspects within the HMI. It can be started with some dialog to check the overall situation, to describe the situation and the relevant actions which were not immediately needed, but would be helpful when urgent events happened. Expert 5 also mentioned the importance of explaining the reasons for takeover to drivers to motivate drivers to react. Expert 6 suggested to provide assistance more quickly in dynamic situations than static situations. In addition, the levels of details with either general information or detailed information can be different in the assistance representation.

For the Q5: “Please try to formulate a Checklist Item to evaluate the safe interaction between driver and automated vehicles.”, Expert 1 and Expert 3 both suggested to display the urgency of the situation using multimodal cues and the numbers of modalities should increase with the urgency of the situation. Expert 2 and Expert 5 mentioned that the driver´s availability had to be tracked continuously on a second level. In addition, they also mentioned that the HMI needs to be designed in a way to ensure driver response ability at any moment and to communicate drivers’ preparation for taking over control in a clear way to feel ready to take back control. In dynamic situations, Expert 2 mentioned that the takeover request had to be given to drivers via an easy and noticeable HMI.

Expert 3, Expert 4 and Expert 6 emphasized that the takeover warning should be clearly perceivable by drivers. Expert 4 further emphasized that the warning should not be too intense and it should be possible to design a warning with a reaction time below one second. Moreover, the system should announce a takeover as early as possible. Expert 5 suggested to check the steering performance after the takeover process to achieve steering stabilization and all the interactions should be standardized to avoid relearning systems of different brands. Expert 6 emphasized that the transferred information from the system such as system state and expectations on drivers should be clearly communicated, so that drivers are always aware of the system state to avoid mode confusion and to be prepared for actions.

For the Q6: “Do you have a specific system in mind? When or how far can it be implemented?”, Expert 3 mentioned that Honda and Mercedes-Benz have announced their SAE Level 3 system and that BMW will release its Traffic Jam Pilot next year. Expert 4 and Expert 5 both mentioned that Mercedes would bring a SAE Level 3 system of the Mercedes EQS to the market. Expert 5 explained that it won't be only a traffic jam assist, instead it will be a highway pilot and able to perform lane changes and also drive at higher speeds. Expert 4 had a driving experience of a Tesla vehicle. He mentioned that although the Tesla system was often promoted as SAE Level 3, it has only SAE Level 2 characteristics. He also found that the Tesla often made wrong decisions, which greatly reduced driver trust. Expert 2 and Expert 6 expressed although Honda stated that they had a SAE Level 3 system with some specific functionality, these functions were still considered as SAE level 2 assistance. Nevertheless, there is clear action on the market going in this direction and the functionality is clearly starting with a traffic jam pilot.

With regard to the functions which should be taken into account in the future and how far they can be implemented in series production vehicles, Expert 3 and Expert 5 both expressed that it would be important to communicate the limitations of ODD to the drivers. Expert 2 mentioned that traffic jam pilots should be able to make lane changes. Without this functionality, the development of automated vehicles is far away from bringing a system to the urban market. In addition, Expert 2 said that a Level 3 system should be more reliable to be able to cope with a real dynamic situation at a speed of 120km/h. Expert 4 stressed that the most important aspect was to keep the monitoring of the vehicle as simple as possible. Trust in the system was important, but this trust must be realistic and shouldn't be based on unrealistic expectations on the system. Expert 6 thought that such a system won't be available in the coming two years, maybe after 5 or 10 years. It will be first implemented in specific scenarios such as parking scenarios.

6.3 Development of Checklist

Based on the results of meta-analysis and the expert interviews, a checklist for the evaluation of the safety of human-machine interaction at SAE Level 3 has been developed.

