Individual assistance during automated driving through driver modelling

Vorgelegt von

M. Sc.

René Alexander Lotz

an der Fakultät V – Verkehrs- und Maschinensysteme

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

\- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende: Prof. Dr. Stefanie Marker

Gutachterin: Prof. Dr.-Ing. Nele Rußwinkel

Gutachter: Prof. Dr. Martin Baumann

Tag der wissenschaftlichen Aussprache: 06. Juli 2020

Berlin 2020

Abstract

The automotive industry has been following the objective of increasing assistance for improved usability and safety over the past decades. Currently, the focus of automotive development concentrates on implementation of automation, promising the possibility of vehicles holding temporary responsibility of the driving task, while the driver is free to focus on other tasks. However, especially in Level 3 automated driving, the human has to act as a fallback, if technical systems fail or reach limits. This transferal of responsibility in any direction can cause miscommunication, incomplete driver mental models and accidents if mismanaged.

Therefore, a number of questions regarding the human-machine-interaction have to be addressed, in order to guarantee safe transitions between humans and vehicles. Firstly, the spectrum of human capabilities demands definition in the context of Level 3 driving. For empirical investigation, this requires monitoring systems for vehicles in order to identify factors influencing driver capability. Secondly, humans need assistance in the transition phases between automation and manual control. Based on the available information for an automated vehicle, a prediction of driver reactions is necessary to understand how a driver can be assisted, to reduce mismanagement of information. This is relevant for individuals as well as all varieties of driver types. Especially professional drivers, e.g. truck drivers, will experience long exposures to automation. Through extensive driving experience, possibly leading to other interaction than non-professional drivers, assistance needs to cater to their requirements. This dissertation focuses on the evaluation of professional driver capabilities in the context of Level 3 take-over. It proposes different computational models that offer the possibility of predicting driver behaviour, in order to improve usability and offer adaptive assistance during transition phases for automated driving. Thereby, assistance is improved and communication between the driver and vehicle is enhanced, to reduce accidents.

In order to examine driver behaviour and capability for Level 3 automation, two highway scenario studies were directed, analysed and utilised for predictive computational driver models. The first study was conducted in a high-fidelity moving-base simulator, displaying critical take-over situations. Amongst others, analysis of the empirical data showed that premature visibility of take-over situations generated significantly quicker reaction times. The data gathered of 88 participants was later utilised to analyse different data-driven computational models to predict reaction times of drivers. Based on remote eye-tracking data, vehicular environmental information and non-driving related task interaction, four different machine learning algorithms were compared in a four-class classification problem. Information was obtained from 238 features, reaching a misclassification rate of 38% for support vector machines.

The empirical data of the simulator study was later validated in a test track study with a prototypic Level 3 vehicle. As premature visibility was a determining factor in the capacity of drivers to react quickly, expectancy was assigned as an independent variable. The statistical analysis displayed no significant effect of expectancy of different take-over situations on the reaction time. Overall, the 3σ-interval of *time to first reaction* of the simulator study ranged from 0 to 2.82 seconds, while the test track study displayed slightly quicker reactions up to 2.34 seconds.

Finally, based on the findings of the two studies, an adaptive assistance during automation was developed. This system is designed to monitor the surroundings of any automated vehicle through vehicular sensors. The system relates environmental metrics to situations in which individual drivers self-initiate take-over. Thereby, it acts as an example for a seamless assistance system that is capable of individualisation and learning to enrich the driving experience and safety of drivers.

Zusammenfassung

In den vergangenen Dekaden hat die Automobilindustrie stets das Ziel verfolgt, nutzerorientierte Fahrerassistenz weiterzuentwickeln, um eine verbesserte Usability und erhöhte Sicherheit zu ermöglichen. Derzeit verspricht dabei ein Technologiesprung zum automatisierten Fahren diese Aspekte weiter auszubauen. Dabei soll Fahrzeugen eine temporäre Verantwortung über die Fahraufgabe zugeteilt werden, während der Fahrer seine Aufmerksamkeit auf andere Tätigkeiten lenken kann. Speziell in Level 3, dem hochautomatisierten Fahren, muss der Fahrer allerdings als Rückfallebene fungieren, falls technische Systemgrenzen erreicht werden. Die Übergabe der Verantwortung kann dabei zu Missverständnissen, unvollständigen mentalen Modellen des Fahrers oder Unfällen führen.

Aus diesem Grund, ist die Adressierung verschiedener Fragen zur Mensch-Maschine-Interaktionen essenziell, um eine Steigerung der Sicherheit in Level 3 Übernahmesituationen zu ermöglichen. Erstens, muss das Spektrum von Fahrerfähigkeiten im Rahmen des hochautomatisierten Fahrens erforscht und definiert werden. Zur empirischen Erforschung sind hierzu an das Fahrzeug angebundene Sensoren notwendig, um einflussreiche Faktoren zu erkennen und Informationen dem Fahrzeug zur Verfügung zu stellen. Zweitens, kann auf Basis dieser Erkenntnisse eine Fahrerassistenz entwickelt werden, die Fahrer in Übernahmesituationen unterstützen. Dabei können die sensorisch erfassten Informationen dazu dienen, eine prädiktive Aussage über Fahrerreaktionen zu generieren und somit Assistenz zur Verringerung von Missverständnissen zwischen Fahrer und Fahrzeug anzubieten. Dies ist sowohl individualisiert als auch generalisiert für verschiedene Fahrergruppen anwendbar. Unterschiedliche Fahrergruppen unterliegen dabei möglicherweise unterschiedlichen Einflüssen, z.B. LKW-Fahrer, die lange Automatisierungsphasen erleben könnten. Durch diese unterschiedlichen Ansprüche an eine Automatisierung zwischen Privatpersonen und professionellen Fahrern, kann sich auch der Anspruch an die Assistenz ändern. Diese Dissertation untersucht speziell die Fahrerfähigkeiten von professionellen LKW-Fahrern im Rahmen des hochautomatisierten Fahrens. Verschiedene rechnergestützte Modelle werden entwickelt, die die Möglichkeit einer Fahrervorhersage untersuchen, um die Usability adaptiv in Übernahmesituationen zu verbessern. Dabei wird stets das Ziel verfolgt, den Informationsaustausch zwischen Fahrer und Fahrzeug zu verbessern, um Übergangsphasen sicherer zu gestalten.

Diese Inhalte werden im Rahmen des hochautomatisierten Fahrens auf Basis von zwei empirischen Studien zu Autobahnfahrten untersucht. Die erste Studie fand in einem Fahrsimulator statt. Die Daten von 88 Probanden wurden anschließend zum Vergleich von vier verschiedenen rechengestützten Algorithmen genutzt. Hierzu wurden die Daten aus einer Blickerkennung, der Umgebungsinformationen und Interaktion des Fahrers mit Nebentätigkeiten in vier Machine Learning Algorithmen eingespeist, um ein Klassifikationsproblem zur Vorhersage von Reaktionszeiten mit vier Klassen zu lösen. Mit 238 verschiedenen Features als Eingangsdaten in eine Support Vector Machine, wurde eine minimale Misklassifikationsrate von 38% erreicht.

Die empirischen Daten aus dem Fahrsimulator wurden anschließend auf einer Teststrecke mit einem prototypischen hochautomatisierten LKW validiert. Als unabhängige Variable wurde der Faktor der Erwartung einer Übernahmesituation untersucht. Die statistische Auswertung ergab hierfür keinen signifikanten Effekt. Insgesamt ergab das 3σ-Intervall der *Time to first reaction* eine Reaktionszeit von 0 bis 2.82 Sekunden in der Simulatorstudie und 0 bis 2.34 Sekunden auf der Teststrecke.

Auf Basis der empirischen Daten entstand zusätzlich eine neuartige Ausführung eines Assistenzsystems. Die Betrachtung verschiedener Umweltvariablen durch Fahrzeugsensorik ermöglicht es hierbei Fahrsituationen zu identifizieren, die den Fahrer zu einer unaufgeforderten Übernahme verleiten. Das System identifiziert Situationen in denen der Fahrer Unsicherheiten durch speziell diese unaufgeforderten Übernahmen an den Tag gelegt hat, um frühzeitig die Entstehung dieser Situationen zu erkennen und den Fahrer vor unkontrollierten Handlungen zu schützen. Hierdurch wird eine kontinuierliche Assistenz dargestellt, die individualisierbar und selbstlernend sich an verschiedene Fahrer- und Umgebungssituationen anpassen kann und so eine neue Ausprägung von Assistenz zum hochautomatisierten Fahren darstellt.

Overview of Papers

This dissertation consists of four papers.

Paper No. 1

Lotz, A., Russwinkel, N. and Wohlfarth, E. (2019). Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F,* 64, pp. 532-551. doi:10.1016/j.trf.2019.06.008

Paper No. 2

Lotz, A. and Weissenberger, S. (2019). Predicting take-over times of truck drivers in conditional autonomous driving. In: Stanton N. (Ed.) Advances in Human Aspects of Transportation. AHFE 2018. *Advances in Intelligent Systems and Computing, 786*, pp. 329-338. doi:10.1007/978-3-319-93885-1_30

Paper No. 3

Lotz, A., Russwinkel, N. & Wohlfarth, E. (2020). Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks. Cognition*, Technology and Work*. doi:10.1007/s10111-020-00626-z (licensed under CC BY 4.0)

Paper No. 4

Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E. (2020). An adaptive assistance system for subjective critical driving situation: understanding the relationship between subjective and objective complexity. In: D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference (pp. 97-108).* Nantes, France. https://www.hfes-europe.org/wp-content/uploads/2019/10/Lotz2019.pdf

Contents

1 Synopsis

Automation has been introduced into many professions over the past century, causing humans to adapt to technological advances. While some tasks have been completely replaced by automation solutions, other work areas currently rely on new collaboration schemes between humans and machines (Romero, et al., 2017). Rudimentary benefits of substituting human interaction in a task with automation include the benefits of reducing repetition of monotonous tasks, increasing safety and precision.