6.3.1 Approach

Based on the notes of the conducted expert interviews, a draft of the checklist items has been driven, especially from experts' opinions regarding Q4 and Q5, which are highly relevant to the development of checklist items. In this draft version of the checklist, each expert's answer in relation to Q4 and Q5 have been formulated to checklist items with the consideration of requirements of assistance systems that can support driver's take over process at SAE Level 3. Initially, 21 checklist items have been driven from the results of expert interviews. These items have been internally reviewed within the department of Human Factors of Ulm University. After a worksheet was created with all important statements of each expert, a search was first made for similar statements from different experts. If there is an overlapping between the checklist items, the similar items will be integrated into one item. It also happened that a checklist item of high relevance occurred only once. In that case the plausibility of the statement was checked by employees of Ulm University. After checking and integrating, 16 checklist items were left for the final version. These checklist items have been further divided into six categories, where each category addresses the same aspect of system requirements for evaluating safe human-machine interaction at SAE Level 3.

6.3.2 Structure of Checklist

The checklist starts with some basic information needed for the evaluation of safe human-machine interaction at SAE Level 3, such as the date, time, expert name, evaluated vehicle or system. Followed are the categories with the corresponding checklist items. For each checklist item, it can be answered with "yes or "no" or with "not applicable". The option "not applicable" can be chosen if the corresponding item cannot be evaluated due to the missing functionality of the system or relevant experience of experts.

After evaluating the relevant items in one category, a summarized rating can be given with three options: "Fulfilled", "Satisfactory", "Not acceptable", based on the Response Code of Practice (Response Consortium, 2006) or the checklist of NAUJOKS et al. (2019c). The option "Fulfilled" can be chosen, if all the items in this category are evaluated with" yes", whereas the option "Satisfactory" is suitable for the assessment where at most one item is answered with "no". The option "Not acceptable" is used when more than one item in this category is rated with "no". Besides these three options, a remark can also be made for each category of requirement, so that a judgment can be noted.

After evaluating all items in the checklist, a final rating for the system evaluation needs to be given from "very good" to "not acceptable":

- Very good - no improvements necessary
- Acceptable with minor options for improvement: Satisfactory with minor options for improvement
- Acceptable with major options for improvement: Satisfactory with major options for improvement
- Not acceptable: Not satisfactory

6.3.3 Category of System Requirements

The 16 checklist items have been further grouped into six categories based on the system requirements for evaluating safe human-machine interaction at SAE Level 3.

The first category is "Requirements for driving at SAE Level 3" which consists of 2 checklist items. This category aims to summarize the important requirements such as driver availability or responsibility to guide the automated driving at SAE Level 3 in general.

The second category is "Requirements for system's takeover request" which has 4 checklist items. As a takeover situation is a typical and sometimes critical maneuver for automated driving of SAE Level 3, this category targets the requirements specific to the system's takeover requests with regard to timing, urgency of the situation and modalities of warnings.

The third category "Requirements for interaction design for Takeover assistance" has totally 5 checklist items. It provides design suggestions for takeover assistance with the consideration of standardization, highlighting, explanation, detection of surrounding obstacles, and suggestion of maneuver.

The fourth category is "Additional requirements for assistance in takeover situations with Dynamic movement" with two checklist items, which aim to provide assistance in the classified dynamic takeover situations.

The fifth category "Requirement for the driver's mutual understandability to the Takeover request" includes 2 checklist items, which targets the requirements to support the mutual understanding between drivers and systems.

The sixth category "Requirements for assistance after driver's takeover" has one item and addresses the requirements after the takeover process is finished.

6.3.4 Checklist Items

The six categories of the system requirements with the corresponding 16 checklist items are listed in the Appendix of this document. As an example, a screenshot of the checklist items in the second category can be seen in the Figure 11.

1	**Requirements for driving at SAE Level 3**	Yes	No	n/a
2.1	System checks the responsibility of the driving task at any moment	☐	☐	☐
2.2	System provides the driver with a permanent display of the traffic situation (e.g., birds eye view or using light display to show where the vehicle is going and where obstacles are dangerously close using a red light)	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

Fig. 11: Example of checklist item for evaluating human-machine interaction at SAE Level 3

Among these 16 checklist items, 5 of them have been mentioned by more than three experts. Three checklist items have been mostly suggested by four experts. They are the "2.3 System provides multi-modal warnings", "4.2 System displays the trajectory movement of the obstacle and shows the potential conflict of points of collisions on real time to the driver, e.g. with light display highlighting all the potential obstacles", and "5.1 The takeover request from the system can be fully perceived and understood by the driver". The checklist items 2.3 and 5.1 are connected, which means that most experts suggest to make use of multimodal warnings for takeover requests, in order to enhance the driver's perception and understanding of the takeover requests. In addition, experts have emphasized the importance of displaying the dynamic movements of the relevant obstacles and potential collisions in dynamic takeover situations.

Two checklist items such as "2.2 System informs the driver of the urgency of the situation (e.g. 5 s in advance for urgent situation and 10 s for non-urgent situation)" and "3.3 System highlight the relevant obstacles with (e.g. flags, marks or flashing lights)" have been suggested by half of the experts. It shows that half of all experts have agreed that highlighting the relevant obstacles and informing about the urgency of the situation could support the driver's takeover process.

6.3.5 Application of Checklist Item

The main purpose of this application is to simplify the evaluation process of a new vehicle concept, which can be used by international experts for the evaluation of human-machine interaction at SAE Level 3 in the context of automated driving. This service is provided online and cross-platform in a user-friendly environment. The scope of the application is, besides a short description of how to use it, the evaluation using the checklist that was developed throughout this project (see Appendix: The Checklist for evaluating Human-Machine Interaction at SAE Level 3), and a page for additional comments that will then be sent to a server or any receiver that has to be specified. The application can be described as a prototype application as a suggested way of providing the checklist online to any customer.

The application is web based and uses HTML, a CSS-Stylesheet, and a Javascript file with the JQuery package. The JQuery package was used to improve the user experience of the application, but can be removed for the final product. As mentioned, we propose that the application has three features: (1) the description of how to use it, (2) the main evaluation of the seven checklist categories, and (3) a field for additional comments in case something was unclear throughout the evaluation.

At the current state, this application is hosted on a private GitLab Pages site. The source code will be handed over together with the results of this project for further use. The application is in a prototypical state and should be prespecified according to the affordances. In the upper part of the application, we propose a navigation bar, that shows the progress and in which category the user is at the current moment (see Figure 12).

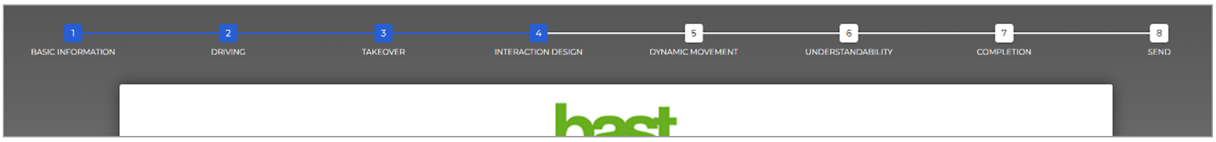


Fig. 12: Navigation and progress bar

There is a “previous” and “next” button to navigate through the pages (see Figure 13). The information that is provided by the user is saved immediately and is not lost while navigating back and forth. Therefore, the user will be free to go back to any previous category without losing any progress.

Fig. 13: Navigation buttons

The instruction page (see Figure 14) will provide a description of how to use and what is the purpose of this application. The current application should only be a suggestion of how to implement it, we therefore suggest to revise it and possibly add more basic information if needed.

After the instructional and basic information page, the seven categories of the checklist can be evaluated by experts. The interaction is optimized for the use on a computer or touch device, such as a tablet or mobile phone. To fill in the needed information, users have to rate the question on a 3-point Likert scale with the answer options “yes”, “no”, and “N/A” (not applicable) (see Figure 15).