However, some monotonous tasks have not been tackled, due to system complexity and superior human performance to adapt to new and dynamic situations. Driving is a prime example of such a task, where thus far human performance has exceeded technical possibility. Nevertheless, automation can potentially improve comfort and reduce human error, even in the driving domain through collaboration (Nyholm & Smids, 2018). This is especially the case in situations where long periods of driving cause a reduction of driver ability or driving situations induce fatigue, e.g. highway travel. However, some driving situations are difficult to automate technically and regardless of current technical maturity, humans exhibit higher performance. Especially in the distinction of human-machine-interaction during automation, a variety of well-established research linking automation levels to human performance has been produced (Sheridan, et al., 1978; Parasuraman, et al., 2000). Multiple factors need to be considered in order to differentiate which driving subtasks can benefit from automation with current technical systems.

In long-haul logistics, truck drivers operate the vehicle in highway scenarios for extensive periods. The truck domain thereby combines the unique aspects of professionals handling vehicles for multiple hours a day, possibly benefitting from automation. At the same time, novel European regulations address the ambitious goal of lowering traffic fatalities by 50% in ten years (European Commision, 2018). Through the development of advanced assistance systems, both the individual driver and societal benefit could be achieved. However, technical and usability questions regarding such assistance systems offering automation need to be investigated to ensure these global benefits.

On highways or federal roads the variability of situations, which an automation function needs to handle, is the most confined compared to other road types. Automation on highways therefore displays the highest potential of technical feasibility. Examples of automation functions being introduced into the market for this scope, include Audi’s AI traffic jam pilot (Audi, 2017) and Daimler Trucks Active Drive Assist (Daimler AG, 2019). By assisting drivers in this spectrum of their work, especially truck drivers, their behaviour will inevitably adapt to the new handling of the vehicle. Truck drivers could benefit from automation for long periods every workday if collaboration is managed correctly.

Although benefits seem to outweigh the disadvantages by introducing automated driving functions into vehicles, it also bears risks. Especially if automation cannot handle all facets of driving, driving situations will arise that require human input. This leads to the accepted theoretical concept of the ironies of automation (Bainbridge, 1983), in which humans are relied upon in situations in which peak performance is difficult. The driver, who at the most had to previously monitor the driving task, now has to act, to assist automation in decisions and/or task execution, without complete human overview over the task. This interpretation of automated driving is defined as Level 3 (SAE J3016). Short-term transitions between automation and human control are required, as the human functions as a fallback entity. Short-term requirements can induce the possibility of novel unaligned handling errors, which would not occur without automation available. Therefore, while automation might reduce errors due to inattention or drowsiness when automation functions are activated, new handling errors can occur in transition phases. This is especially the case, if the system is designed outside of the scope of human capability. The automation function should therefore be designed in a fashion to optimally incorporate the human in all automation phases. This especially includes those situations in which automation itself reaches technical limits. Ideally, the function should also assist in any transition, even if automation has reached its limitations.

Through the integration of automation in the driving domain, driver behaviour during automation and in the transition back to manual control will adapt compared to pure manual driving. Requiring a driver to transition between manual and automated vehicle control when automation fails, can have a profound influence on the way this behaviour is adapted with automation. Based on experience, daily exposition of situations, vehicle type and other influences, this behaviour might change differently for individuals. While the main goal of automated driving is to increase safety and comfort, ill-advised transition protocols can decimate all of these benefits and cause accidents. Depending on the human, the situation in which take-over is requested as well as external environmental factors, transition behaviour will differ (Vogelpohl, et al., 2016). Only by gathering information of driver behaviour during automated driving, transition protocols can be improved to accommodate individual driver's transition capability. This may include the current driver attention and interaction with tasks other than the driving task. Therefore, a good system design should always consider human variability, as strengths and weaknesses are never constant for any human. This follows the design guideline towards the question how to get the human and the machine to work together successfully (Dekker & Woods, 2002).

Three main aspects have to be considered when designing the transition phases between automation and manual driver control. What is a driver capable of managing physically and cognitively? What does a driver have to do? What does a driver optimally want or how should transition be designed for the driver? In order to consider different driver states and intentions, a method of generating this information is required.

Especially for the third question, an understanding of the human is required. Previously, only the human was able to monitor the machine, with the machine having no data input on human status during automated driving in Level 3. By continuously monitoring the driver and generating driver capability predictions, time critical transition phases can potentially be deescalated based on these prognoses. The prediction models thereby function as basis for the machine's understanding of the human, in order to adapt the transition phase and thereby maximize the collective human-machine output by reducing the risk of human error. The bilateral combination of human understanding of the machine and a machine's understanding of the human, can improve the ability of both partners, especially in time critical take-overs of Level 3. This thesis will not concentrate on the adaptations of any automation system, but on the methodology of forming prototypic driver prediction models that can aid this decision.

Two major empirical studies are conducted and presented in this thesis. They concentrate on the investigation of truck driver behaviour during time critical transition phases of Level 3 automation. By developing a driver monitoring system, physiological driver data is gathered during the course of these studies. There are different approaches that can result in methodologies allowing application of this knowledge to improve driver assistance in transition phases of automation. Based on the obtained empirical data, this thesis will investigate the application of data-driven computational modelling approaches. The predictions by the computational models shall function as a basis to predict driver behaviour in transition phases. Thereby, emerging from the first two abovementioned aspects regarding driver capability and transition demand, the third question addressing desired driver assistance can be developed.

The overall research question of this thesis combines all of these aspects into one elemental question. How can computational driver models improve the understanding of individual demand for assistance during automated driving transitions?

This thesis will address the three rudimentary questions introduced above, with the aim of ultimately focusing on individual driver assistance during transition between automation and manual control. Fundamentally, understanding of the human-machine-interaction during all phases of automated driving is required. This especially includes experimental research on driver behaviour during critical transition phases, i.e. take-overs, of Level 3 automation. Therein, empirical data forms the basis to answer the abovementioned three main questions. Different computational driver models will be developed and evaluated, which are designed to predict human driver behaviour adaptively and learn depending on human adaptation. With the insight gained, it shall be discussed what type of assistance a driver requires for Level 3 automation and in what form assistance systems may alter throughout the task, to design safe and adequate transition phases during automation. Through the adaptation of these systems, it is the goal of incorporating drivers into a fluid and seamless human-machine-interaction for automated driving. Thereby, the main objective is to increase the safety during automated driving transition, by understanding influences that reduce the transition capability of any driver.

In order for the driver to be guided back to the driving task, possible influencing factors hindering or facilitating a transition require investigation. These factors need to be considered in the computation models in order to make adequate predictions of driver behaviour. What types of influencing factors are there and how do they affect drivers during transition?

Multiple independent effects have been analysed in recent studies regarding automated driving, especially for self-driving passenger cars (Zhang, et al., 2019). The variety of influences range from vehicular to environmental as well as driver and human-machine-interface factors (Vogelpohl, et al., 2016). The variety of factors within these groups characterizes the large spectrum of impacts, which require an adequate structure. The comparison of the plethora of studies is complex, as experimental setups and interaction effects can cause false conclusions and necessary detail is often not reported.

In this context, it should also be noted that most of the abovementioned studies were conducted in simulator experiments. The comparability of simulator studies has been discussed in the literature with regard to risk perception (Carsten & Jamson, 2011; Bellem, et al., 2016). Actual highly critical take-over situations are extremely difficult to realize in empirical studies due to ethical restrictions, possible injuries and property damage. A comparison of take-over behaviour for Level 3 between a simulator and real-world study, in the truck context, is absent in current published research. The transferability between the virtual and real-world environment is another aspect, which requires investigation. It is important to consider the differences between simulators and real-world application to guarantee correct causality in the system design and validate empirical data.

The investigation of all combinations of influencing factors is impossible in controlled experimental environments. Some individual isolated effects might not be consistent for all driving situations. This calls for a data-driven approach that introduces a methodology of collecting all this data of the different influencing factors in all situations and embedding predictions into the automation function. In order for a data-driven approach to function, sensors need to be implemented, to accomplish driver and situational monitoring. Based on these data inputs, the information needs to be aggregated, to ultimately assert driver take-over capability or not. Typically, driver take-over capability is expressed in quantitative driver reaction times. This addresses the first two abovementioned aspects of what a driver is physically and cognitively capable of managing in transition and what has to be done. It disregards what the driver wants during transition, as this is difficult to quantify in singular metrics.

Data-driven computational models offer the possibility of accumulating different factors and analysing sensor data to consider underlying effects. All incoming data thereby needs to be referenced to an output value. Depending on which of the three abovementioned research aspects should be answered with such a model, different output values can be considered. When concentrating on the first two aspects, how can the driver manage and how transition is accomplished, the investigation of reaction times is a viable solution to express driver capability. The primary question that needs to be answered with regard to these two questions is, how much time does the driver require for take-over?