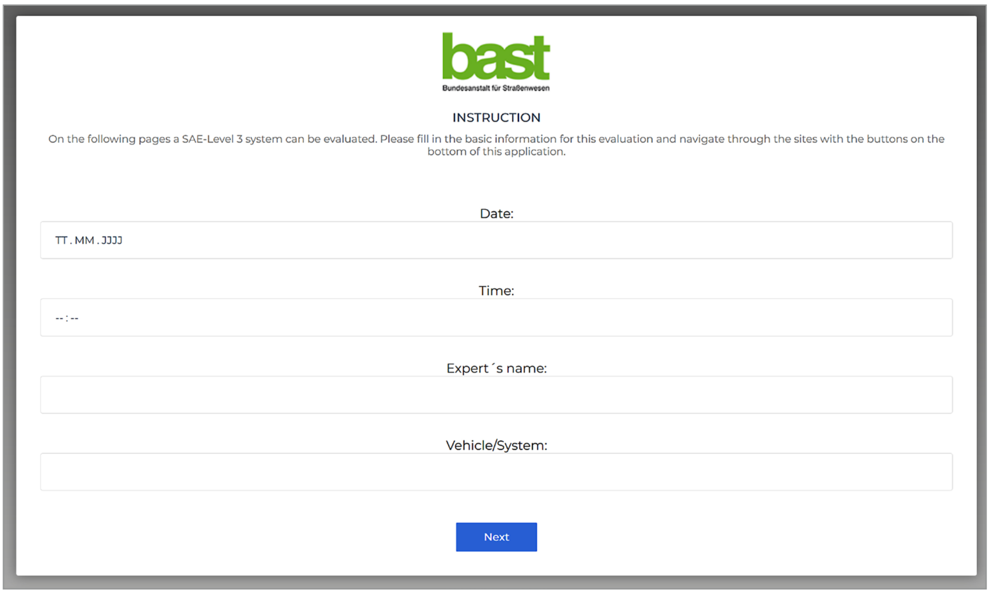


Fig. 14: Instruction page of the web application

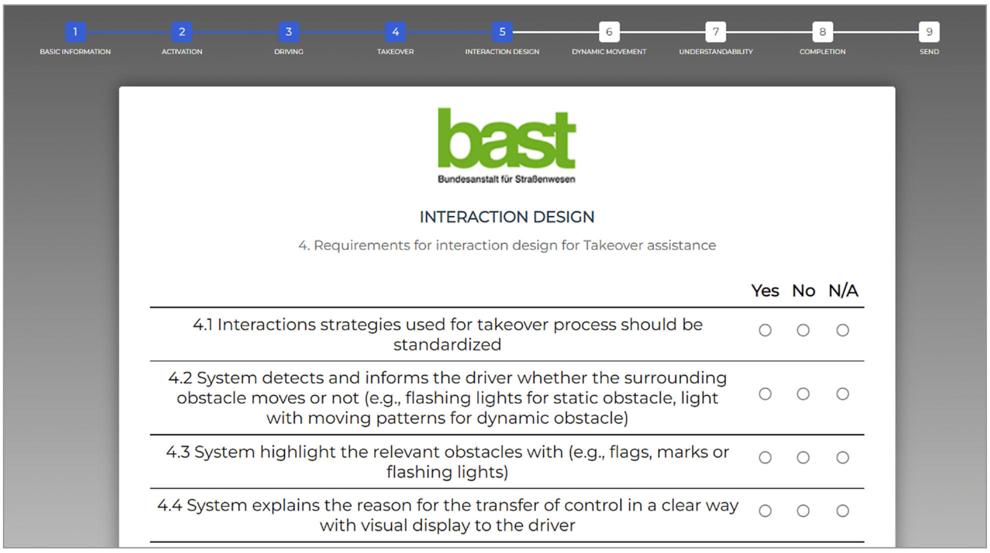


Fig. 15: Checklist items in category four “interaction design”

After evaluating all corresponding items in one category, an overall evaluation of this category (“fulfilled”, “satisfactory”, “not-acceptable”) (see Figure 16) and additional comments can be provided under this checklist.

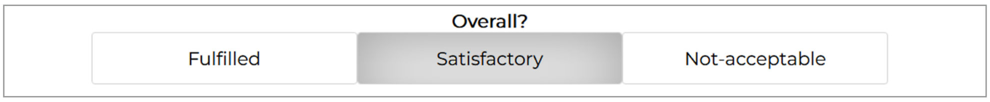


Fig. 16: Overall evaluation of this category

After evaluating the seven categories, the user can rate the whole system and has the option to specify further comments and send the results (see Figure 17) and provided information of the evaluation to a yet to be specified server address.

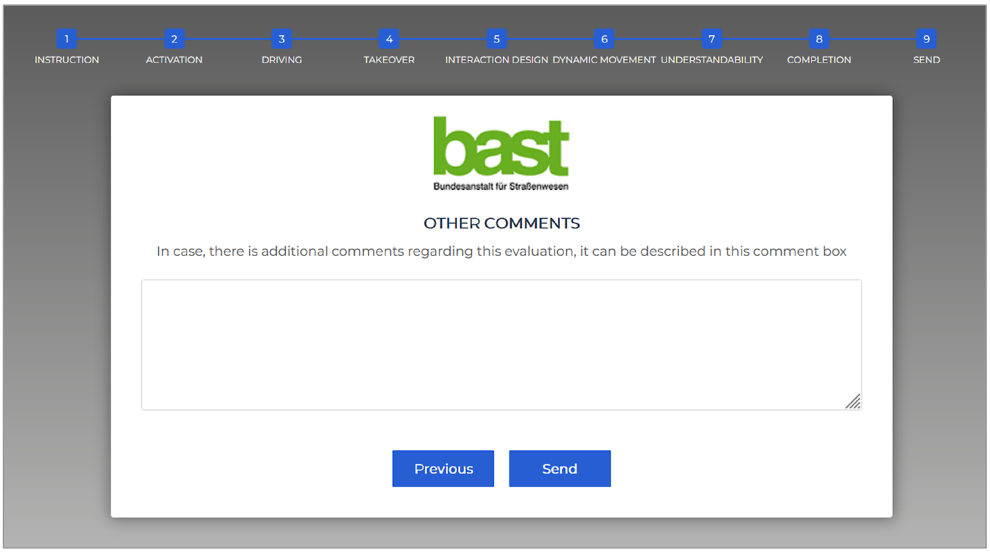


Fig. 17: Final page and send option

6.4 Summary

Based on the results of meta-analysis and expert interviews, 16 checklist items assigned into six categories have been developed to evaluate the safety of human-machine interaction at SAE Level 3. In addition, these checklist items have been further developed to a prototype application as a suggested way of providing an online version of the checklist, which can be easily accessible and used for evaluating the human-machine interaction at SAE Level 3 in the context of automated driving.

7 Conclusion

With the aim of identifying suitable indicators and criteria for evaluating the safe human-machine interaction for SAE level 3 systems up to 60 km/h in the context of automated driving, this research project has started with a focus group interview to identify relevant publication channels and list of keywords regarding indicators for the evaluation of human-machine interaction at SAE Level 3. Based on the identified list of keywords, literature reviews have been conducted to extract relevant publications from the identified publication channels. These selected papers have then been used for meta-analysis to study the influence of different takeover situations on takeover performance measured by the categories of takeover time, takeover quality and subjective workload. The results of meta-analysis indicate drivers' takeover performances are different in static and dynamic situations. Expert interviews have been conducted to help interpret the results of meta-analysis and develop checklist items. In the end, 16 checklist items assigned in six categories of system requirements haven been developed and can be used by international experts to evaluate the safety of the human-machine interaction of SAE Level 3 systems up to 60 km/h in production vehicles. This checklist has been further developed to an online application, which can be used as an easy-to-implement and efficient evaluation procedure in relation to the traffic safety relevant interaction quality of the systems.