This question examines what can be expected from a driver when a take-over warning is initiated and the driver has to commence with the driving task. Thereby, take-over requests are typically implemented with a multimodal approach of optical-acoustic warnings (Bellet, et al., 2019). This quantitative approach is interesting from an engineering perspective, as resulting information can be utilized for the system design of the automation function according to usability guidelines (DIN EN ISO 9241, 2006). Interested in the measurable output a driver generates at take-over, the automation function needs to be designed according to human cantered guidelines. While keeping the technical applicability in mind, this technical solution for take-over should always be controllable by human drivers. Therefore, reaction times are essential. Subjective perception of take-over criticality and their effect on driver workload (Wickens, 2008) are not considered in this approach. However, such abstract psychological constructs apply and might affect reaction times subliminally. A first set of computational models should focus on driver take-over capability based on reaction times. This has been widely adapted in the literature to express influences of different factors to better understand the mechanisms of these effects on human cognition (Petermann-Stock, et al., 2013; Feldhütter, et al., 2016; Gold, et al., 2016; Zeeb, et al., 2016).

In any take-over situation characterized through system limits, the criticality can vary immensely through variation of interface, environment, driver and vehicle parameters. Depending on the environment and current trajectory, criticality of the situation depends on the time to collision with any object in the surrounding (Damböck, 2013). Therefore, situations can be extremely critical, if objects in the current trajectory path are imminently in front of the vehicle at the point in time when a warning is initiated, i.e. request to intervene. These take-overs become particularly difficult and reduce the amount of time drivers have to react, forcing them to act at peak performance. While not all take-over situations in Level 3 automation will reach these levels of criticality, the investigation of such situations can quickly frame the limitations of human capability. By analysing these situations specifically, the range of human actions can be described, which in turn affect human-machine-interaction for all take-over situations.

It should also be noted that non-driving related tasks during automation are allowed according to the definition of Level 3 automated driving. This thesis will limit all investigation to tasks on fixed digital screens in the vehicle, disregarding tasks executable on handheld devices. Research has shown that cognitive load, glance behaviour and take-over performance deteriorates with handheld devices (Owens, et al., 2011).

If research of the considered affecting factors reveals that safe transition between the self-driving vehicle and driver is not possible for a subset of situations, alternative solutions need to be implemented in the technical systems. This can include safe state mitigation of the vehicle, e.g. parking at the side of the road. Alternatively, the transition phase can also be adapted to improve the flow of information between human and machine, to ultimately increase human take-over performance. Universally, miscommunication and ironies of automation need to be obverted for self-driving vehicles, as drivers and individuals in the environment could be harmed through inadequate take-overs.

So is it possible for computational models to find underlying interactions between driver behaviour and influencing factors? Especially in the context of this thesis, this can be summarized to the following research question. Can a system identify driver availability and take-over ability, in order to possibly predict human driver interaction?

This would allow for an adaptive assistance system that considers individual and contextual differences. Thereby the third research question is addressed; what does the driver ideally want at take-over? By sharing the responsibility of the driving task, the driver assistance system also needs to adapt to these circumstances. Especially in situations in which a transition of responsibility is required due to technical system boundaries, information of the other partner handling the driving task is crucial for higher safety and usability.

The most obvious possibility of allowing the vehicle to form an understanding of the driver is by fitting it with sensors monitoring the driver (Braunagel, et al., 2017). By combining different intrusive and non-intrusive sensors, a driver monitoring system can be implemented to gather information during critical take-over situations. In the scope of this thesis, only remote sensors will be utilized for driver monitoring.

The approach of classifying take-over reaction times based on features generated from vehicular environmental data and non-intrusive remote driver monitoring cannot explain the causation of human behaviour. This methodology is inherently different to computational models enabling cognitive modelling, such as ACT-R (Anderson, et al., 2004). Deng et al. (2019) give an example of modelling take-over reaction times with a derivative of ACT-R (QN-ACTR) as a possible solution of generating knowledge on the causation of driver behaviour. A disadvantage of utilizing a cognitive architecture during the analysis of driver behaviour in new scope of driving is that the solution would not be scalable to other situations. Therefore, an explanation as to the reasoning behind the driver behaviour is not given in this thesis. Rather, the empirical data collected serves as the basis for future research of cognitive modelling. Without this data, experienced modellers of cognitive architectures will not have the basis to validate their models. If different effects are identified, different models can be tested against each other to describe possible behaviour causation. Data-driven approaches are more compatible in scalability while utilizing the same methodology and only adjusting data inputs and outputs. Although human cognition is not explained explicitly in the presented computational models, their validity is based on the capability to predict outcomes based on driver behaviour. Explanations on driver behaviour are therefore discussed and inferred, based on empirical data of this thesis.

A final discussion point of this thesis regards predictive assistance systems. Based on a vehicle enabling Level 3 automated driving, what information does a driver need during critical take-over? How should an advanced driver assistance system be configured, to assist the driver during automated driving?

The prime objective of a model-based prediction is to generate information allowing the adaptation of a system, based on validated information. Prediction functions are ideally based on data and can increase safety and/or usability. While there are multiple models predicting driver behaviour to improve aspects of the driving task, e.g. fuel efficiency (Bender, et al., 2014), Level 3 systems have the added difficulty of trying to predict measures that influence driver's safety. Ideally, systems would be capable of adapting their assistance based on the driver's needs. However, high accuracy is paramount for such assistance systems, to reduce the possibility of generating a situation that is worse than without adaptive assistance. Therefore, this thesis also investigates the possibility of a new type of assistance system, with the objective of predictively warning a driver individually based on individual interpretation of driving situations. This can improve collaboration between human and machine during automated driving and progress individualized advanced driver assistance.

Summary of papers

Four papers are included in this dissertation, three of which have previously been published (**Paper No. 1:** Lotz, Russwinkel & Wohlfarth (2019); **Paper No. 2:** Lotz & Weissenberger (2019); **Paper No. 3:** Lotz, Russwinkel & Wohlfarth (2020); **Paper No. 4:** Lotz, Russwinkel, Wagner & Wohlfarth (2019)). An overview of the contents of these four publications are introduced in the following, demonstrating how the publications address the previously introduced research questions.

Paper No. 1: Lotz, A., Russwinkel, N. & Wohlfarth, E. (2019). Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F*, 64, pp. 532-551.

Paper No. 1 presents results of a driving simulator study on conditional automated driving (Level 3; SAE J3016, 2018), conducted in a moving based simulator with professional truck drivers. The emphasis of the paper is to investigate critical take-over situations with the focus of identifying the influence of learning, non-driving related tasks, automation duration and type of take-over situation on behavioural driver reactions.

When considering Level 3 driving, an automation function is defined with certain system boundaries in which a driver needs to regain control. The paper accentuates that there are two possible approaches to specify systems in order to adequately consider driver reaction times for warning times. The first approach consists of globally warning times, which do not consider individual driver reaction times that can depend on influencing factors. A second approach proposes individualized reaction times that are considered based on influencing factor types. This can lead to an adaptive automation function based on influencing factors. In order to realize the possibility of such a driver assistance system, the effects different influencing factors cause need to be identified. This is the primary aim of Paper No. 1.

The paper explores two main aspects. Firstly, what type of behavioural reactions do professional truck drivers exhibit in critical take-over situations? Secondly, how quick are said reactions? This closes the gap between large quantities of research on take-over scenarios in the passenger car context to the truck context. Comparisons and explanations for variations are also drawn to results in the car context.

As outlined above, the paper limits the investigation to critical take-over situations. Through longer times to collision in take-over investigation, drivers are not forced to react close to their maximal capabilities. The expectation is formed that if longer times to collision were allowed, the driver would organize this time in any way they see fit. By presenting difficult take-over situations, the professional drivers should act at peak performance to circumvent critical situations. This allows the investigation of minimum reaction times and permits to identify what capability is to be expected of professional truck drivers in such situations. The gathered information is extremely valuable, as an adaptive automation function could identify which criticality would exceed driver capabilities and alternative strategies could be engaged promptly by a system, to increase safety.

A total of 88 professional truck drivers generated 768 take-overs. Two different non-driving related tasks are presented. One task consists of a video where no motoric interaction is required, while the other is a game requiring declarative knowledge and motoric interaction. Through the presentation of nine take-over scenarios, learning between the trails is also addressed. Gaze behaviour shortly before and after take-over is considered to investigate driver behaviour in critical take-overs.

The empirical results expose that the type of non-driving related task and automation duration have no significant effect on take-over times within the limited duration considered. Learning between trails is observed, especially in the first three take-overs. Without respect to the considered independent variables, the reaction times of this study are lower than those of published passenger car studies. The differences in take-over situation types is also discussed, finding that premature visibility of situations causes exceptionally quicker reaction times.

The analysis of gaze behaviour at take-over also shows that glances to the rear were delayed in critical take-over situations, even though lane switching was required in one situation. It seems that take-over situations were successfully designed as highly critical, as some drivers did not achieve safe manoeuvring between lanes. However, basic success criteria such as obverting accidents were handled correctly.

This comprehensive study forms the basis for the analysis of predicting driver take-over capability through data-driven computational models. Statistical analysis of the abovementioned independent study variables allows the in-depth discussion of these well-established influencing factors, compared to published Level 3 studies. However, the data holds further features that can be analysed in data-driven computational models to investigated different approaches of modelling driver take-over capability. The data obtained from this study serves as the input data for Paper No. 2. A total of 755 of 768 take-overs were utilized for Paper No. 2, as explained in further detail in the following section.

Paper No. 2: Lotz, A. & Weissenberger, S. (2019). Predicting take-over times of truck drivers in conditional autonomous driving. In: N. Stanton, (Ed.) Advances in Human Aspects of Transportation. AHFE 2018. *Advances in Intelligent Systems and Computing,* 786, pp. 329-338.