The international experts have given valuable and professional opinions on understanding driver's takeover performance in different takeover situations, and they have also expressed their concerns about the possibility and timing of implementation of some automated functions in series production vehicles, which can be taken into account for the development of assistance systems for SAE Level 3 in the future. The contents of the checklist provide important insights into the development of advanced driver assistance systems and evaluation for the safe human-machine interaction in the context of automated driving at SAE Level 3. Furthermore, the suggested important issues addressing the human-machine interaction at SAE Level 3 from expert interviews and checklist items can be used to help develop further vehicle technology regulations at UN ECE level.

Reference

ERIKSSON, A. & STANTON, N. A. (2017). Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. Human Factors: The Journal of the Human Factors and Ergonomics Society, 59, 689–705. https://doi.org/10.1177/0018720816685832

GOLD, C.; KÖRBER, M.; LECHNER, D. & BENGLER, K. (2016). Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. Human factors, 58(4), 642-652.

HART, S. G. & STAVELAND, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press.

KERSCHBAUM, P.; LORENZ, L. & BENGLER, K. (2015, June). A transforming steering wheel for highly automated cars. In 2015 ieee intelligent vehicles symposium (iv) (pp. 1287-1292). IEEE.

KÖRBER, M. & BENGLER, K. (2014). Potential individual differences regarding automation effects in automated driving. In C. S. González González (Ed.), Proceedings of the XV international conference on human computer interaction (pp. 1–7). Association for Computing Machinery.

LEWIS, S. & CLARKE, M. (2001). Forest plots: Trying to see the wood and the trees. British Medical Journal 322: 1479–1480.

National Highway Traffic Safety Administration. (2017). Automated driving systems 2.0: A vision for safety. U.S. Department of Transportation.

NAUJOKS, F.; WIEDEMANN, K.; SCHÖMIG, N.; HERGETH, S. & KEINATH, A. (2019c) Towards guidelines and verification methods for automated vehicle HMIs. Transportation Research Part F: Traffic Psychology and Behaviour, 60, 121-136.

RADLMAYR, J. & BENGLER, K. (2015). Literaturanalyse und Methodenauswahl zur Gestaltung von Systemen zum hochautomatisierten Fahren [Literature Survey and Description of Methods for the Development of Highly Automated Driving]. Forschungsvereinigung Automobiltechnik E.V. https://www. researchgate.net/profile/Jonas_RADLMAYR/publication/292987120_Literaturanalyse_und_Methodenauswahl_zur_Gestaltung_von_Systemen_zum_hochautomatisierten_Fahren/links/56b46eac08ae1f8aa4549626.Pdf

Response Consortium. (2006). Code of Practice for the Design and Evaluation of ADAS. RESPONSE 3: A PReVENT Project.

ROSENTHAL, R. & DIMATTEO, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. Annual Review of Psychology, 52, 59–82. https://doi.org/10.1146/annurev.psych.52.1.59

SAE International. (2018). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles(J3016). https://www.sae.org/standards/content/j3016_201806/

SAE International. (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016) https://www.sae.org/standards/content/j3016_202104/

TEOH, E. R. (2020). What's in a name? Drivers' perceptions of the use of five SAE Level 2 driving automation systems. Journal of safety research, 72, 145-151.

WEAVER, B. W. & DELUCIA, P. R. (2020). A systematic review and meta-analysis of takeover performance during conditionally automated driving. Human factors, 0018720820976476.

ZHANG, B.; DE WINTER, J.; VAROTTO, S.; HAPPEE, R. & MARTENS, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. Transportation research part F: traffic psychology and behaviour, 64, 285-307.

Figures

Tables

Appendix

The Checklist for evaluating Human-Machine Interaction at SAE Level 3

Please fill in the following questions:

1. Date: _____ 2. Time: ______ 3. Expert's name: ______ 4. Vehicle/system: _____

1	**Requirements for driving at SAE Level 3**	Yes	No	n/a
1.1	System checks the responsibility of the driving task at any moment (mode awareness)	☐	☐	☐
1.2	System provides the driver with a permanent display of the traffic situation (e.g. birds eye view or using light display to show where the vehicle is going and where obstacles are dangerously close using a red light)	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

2	**Requirements for system's takeover request**	Yes	No	n/a
2.1	System announces a takeover within 10 seconds	☐	☐	☐
2.2	System informs the driver of the urgency of the situation (e.g. a warning with red colour with 5 s in advance for urgent situations and orange colour with 10 s for non-urgent situations)	☐	☐	☐
2.3	System provides multi-modal warnings (e.g. auditory, visual, tactile)	☐	☐	☐
2.4	The numbers of warning modalities increase with the urgency of the situation (single modality for non-urgent situations and multi-modals for urgent situations)	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

3	**Requirements for interaction design for Takeover assistance**	Yes	No	n/a
3.1	Interaction strategies used for takeover process should be standardized in terms of implementation (e.g. modalities of takeover request and driver's confirmation)	☐	☐	☐
3.2	System detects and informs the driver whether the surrounding obstacle moves or not (e.g. flashing lights for static obstacle, light with moving patterns for dynamic obstacle)	☐	☐	☐
3.3	System highlights the relevant obstacles (e.g. flags, marks in the left side mirror or flashing ambient lights in the front of the vehicle such as A-pillar or dashboard)	☐	☐	☐
3.4	System explains the reason (e.g. error code) for the transfer of control in a clear way to the driver e.g. with a visual display or auditory information	☐	☐	☐
3.5	System suggests possible maneuvers to the driver (e.g. visual display in urgent situations)	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

4	**Additional requirements for assistance in takeover situations with Dynamic movement**	Yes	No	n/a
4.1	System gives warnings to the driver in a simple, noticeable and adequate way (e.g. manufacturers style guides)	☐	☐	☐
4.2	System displays the trajectory movement of the obstacle and predicts the potential conflict of points of collisions on real time to the driver, e.g. with light display highlighting all the potential obstacles	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

5	**Requirement for the driver's mutual understandbility to the Takeover request**	Yes	No	n/a
5.1	The takeover request from the system can be fully perceived and understood by the driver (expected high situation awareness of the driver measured by the error rate)	☐	☐	☐
5.2	The driver expresses that he/she is ready and able to take back control in a clear way according to UN-R157,6.2.5.2	☐	☐	☐

Fulfilled	Satisfactory	Not-acceptable

Comments:

6	**Requirements for assistance after driver's takeover**	Yes	No	n/a
6.1	System checks the driver's steering performance to achieve steering stabilization by adaptive steering similar to Active Front Steering (e.g. BMW AFS)	☐	☐	☐

Comments:

Final Rating:

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

Schriftenreihe

Berichte der Bundesanstalt für Straßenwesen

Unterreihe „Fahrzeugtechnik“

2020

F 132: Handbuch Barrierefreiheit im Fernbuslinienverkehr
Boenke, Grossmann, Nass, Schäfer € 17,50

F 133: Lkw-Notbremsassistenzsysteme
Seiniger, Heinl, Bühne, Gail € 15,50

F 134: Stationär-Geräusch von elektrisch angetriebenen Fahrzeugen
Altinsoy, Lachmann, Rosenkranz, Steinbach € 19,00

F 135: Abweichungen von der akzeptierten Fahrleistungsschwelle in automatisierten Fahrsituationen
Voß, Schwalm € 18,00