Based on the data gathered in Paper No. 1, Paper No. 2 presents a comparison of machine learning mechanisms to predict take-over reaction times. Data considered consist of features generated from eye-tracking, user interaction with non-driving related tasks and vehicular data during Level 3 automation. A total of 238 different features serve as input for the algorithm comparison.

Paper No. 2 specifically explores if machine learning methods generate beneficial predictions of reaction times based on the abovementioned data. A main aim of this paper is to investigate the ability to classify how quickly a truck driver will react based on features and information perceivable through remote sensors. This specifically addresses the two research questions regarding driver ability to manage take-over and the extent of successfully handling take-over. The focus is strictly based on remote driver monitoring data and environmental vehicular data from Paper No. 1. By exploring this research question and finding promising prediction methods, automated vehicles could generate a prediction of human-machine-interaction based on remote sensors. This would again make an adaptive automation function possible. Thus, assisting a driver in the complex and difficult transition phase of regaining control of the vehicle would possibly cause less human error. Alternatively, if human reactions were to be too slow, based on the prediction method, fallback strategies could be initialized promptly. The analysis of the computational models thereby also addresses the differences in accuracy of an online and offline prediction system, having different features available during or after take-over.

A clear limitation of analysing machine learning algorithms in this publication is the available dataset. Typically, data-driven approaches require a large dataset for the definition of any prediction model. The results of this comparison are limited to this extent. A total of 755 take-over situations were utilized from the study presented in Paper No. 1. Four different machine learning algorithms are compared in Paper No. 2. Included in this selection are two methods that allow weighting of input features, while two methods are rigid in their weighting scheme. It is examined how information that is available after take-over improves predictability of driver take-over reaction times through offline prediction. Apart from the investigation of fundamental machine learning methodology for this applied problem, the influence of singular features on the accuracy of reaction time prediction are investigated. The spectrum of reaction times is partitioned into four classes.

The results display a clear benefit of the offline dataset, which consists of additional information generated after take-over. Due to more information being available after take-over, accuracy of all algorithms improve. This benefit mostly stems from behavioural eye-tracking data that bears a lot of information in the take-over process. Eye movements towards the driving scene are generated at take-over and thereby can only be calculated after a take-over request is initiated, making the feature part of the offline dataset. At the same time the data gathered in Paper No. 1 did not suffice for the definition of clear reaction time classes. Inter- and intra-individual differences were too large for clear identification of influencing factors by all machine learning algorithms.

This leads to the conclusion that either more information is required to achieve an accurate application of the presented algorithms or that the features generated do not offer enough information for precise prediction. Methodologically, the algorithms presented in this paper would not need to be altered in order for more empirical data to be analysed and reach a conclusion on this question. This highlights the main benefit of scalability of data-driven computational models for driver take-over prediction. The complete efficiency of the models can, however, only be precisely defined with more data that are available.

Paper No. 3: Lotz, A., Russwinkel, N. & Wohlfarth, E. (2020). Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks. *Accident Analysis & Prevention*. doi: 10.1007/s10111-020-00626-z

The purpose of Paper No. 3, in the scheme of this thesis, is to validate the data gathered in the study of Paper No. 1 outside of a simulator. A prototypic vehicle was fitted with an automation function, allowing the application of non-driving related tasks during driving on a test track. Thereby, a prototypic Level 3 vehicle was developed, to investigate further influencing factors with reduced criticality, to ensure safe operation. The driver monitoring sensor setup of Paper No. 1 is reutilized for this experiment. All participants are professional truck drivers and the experiment is conducted on a private test track with limited unregulated traffic.

Two primary research questions are addressed. Firstly, is the data gathered in the driving simulator environment of Paper No. 1 comparable to real-world environments? Secondly, how does the expectation of a take-over situation influence the driver in a real-world critical take-over situation?

To guarantee a high familiarity with the vehicle, i.e. automation function and optical-acoustic warning signal at take-over, a prolonged training phase is introduced. Drivers complete five laps (2 training, 3 experimental) of a 12 km round circuit, during which multiple take-overs are experience. Parallel, this procedure reduces the possibility of recording a similar learning effect in the first few trials, as experienced in the study conducted in Paper No. 1. Three types of take-over situations with varying criticality and expectancy are compared. With the possibility for drivers to predict take-over for one of these situations, the effect of take-over expectancy is investigated. Additionally, take-over criticality is examined.

The experimental procedure on the test track requires the additional consideration towards uncontrollable environmental factors such as weather and daytime. Each driver drove a distance of 60 km on a high speed oval test track (ATP Automotive Testing Papenburg GmbH, 2018). The drives were conducted continuously between early morning hours 8:00 and night time 19:00 in December. This caused a small group of drivers to drive during the night-time. No effect towards reaction times was found for this uncontrolled group of drivers. Similarly, precipitation caused no significant difference in reaction times.

Results display that the two levels of criticality present a low significant effect. This could be caused by a kinaesthesia reaction, as criticality was induced through lateral acceleration of the steering. Finally, the overall expectancy a driver can hold towards a take-over resulted in no significant quicker reactions of *time to first reaction* nor *time to hands on steering*.

Finally, the comparability of simulator data and test-track data from real-world driving are associated. Limitations of both scenarios are discussed, focusing on the application of a Level 3 system on public roads.

Paper No. 4: Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E. (2020). An adaptive assistance system for subjective critical driving situation: understanding the relationship between subjective and objective complexity. In: D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference*. Nantes, France.

Paper No. 4 presents an adaptive driver assistance system that continuously monitors the surrounding environment of an ego-vehicle in which a driver has a Level 3 automation activated. It provides a framework and conceptual simulation model that offers an answer to how drivers can be assisted in driving situations, which they subjectively find too complex for an automation function to handle. This type of assistance system relies on the prediction of subjective perception of the driving environment and trust in automation, an inherently different aspect than previous reaction time metrics. The concept forms a new type of driver assistance system, warning not of technical malfunction or system boundaries, but providing a suggestion on subjective perception of the environment. Notably, the introduced conceptual model only functions in situations in which the driver does not hold control of the lateral and longitudinal driving task and self-initiates take-over. Paper No. 4 discusses how such a driver assistance system gathers information on drivers and infers their subjective awareness of the environment, without the necessity of sensors monitoring the driver.

A second major aim of Paper No. 4 is to provide a platform for future investigation of perceived driving complexity. The differentiating concept of objective and subjective complexity is introduced in this paper. Objective complexity is thereby defined as the mathematical description of a driving environment in which the ego-vehicle forms the origin of the coordinate system and metrics such as distances, velocities and collision times are referenced between objects. Paper No. 4 defines subjective complexity as the perceived complexity of a human driver, including psychological influences such as workload, fatigue and situation awareness.

The complexity prediction system introduced in Paper No. 4 utilizes vehicular sensor data to monitor the direct periphery of an ego-vehicle. Through a signal filtering and subsequent calculation of kinematic variables for each object in the environment, a mathematical description of the environment is determined. Based on this description, critical objects are identified and compared to previous traffic situations at take-over. When a driver intervention occurs the system records the kinematic environmental variables of the critical objects and thereby learns driving situations, which the driver feels the need to take-over control. A feature of this methodology is the possibility of having a fully adaptive assistance system catering to subjective and environmental variation. The complexity prediction system intrinsically captures these traits.

The conceptual assistance system is evaluated with a rudimentary database, simulating driver request for assistance when sideswipe manoeuvres occur during Level 3 automation. A qualitative analysis of the introduced assistance system displays a preliminary high functionality of the concept based on a training set of 105 and test set of 13 sideswipe manoeuvres. Further data of self-initiated take-overs is required for a comprehensive analysis of the functionality.

Assistance introduced in Paper No. 4 presents a possibility of collecting data on what a driver wants during take-over. The level of assistance a driver wants is not defined, however, when assistance is relevant from a driver's point-of-view is documented through the system. This is important, as different assistance models can then be developed to improve usability and support drivers in these situations. Ultimately, this offers the human-machine-interaction to be adapted towards individual assistance and thereby reduce errors occurring at self-initiated take-over.

Bibliography

Anderson, J. R. et al., 2004. An Integrated Theory of the Mind. *Psychological Review*, 111(4), pp. 1036-1060.

ATP Automotive Testing Papenburg GmbH, 2018. *Automotive Testing Papenburg.* [Online]
Available at: http://atp-papenburg.de
[Accessed 20 02 2019].

Audi, 2017. *Audi Technology Portal.* [Online]
Available at: https://www.audi-technology-portal.de/en/electrics-electronics/driver-assistant-systems/audi-a8-audi-ai-traffic-jam-pilot
[Accessed 26 02 2019].

Bainbridge, L., 1983. Ironies of automation. *Automatica*, Issue 6, pp. 775-779.

Bellem, H. et al., 2016. Can We Study Autonomous Driving Comfort in Moving-Base Driving Simulators? A Validation Study. *Human Factors*, 59(3), pp. 442-456.

Bellet, T. et al., 2019. From semi to fully autonomous vehicles: New emerging rist and ethico-legal challenges for human-machine interactions. *Transportation Research Part F Traffic Psychology and Behavior*, May, Issue 63, pp. 153-164.

Bender, F. A., Uzuner, H. & Sawodny, O., 2014. An Adaptive Driver Model for Driving Cycle Prediction in the Intelligent Truck. *Proceedings of the 19th World Congress The International Federation of Automatic Control*, 24-29 August, pp. 10682-10687.

Braunagel, C., Rosenstiel, W. & Kasneci, E., 2017. Ready for Take-Over? A New Driver Assistance System for an Automated Classification of Driver Take-Over Readiness. *IEEE Intelligent Transportation Systems Magazine*, 9(4), pp. 10-22.