2021

F 136: Kamera-Monitor-Systeme als Fahrerinformationsquelle
Leitner, Oehme, de Silva, Blum, Berberich, Böhm
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 137: Konzept für die Erzeugung eines ISO-konformen UML-Modells und Generierung eines GML-Applikationsschemas für DATEX II zur Verbesserung der Interoperabilität
Lauber, Steiger, Kopka, Lapolla, Freudenstein, Kaltwasser
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 138: Grundlagen zur Kommunikation zwischen automatisierten Kraftfahrzeugen und Verkehrsteilnehmern
Schaarschmidt, Yen, Bosch, Zwicket, Schade, Petzold € 16,50

F 139: Einfluss von Notbremssystemen auf die Entwicklung von Lkw-Auffahrunfällen auf Bundesautobahnen
Straßgütl, Sander € 14,50

F 140: Reibwertprognose als Assistenzsystem
Leschik, Sieron, Gregull, Müller, Trapp, Brandenburg, Haalman, Terpstra
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 141: Methoden für die Bewertung der Mensch-Maschine-Interaktion beim teilautomatisierten Fahren
Schömig, Wiedemann, Julier, Neukum, Wiggerich, Hoffmann € 18,00

F 142: Schräglagenangst
Scherer, Winner, Pleß, Will, Neukum, Stanglmayr, Bäumler, Siebke, Prokop € 14,50

2022

F 143: Unfallverletzungen in Fahrzeugen mit Airbags
Holtz, Heidt, Müller, Johannsen, Jänsch, Hammer, Büchner
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 144: Entwicklung eines Verfahrens zur Generierung eines Safety Performance Indikators aus der Bewertung von Euro NCAP
Bäumer, Hautzinger, Pfeiffer
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 145: Regeneration von Partikelfiltern bei Benzin- und Dieselkraftfahrzeugen
Langwald
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 146: Analysis of options for the creation of safety-related traffic information based on vehicle-generated data
Margalith, Sickenberger, Wohak
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 147: Automatische Notbremssysteme für Motorräder
Merkel, Pleß, Winner, Hammer, Schneider, Will
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 148: Analyse glättebedingter Unfälle von Güterkraftfahrzeugen mit mehr als 12 t zulässigem Gesamtgewicht
Müller, Thüring, Jänsch, Epple, Kretschmer, Gottwald, Oehring, Winkenbach
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 149: Evidenzorientierte Ableitung von sicherheitsrelevanten Grundszenarien für die Fahrdomäne Bundesautobahn
Weber, Eckstein, Tenbrock, König, Bock, Zlocki
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

2023

F 150: Fahrerassistenzsysteme für die Geschwindigkeitsreduzierung bei schlechten Bedingungen
Pohle, Günther, Schütze, Trautmann
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 151: Integration von öffentlichem und privatem Parkraummanagement
Höpping, Jonas, Becker, Krüger, Freudenstein, Krampe, Godschachner, Inninger, Scholz, Hüttner, Grötsch, Stjepanovic
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 152: On-Board-Diagnose (OBD) – Analyse der OBD in Bezug auf zukünftig verfügbare Emissionsdaten für die Periodische Technische Inspektion (PTI)
Hausberger, Matzer, Lipp, Blassnegger, Hametner, Prosenc
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

2024

F 153: Zusammenstellung geeigneter Sicherheitsindikatoren für die Bewertung der Mensch-Maschine-Interaktion von Level 3 Systemen
Yan, Pichen, Schmitz, Sklorz, Baumann
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

F 153b: Compilation of suitable safety indicators for the evaluation of Human-Machine Interaction of level 3 systems
Yan, Pichen, Schmitz, Sklorz, Baumann
Dieser Bericht liegt nur in digitaler Form vor und kann unter https://bast.opus.hbz-nrw.de/ heruntergeladen werden.

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

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

www.schuenemann-verlag.de

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

ISSN 0943-9307
ISBN 978-3-95606-771-6

www.bast.de