Carsten, O. & Jamson, A. H., 2011. Driving simulators as research tools in traffic psychology. In: B. E. Porter, ed. *Handbook of traffic psychology.* s.l.:Academic Press, pp. 87-96.

Daimler AG, 2019. *Daimler Global Media Site.* [Online]
Available at: https://media.daimler.com/marsMediaSite/ko/en/43216408
[Accessed 08 July 2019].

Damböck, D., 2013. *Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme,* München: Technische Universität München, Lehrstuhl für Ergonomie.

Dekker, S. W. A. & Woods, D. D., 2002. MABA-MABA or Abracadabra? Progress on Human-Automation Co-ordination. *Cognition, Technology & Work*, 4, pp. 240-244.

Deng, C., Cao, S., Wu, C. & Lyu, N., 2019. Modeling Driver Take-Over Reaction Time and Emergency Response Time using an Integrated Cognitive Architecture. *Transportation Research Record: Journal of the Transportation Research Board*, 16 July.

DIN EN ISO 9241, 2006. *Ergonomics of Human System Interaction,* s.l.: International Organization for Standardization.

European Commision, 2018. *General Safety Regulation - 2018/0145,* Brussels: Regulation of the European Parliament and of the Council.

Feldhütter, A., Gold, C. & Bengler, K., 2016. How the Duration of Automated Driving Influences Take-Over Performance and Gaze Behavior. *Kongress der Gesellschaft für Arbeitswissenschaft*.

Gold, C., Korber, 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), pp. 642-652.

Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E., 2020. An adaptive assistance system for subjective critical driving situation: understanding the relationship between subjective and objective complexity. In: D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference*. Nantes, France.

Lotz, A., Russwinkel, N. & Wohlfarth, E., 2020. Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks. *Accident Analysis & Prevention*.

Lotz, A., Russwinkel, N. & Wohlfarth, E., 2019. Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F*, 64, pp. 532-551.

Lotz, A. & Weissenberger, S., 2019. Predicting take-over times of truck drivers in conditional autonomous driving. *Advances in Intelligent Systems and Computing*, pp. 329-338.

Nyholm, S. & Smids, J., 2018. Automated cars meet human drivers: responsible human-robot coordination and the ethics of mixed traffic. *Ethics and Information Technology*, pp. 1-10.

Owens, J. M., McLaughlin, S. B. & Sudweeks, J., 2011. Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention*, May, Issue 43, 3, pp. 939-947.

Parasuraman, R., Sheridan, T. & Wickens, C., 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Issue 30, pp. 286-297.

Petermann-Stock, I., Hackenberg, L., Muhr, T. & Mergl, C., 2013. Wie lange braucht der Fahrer - eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. *6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren*.

Romero, D. et al., 2017. The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems. *Advances in Production Management Systems. Initiatives for a Sustainable World*, March, pp. 977-686.

SAE J3016, 2018. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.*

Sheridan, T., Verplank, W. & Brooks, T., 1978. *Human and Computer Control of Undersea Teleoperators,* Cambridge, MA: MIT Press: Man-Machine Systems Laboratory, Department of Mechanical Engineering.

Vogelpohl, T. et al., 2016. *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung,* Berlin: Gesamtverband der Deutschen Versicherungswirtschaft e.V..

Wickens, C. D., 2008. Multiple Resources and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 1 June, 50(3), pp. 449-455.

Zeeb, K., Buchner, A. & Schrauf, M., 2016. Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, 92, pp. 230-239.

Zhang, B. et al., 2019. Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F*, 64, pp. 285-307.

2 Paper No. 1

Lotz, A., Russwinkel, N. and Wohlfarth, E. (2019). Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F,* 64, pp. 532-551.
doi:10.1016/j.trf.2019.06.008

Transportation Research Part F 64 (2019) 532–551

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Response times and gaze behavior of truck drivers in time critical conditional automated driving take-overs

Alexander Lotz [a,b,*], Nele Russwinkel [b], Enrico Wohlfarth [a]

[a] Daimler AG, TP/VES HPC T332, 70546 Stuttgart, Germany
[b] Technische Universität Berlin, Marchstr. 23, Sekr. MAR 3-2, 10587 Berlin, Germany

A R T I C L E I N F O

Article history:
Received 4 April 2018
Received in revised form 20 June 2019
Accepted 23 June 2019

Keywords:
Conditional autonomous driving
Driving simulator
Non-driving related tasks
CAD duration
Take-over situations
Truck

A B S T R A C T

The desire to enable conditional automated driving (CAD) in the near future, entails the challenge to manage drivers' safe transitions from automation back to manual control. Several factors have been considered in recent years in the passenger car context, while the truck has largely been disregarded. For the first time take-over behavior of heavy-duty truck drivers in time critical take-overs is considered in CAD research. This study analyzes the effect of non-driving related tasks, CAD duration, take-over situations and number of take-overs on reaction times of truck drivers. Gaze behavior was tracked with a remote eye-tracker; reaction times and driver interaction during CAD drives was recorded and analyzed. Two different non-driving related tasks were presented in nine unique take-over situations, while also controlling for the duration of CAD. Contrary to assumption, no influence of non-driving related tasks or CAD duration on reaction times is found. Notably, different reaction times are recorded due to the nine unique take-over situations. Finally, it is shown that our take-over times decrease over the course of the experiment and are far lower than other published reaction times (M = 1.35 s) in the passenger car context. The findings are discussed and implications with regard to other published studies are drawn.

1. Introduction

Current advanced driver assistance systems (ADAS) generate the possibility of automated driving. While the vehicle is assigned to safely navigate through the environment for this driving scenario, conditional automated driving (Level 3; SAE J3016, 2018) requires the driver to transition back to full control of the vehicle once system limits are reached. Due to the complexity of sharing the driving task, car manufacturers are currently investigating the implementation of Level 3 on highway and interstate roads. Although driving on a highway may occur at much greater average speeds, the roads and behavior are extremely regulated, making the technical controllability simpler than in urban scenarios. Higher automation levels (Level 3 & 4) allow drivers to shift their attention away from the driving task (Gasser, Arzt, Ayoubi, Bartels, Bürkle, & Lotz, 2012) and potentially focus on non-driving related tasks (NDRT). This transition phase between the driving and non-driving related task creates a time and safety critical challenge. Generally, the human perception of the environment and relevant factors requires time and can cause higher error rates (Allport, Styles, & Hsieh, 1994). Prolonged reaction times have also been identified specifically for increased automation levels in vehicles (Young & Stanton, 2007). The question that arises when designing a Level 3 system is; how much time does the driver need to process information in Level 3 and regain control

* Corresponding author at: Daimler AG, TP/VES HPC T332, 70546 Stuttgart, Germany.
E-mail addresses: rene_alexander.lotz@daimler.com (A. Lotz), nele.russwinkel@tu-berlin.de (N. Russwinkel).

https://doi.org/10.1016/j.trf.2019.06.008
1369-8478/

of the vehicle safely (Gold, Damböck, Lorenz, & Bengler, 2013)? Different aspects of take-over behavior have recently been investigated for the qualitative and quantitative description of take-overs. These aspects can broadly be categorized into four classes, *driver, environment, vehicle* and *human–machine interaction (HMI),* according to Vogelpohl, Vollrath, Kühn, Hummel, and Gehlert (2016), which will be described in Section 1.1.

Typically, if the Level 3 technical system is unable to solve a driving situation and the vehicle is travelling with any speed, take-over is time critical due to obstacles in its path. At higher speeds, obstacles or humans in the environment and/or dangerous roads also makes situations safety critical. Two different approaches can be adopted to define adequate take-over times. Either a globally valid fixed time is chosen for all take-over situations, or conditional times depending on the abovementioned factors *driver, environment, vehicle and HMI* are assigned. For the second approach, the *driver* and *environment* factors have to be monitored continuously to adjust the adequate take-over time accordingly, as *vehicle* and *HMI* factors usually remain constant within one vehicle. This is a paradigm shift in the traditional human-machine relationship in which the human no longer needs to monitor the state of the vehicle or machine, but the vehicle has to monitor the environment and the driver, roles are reversed. Some of the *driver* factors can be measured through remote measurement systems, while most cognitive factors cannot or require on-body sensors. Such determinable factors incorporate gaze behavior, physical movements of the driver and NDRT interaction (Ohn-Bar, Martin, Tawari, & Trivedi, 2014). Since understanding of the influencing factors on the driver's reaction is needed, better measurement techniques and algorithms are required that potentially predict the right time interval for secure human behavior (Lotz & Weissenberger, 2019). The second approach for the definition of adequate take-over times is a more elegant, however, technically more complex due to additional sensors and algorithms necessary for driver observation.

The aims of the study are to investigate influencing factors which possibly cause varying behavioral driver reactions. The independent variables chosen for this study are categorized as environmental factors and incorporate the temporal duration of automation and type of NDRT. In addition, it is our goal to identify if learning behavior can be established in the obtained data. The study aims also include the analysis with regard to take-over quality in temporal critical take-over situations. Gaze behavior during take-over will be consulted for the determination of visual reaction times as well as the construction of situation awareness as a measure of take-over quality. Within this study situation awareness refers to the perception, comprehension and projection of environmental objects according to Endsley (1995). Based on the findings, applicability of the abovementioned second approach is discussed.

Notably, "driver availability is not just an absolute description of the current driver state but a relative assessment that is influenced by the situational demands" (Marberger, Mielenz, Naujoks, Radlmayr, Bengler, & Wandtner, 2017). The underlying question that needs to be answered when defining take-over time reserves is, which factors influence reaction times, as previously stated by Gold (2013). While a global fixed take-over time might cause safety issues due to extensive complex situations, variable take-over times could also have a negative effect, by warning too late or misjudging situations leading to safety critical take-over. The identification of relevant factors to allow for variable warning times is paramount. Additionally, this study should reproduce some of the findings regarding the influencing factors automation duration and NDRT from the passenger car for the truck domain, closing the gap of research between domains. Strongly differentiating NDRT (e.g. reading on a fixed screen in the cockpit and playing an instrument) surely present different reaction times, due to different modalities being required for both tasks and objects causing physical hindrance in the take-over behavior. Due to the amount of possibilities, conducting validation of each feasible NDRT and its influence on take-over is impossible. It would be favorable if observation of temporal short behavioral driver operations, e.g. cascade of fixations or movements, before take-over situations exhibit evidence about take-over times. These smaller behavioral operations would increment to a large NDRT with operations being exchangeable between different tasks, i.e. fixation of a screen and touching the surface. Previous work has shown that information is present in the data, however, misclassification rates are too high for safe predictions of take-over times (Lotz & Weissenberger, 2019).

To investigate the four groups of factors by Vogelpohl et al. (2016), a simulator study was designed to determine if behavioral factors can be identified for differences in take-over behavior. Three relevant take-over times according to Damböck (2013) will be reported in the following sections. The take-over procedure typically incorporates the shift of visual attention towards the driving scene, forming motor readiness and showing a reaction at the steering wheel and/or ac-/decelerator pedals. This study presents results from the analysis of the four main influencing factors; *NDRT*, *situation type, trial* and *CAD duration*, to investigate the possibility of designing variable warning times. Eriksson and Stanton (2017) give an overview of 25 publications investigating take-over behavior in automated vehicles. Reaction times at the steering wheel or pedals vary between 1.14 and 15 s, with the majority of reaction times measuring between 2 and 3.5 s. Vogelpohl et al. (2016) cite a reaction time bandwidth of 3–8 s. These references will be consulted for comparison to the results obtained in this study to draw the bridge between the truck and passenger car domain. Table 1 holds acronyms utilized in this study.

1.1. Factors influencing take-over behavior

As mentioned previously, Vogelpohl et al. (2016) define four classes of influencing factors on take-over behavior that will be introduced and discussed in the following.

Table 1
Acronyms within publication for further reference.

Acronym	Definition
ADAS	Advanced Driver Assistance System
CAD	Conditional Automated Driving (Level 3)
HMI	Human-machine interaction
NDRT	Non-Driving Related Task
RtI	Request to Intervene (a.k.a. Take-over Request TOR)
TTC	Time to Collision
TTEoR	Time to Eyes on Road
TTFR	Time to First Reaction
TTHoS	Time to Hands on Steering

1.1.1. Driver factors

Firstly, *driver factors* are a prime subject matter when analyzing Level 3. Included in this factor class are behavioral reactions drivers demonstrate while driving or engaging in NDRT. The *driver* factors entail all driver knowledge and personal characteristics that can vary over long- or short-term. Fatigue or situational awareness are part of this first class. Most of these characteristics are limited to internal abstract psychological constructs, e.g. trust in automation, or anthropological measures that can influence take-over. These factors are difficult to measure and manipulate objectively due to their level of abstraction. Several studies have investigated widely accepted constructs in the passenger car context. Specifically, fatigue does not seem to negatively affect the drivers' take-over capabilities (Weinbeer, Muhr, Bengler, Baur, Radlmayr, & Bill, 2017). Although task-related fatigue can be induced through NDRT and measured through eye closure and head movement changes (Jarosch, Kuhnt, Paradies, & Bengler, 2017; Schmidt, Braunagel, Stolzmann, & Karrer-Gauß, 2016), recent findings also present results in which engaging and activating NDRT can manage and reduce driver drowsiness (Weinbeer, Muhr, & Bengler, 2019). Typically, the train of thought when conducting a Level 3 experiment, is to have very few take-over situations to exclude the possibility of learning. Instead of trying to exclude any possibility of learning, this study will present multiple take-over situations. Therefore, learning effects are expected to reduce the time to first reaction, as defined in Section 2.1 (Hypothesis 1).

1.1.2. Environmental factors

The second class, *environmental factors*, are widely considered by previous published studies regarding take-over behavior during Level 3, e.g. Damböck (2013), Merat, Jamson, Lai, Daly, and Carsten (2014), Gold, Körber, Lechner, and Bengler (2016). These factors are primarily investigated as they are easily manipulated in experimental designs and can be measured objectively. Included in this class are factors such as type of driving environment, criticality of take-over situations, CAD duration and NDRT. As the effects of a NDRT may influence a driver, but are induced through the environment, we chose to define NDRT as an *environmental factor*, altering the original definition of Vogelpohl et al. (2016). Most current Level 3 studies are limited to highway scenarios and Gold et al. (2016) find a negative influence of higher traffic densities on take-over performance in the passenger car context. Also, higher complexities of take-over situations seem to elongate take-over times (Damböck, 2013). Prolonged take-over times for high situation complexity can be explained cognitively through the multitude of objects that need to be perceived and processed, leading to more complex planning processes and reactions as proposed by Endsley (1995). While cognitive processing of the environment cannot be monitored objectively, the surroundings that are processed by the driver can be measured and described through camera systems, radars and lidars in a mathematical sense. However, a clear description method of vehicle surroundings in take-over situations has not been generated thus far (Schneider, 2009), making it difficult to assign environmental complexities to differentiating take-over performance.

Feldhütter, Gold, and Bengler (2016) examined CAD durations of five and twenty minutes, findings showed temporal differences in reaction times of first gaze movement. Typically CAD durations in current publications focus on times lower than five minutes, extremely quick CAD durations being only a few seconds (Louw, Merat, & Jamson, 2015). CAD duration seems to have a negative effect on take-over as glances to the driving scene shorten (Feldhütter et al., 2016) as well as leading to fatigue (Neubauer & Matthews, 2012). This effect is explained through the loss of situational awareness, leading to higher effort and time necessary to perceive the environment. Additionally, autonomous driving appears to have an effect on the quality of take-over after motoric, visual and cognitive control has been obtained (Brandenburg & Skottke, 2014). Multiple studies also address the effect of different NDRT as independent variables, as drivers tend to reallocate attention away from driving-related activities and focus on a NDRT variant (Naujoks, Purucker, & Neukum, 2016). Typically, these studies require participants to perform standardized abstract NDRT [i.e. n-back task, Surrogate Reference Task (Radlmayr, Gold, Lorenz, Farid, & Bengler, 2014), Tracking Task (Damböck & Bengler, 2012), Twenty-Question-Task (Petermann-Stock, Hackenberg, Muhr, & Mergl, 2013)]; entertainment NDRT [i.e. watching video, playing games, listening to music (Zeeb, Buchner, & Schrauf, 2016)]; vehicle related NDRT,[i.e. interaction with navigation system (Wulf, Zeeb, Rimini-Döring, Arnon, & Gauterin, 2013) and speaking on the phone (Naujoks & Neukum, 2017)]. The abstract tasks of the abovementioned list of NDRT can be scaled and controlled in their complexity, e.g. n-back task by increasing the 'n' (Kirchner, 1958). Findings display contradicting results when analyzing the effect of NDRT on reaction times, sometimes displaying an effect of type of NDRT (Petermann-Stock, 2013) and sometimes not (Gold et al., 2016; Radlmayr et al., 2014). Comparison between the studies is difficult, as other parameters are varied throughout the experimental setup between studies. It seems that the execution of NDRT in general lead to longer take-over times as opposed to take-overs in which drivers continuously monitor the driving task (Vogelpohl et al., 2016). Zeeb, Buchner, and Schrauf (2015) find no effect of visual distraction on motor readiness of hands at the steering wheel. Nomadic handheld devices are not examined by Zeeb et al. (2015), as neither this study will, due to the fact that these devices can reduce motor readiness if the hands are not free and take-over warnings are difficult to synchronize on nomadic devices.

As environmental factors are the prime group of influencing factors being analyzed in current studies, this generates a large network of complexity. For the truck context none of these factors have been investigated and a comparison to the passenger car context should be formed. The present study will consider two different NDRT which are expected to generate differences in take-over times as presented by Petermann-Stock (2013) (Hypothesis 2). Three different CAD durations are also examined while keeping traffic densities constant. It is expected that different CAD durations have an effect on reaction times (Hypothesis 3). To examine the possibility of different take-over situation types, multiple highly temporal situations will be explored to exclude the possibility of learning reactions for one specific situation when repeated. A study from Walch, Lange, Baumann, and Weber (2015) presented results in which reaction times were investigated with regard to warning time criticality as well as take-over situations, however, with only a significant effect between some warning types and situations. Take-over situations in the present study are categorized in three groups, i.e. obstacle avoidance, stabilization and braking & stabilization scenarios. Resulting from these three groups and the variation in driver responses needed to perceive and plan the take-over reaction, significantly different take-over times will be observed for different take-over situations (Hypothesis 4). A hypothesis with regard to which type of situation generates the quickest take-over times is not postulated.

1.1.3. Vehicle factors

The third factor class, *vehicle factors,* include technical ADAS versions, fallback systems, and visual field of view from the cockpit within a vehicle. These factors have not been explicitly regarded as independent variables in the abovementioned publications, but are inherently different between experiments. Therefore, a comparison between the scientific discoveries needs to consider the difference in vehicle parameters. It is unrealistic to compare all of the vehicles utilized in the listed publications regarding take-over behavior in Level 3, however, especially fallback systems possibly affect behavior considerably. To the best of the authors' knowledge no current study compares different reaction types in different vehicles. This class of factors is held constant in this experiment and no hypotheses are generated. By utilizing only one truck, i.e. vehicle factors are held constant, the abovementioned environmental factors can be analyzed and a comparison can be drawn between the group of all passenger car studies which also had constant vehicle factors

1.1.4. Human-machine interaction factors

The fourth class, *human-machine interaction factors*, regards the warning modalities, interaction design and NDRT layout within the vehicle (Naujoks & Neukum, 2017). This set of parameters is not investigated explicitly in the present study and is held constant throughout experimentation. While iteration of these parameters can affect take-over behavior (Naujoks, Mai, & Neukum, 2014), vehicle interior design always needs to be taken into account when arranging a HMI. As HMI concepts are different among different vehicles of one automobile manufacturer and are changed with new models, influencing factors also vary. To reduce any effect of HMI factors, this study will only regard one HMI setup and keep all interaction identical between participants.

1.2. Truck context

Predominantly driven by a trained group of professionals, as opposed to occasional drivers in the car context, truck drivers are responsible for vehicles with a higher mass that potentially cause higher safety critical situations if mismanaged. German legislation passed a law forcing new trucks with a mass greater than eight tons to have an emergency brake assist as of 2015, reducing the current speed by 10 km/h (Bundestag, 2016), similar actions are discussed within the European Union. Automation in the truck context thus far always incorporates a human in-the-loop, who is assisted in certain situations through specific ADAS, e.g. lane departure warning systems or automatic transitions.

When considering reactions in take-over behavior between truck and car contexts, behavior may vary between drivers of the two vehicle types. Accordingly, a multitude of factors of truck driving may influence the take-over behavior differently when compared with drivers of passenger cars. Firstly, *driver factors*, e.g. experience or trust in vehicle, and intrinsic driving motivation are essentially different. In Germany and Britain an average of roughly 39 min are spent in passenger cars (Department for Transport, 2015; Streit, Chlond, Weiß, & Vortisch, 2014), while professional truck drivers are allowed to drive up to ten hours according to German law. This effect is also recorded by the total distance passenger cars and trucks cover annually in Germany; cars: 13,341 km, trucks: 98,809 km (Kraftverkehrsstatistik, 2016). Similar statistical differences are documented in other countries (Foundation, 2017; U.S. Department of Transportation Federal Highway Administration, 2014). In addition, a larger portion of time is spent on highways by truck drivers, 50% of annual mileage in Germany (Bueringer, 2007; Bundesamt, 2013). This allows professional truck drivers to practice and learn driving maneuvers over an extensive period of time with a higher reoccurring frequency. Secondly, the *environmental factors* are largely similar for truck drivers compared to passenger car drivers. A difference being, that the situations truck drivers experience are different as a majority of their travel is limited to the right lane on highways. An additional hypothesis is therefore generated, truck drivers will take back control quicker than passenger car drivers due to more experience and training (Hypothesis 5). Thirdly, when regarding *vehicle* factors, depending on the truck model, drivers have a field of view approximately two meters higher above tarmac than those of passenger cars. Vision is more confined for close ranges in trucks, due to a large blind spot imminently in front of the bumper and to the rear of the cabin. Accelerations and velocities are lower, however, sharp decelerations must be avoided to reduce the risk of damaged cargo or rear-ending of following vehicles. Fourthly, different vehicles also present different *HMI factors,* which may influence drivers differently in the two vehicle types. Without a controlled experimental comparison between different setups and interaction modalities a comparison is difficult and has thus far not been conducted.

When taking all four factor classes into account, Vogelpohl et al. (2016) conclude that complexity of take-over situations, warning modalities and the warning times before take-over affect take-over behavior predominantly. All current published studies conducted on time critical take-over behavior of Level 3 automated driving, i.e. CAD, were performed in the passenger car context, investigating different *environmental* and *HMI* factors as independent variables (Damböck, 2013; Feldhütter et al., 2016; Louw et al., 2015; Radlmayr et al., 2014; Vogelpohl et al., 2016). Take-over behavior in the truck context has thus far been investigated by Zhang, Wilschut, Willemsen, and Martens (2017) in platooning scenarios with no time criticality. Different reaction times were measured by controlling for the interaction with the truck and allowing for an NDRT. Unfortunately these findings are not comparable to our scope, as no time criticality was introduced. Similarly the workload and trust in automation has been investigated in platooning scenarios (Hjälmdahl, Krupenia, & Thorslund, 2017). Interestingly, Hjälmdahl et al. (2017) found that drivers tend to overestimate their situational awareness in platooning scenarios. Further research focusing on the design of an HMI in the truck context has been conducted by Richardson, Flohr and Michel (2018) and for military vehicles by Baltzer, Rudolph, López, and Flemisch (2016). Socio-economic factors and the scope of NDRTs on drivers and hauliers was investigated by Janson and Lindqvist (2017). The lateral partial automation of a truck was tested by Schermers, Malone, and Van Arem (2004), identifying that drivers assisted were less likely to change lanes compared to manual drivers in critical take-over situations which required braking. While the truck context has been regarded in several aspects of automated driving, research focusing on time critical take-over scenarios in Level 3 with NDRT has not been addressed.

1.3. Hypotheses

The scope of this study is to record behavioral data of truck drivers currently operating a semi-truck with an activated Level 3 automation function, SAE Level 3. This incorporates the data collection of reaction times and eye-tracking data. The measurements of the body posture are not included in this publication but are considered in Lotz and Weissenberger (2019). A moving-based simulator will be used to manipulate environmental factors, while HMI and vehicle factors are kept constant (see Fig. 1). The variation of said factors will be measured through different camera-based systems and referenced on reaction times. This type of experimental setup is thus far novel within the truck context and will present data on take-over behavior for highly trained drivers that can be compared to established Level 3 research. The comparison can thus far only be generated superficially to other publications in the field. Two different NDRT are chosen which are hypothesized to distract drivers over longer periods of time, making it easier to compare trials in the beginning and end of the experiment as attention should be held almost constant. When presenting many take-overs, drivers will likely display attentive behavior towards the driving task after multiple take-over repetitions and will try and uphold higher situation awareness, with less attention being allocated towards NDRT. Typically, studies with Level 3 only present between one to four take-overs which bears the danger of learning behavior influencing reaction times in early trials. In order to observe learning behavior, short

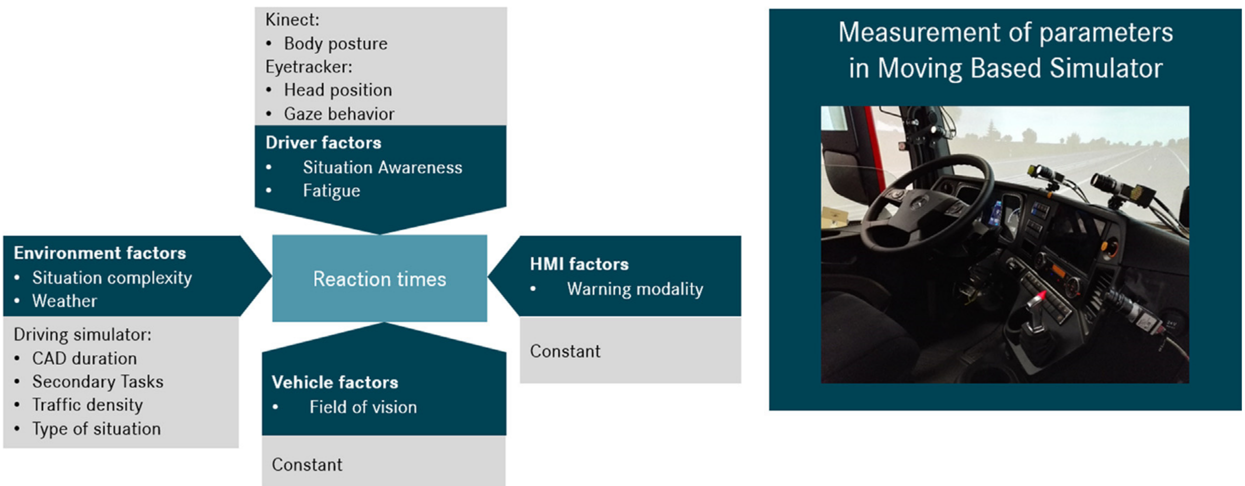

Fig. 1. Influencing factors for the current experimental setup including measurement techniques.

CAD durations were chosen for this study which is coherent with typical CAD durations of less than five minutes (Nilsson, Laine, & Jacobson, 2017).

To exclude the possibility of participants recognizing identical take-over situations and learning situation specific behavior, nine different take-over *situation types* were defined for each of the nine take-overs in the experimental study.

Based on the description of previous research within the Level 3 domain, five primary hypotheses are generated. The first four hypotheses are expected to have a direct effect on reaction times for time critical take-over situations. If an influence is found, this could lead to the possibility of adaptive warning times.

(1) Learning behavior will be observed due to first time use of the automation function. Reaction times will decrease as interaction for first time users of the automation function will adapt behavior.
(2) NDRT *game* will prolong take-over times compared to *video*, as definite visual attention is required to successfully solve tasks. The *video* presents information in two modalities (auditory and visual), while the *game* can only be solved if visual and motoric modalities is allocated. Both tasks require cognitive resources additionally.
(3) Longer *CAD durations* will prolong take-over times for TTEoR as well as TTHoS and TTFR, as argued by Feldhütter et al. (2016).
(4) Due to the complexity of the surrounding and the take-over situations different take-over times are expected.
(5) Average reaction times will be lower than those reported of passenger car drivers. This hypothesis can only be compared quantitatively to published studies as no pure passenger car drivers were examined. Additionally, the *driver, environmental, HMI* and *vehicle* factors of published studies may differ to this study.

2. Materials & methods

2.1. Reaction time definition

The following reaction times are defined for a comparison of independent variables (Damböck, 2013).

TTEoR: The Time To Eyes on Road is a measure for cognitive processing, as a visual and auditory request to intervene (RtI) has to be processed by subjects and eyes have to be moved towards the road. Times can vary significantly due to the visual path subjects choose to address. Theoretically, if information regarding take-over is searched for in the instrument cluster before the environment is focused on, TTEoR rises drastically due to the additional saccade and fixation towards the instrument cluster.

TTHoS: Time To Hands on Steering expresses the motoric reaction time necessary for subjects to regain lateral control of the vehicle. It is not possible to distinguish whether subjects begin with motoric reaction after eyes fixate the road or if motoric reaction is initialized simultaneously with eye movement.

TTFR: Time To First Reaction measures the time between RtI and a steering wheel input greater 1 Nm or 5% accelerator/decelerator pedal position change. At RtI the trajectory and speed of the automation function was kept constant until one of the abovementioned thresholds were activated. This caused the automation function to deactivate. It could be possible, that the TTFR is lower than TTHoS, as reactions at the pedals are also an input for TTFR.

2.2. Participants

The sample size for this experiment consisted of 95 participants, four female and 91 male. This gender distribution is similar to the German truck driver population, with 1.7% professional female drivers employed in Germany (Statista, 2017). Age distribution was controlled for in the sampling and reflects the demographic distribution of truck drivers in Germany. A total of 88 participants completed the course and were considered for statistical analysis, no one was involved in a crash. Exclusion of 7 participants occurred due to technical synchronization failures of sensors. The primary vehicle driven in everyday use by 71 of the 88 participants was a Mercedes truck. Mean age of the participants ($n = 88$) was $M = 42.6$ years ($SD = 9.61$ years) with a range from 23 to 62 years. All participants had an annual minimal mileage of 20,000 km/year and a mean of $M = 73{,}823.8$ km/year ($SD = 45{,}378.5$ km). A total of 38 participants had corrected vision, wearing contact lenses or glasses, participants with progressive lenses did not take part in the experiment. Overall, 55% of participants were familiar and used speedometers daily (Adaptive Cruise Control = 34%). As personal preferences of ADAS usage varies, even for professional drivers, this factor could not be considered during sampling. To improve controllability for the Level 3 automation function, interaction was identical to that of Mercedes-Benz speedometers. An influence of Mercedes familiarity was not found in the data, as the automation function was completely new for all participants. Physical wellbeing was addressed and queried prior to the start of the experiment and monitored during procedure.

2.3. Experimental design

All participants were presented with nine different take-overs (trials). Each trial consisted of one of the nine *situation types*, which required either an obstacle avoidance (i.e. Fog & Warning Triangle, Stranded Vehicle, Lost Cargo, Lane Narrowing and Construction), stabilization of the vehicle (i.e. Loss of Lane) or braking (i.e. Cut in Left, Cut in Right and Lead Braking), the situations types can be found in Fig. 2. Ten different variants of the sequence of *situation types* were designed, of which each

Fig. 2. Nine different take-over situations categorized into three basic maneuvers that were required to circumvent a crash. Times to collision (TTC) are displayed above each situation.

participant drove one, see Fig. 3. In each variant the nine take-over situations were divided into three blocks. Within each block the trials had a *CAD duration* of each 30, 120 or 240 s, levels were randomized within each block. Two different tasks, *video* or *game* were presented during the experiment. During one of the three blocks (3 take-over situations) the *video* was shown, during the other two blocks (6 take-over situations) a geography *game* was played, which is described in Section 2.4.3. Procedure requires all participants that are involved in a collision to terminate the experiment. In order to reduce the amount of take-over collisions, we chose to keep the left lane immediately before and during take-over idle to reduce crash risk at RtI.

This results in an experimental design with the unbalanced within-subject factors *CAD duration* [30, 120, 240 s] and *NDRT* [video, game] as well as the balance within-in subject factors *trial* [1–9] and *situation types* [9 different situations]. The comparatively high amount of independent variables and factors were chosen to address the study aims of identifying influencing factors on the take-over behavior. This resulted in a non-orthogonal experimental design. A full iteration of all variants would have required 9! = 362,880 variants and comparability is limited to this extent.

Each of the nine trials started when the automation function of the vehicle was activated through the driver and the Level 3 commenced. Due to individual temporal activation of the automation function, CAD duration was not fixed, but was distributed around 30 s (M = 37.0 s; SD = 7.0 s), 120 s (M = 121.4 s; SD = 7.5 s) and 240 s (M = 243.6 s; SD = 8.4 s). Simultaneously, the NDRT was activated on the mounted tablet in the central console. Due to the occurrence of one of nine take-over situations, a take-over signal was displayed for three seconds after which the tablet was blackened and locked. This take-over signal or RtI consisted of visual and auditory warnings in the instrument cluster and tablet. The visual icon displayed the outline of a set of hands holding a steering wheel colored red with the text “Hands On” written beneath. Simultaneously, a sound was triggered at 1200 Hz and 200 bpm until a TTFR was measured. During manual driving, after the RtI, the display

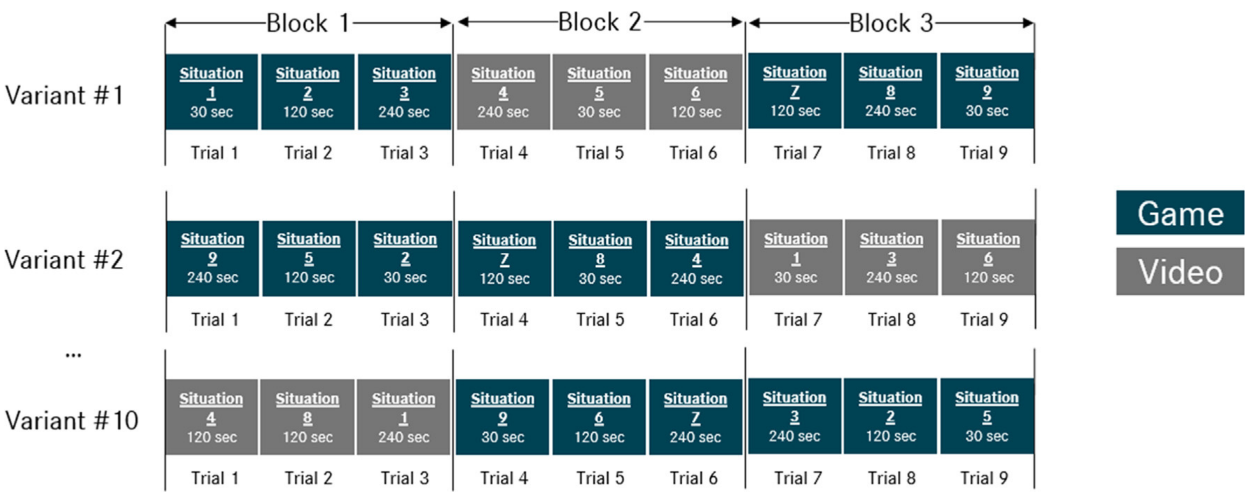

Fig. 3. Experimental procedure of the experiment with three exemplary variants. Nine trials were driven by each participant that had different factors of the independent variables.

was locked and darkened. The vehicle held the current steering angle and speed at the RtI until driver took back control. After take-over a 50 s manual drive commenced until drivers were asked to activate the automation function again, initiating the next trial. Within this manual drive, after 20 s, a verbal question was presented to the participants through loudspeakers in the cabin with regard to the perceived temporal criticality of the prior take-over. The experiment was designed to verbally inquire criticality in temporal proximity to the RtI. In order to shape the query as non-invasive and simple as possible, subjects were asked in regard to the temporal effect of the RtI. The additional aspect of non-fluent native speakers was also taken into account and the question was phrased with regard to take-over time. The question requested a categorization of temporal criticality: *"Was the time for take-over sufficient, short or much too short?"* Graphical representation of the experimental design can be found in Fig. 3.

Dependent variables consisted of take-over times on the basis of Damböck (2013), *time to eyes on road* (TTEoR), *time to hands on steering* (TTHoS) and *time to first reaction* (TTFR). The TTFR is the main measure of reaction as it represents the first oversteering of the automation function by the driver.

2.4. Apparatus

2.4.1. Driving simulator and track

Venue of this experiment was the Mercedes-Benz Moving-Base-Simulator, allowing 360° simulation of driving related situations with realistic forces and accelerations of complete vehicles. A Mercedes Benz Actros truck cabin was mounted within the simulator dome, which was connected to a hexapod on a carriage enabling translatory motion as well as limited pitch, yaw and roll. All human-machine-interaction and actuator inputs from within the cabin are translated with a realistic engine to movements of the composite dome, carriage and visual display of the 360° projection screens. The interior of the cabin was equipped with a capacitive steering wheel designated to record the TTHoS.

The activation of the Level 3 system was managed through the SET-button on the right-hand side of the steering wheel. Through activation with the SET-button on the steering wheel the automation function and therefore Level 3 was initiated, this was displayed at all times in the instrument cluster with a blue icon. During Level 3 the driver could overrule lateral or longitudinal control at any time without deactivation of the automation function. As soon as overruling was completed and all controls were let go, automation reactivated itself. An exception being if the brake pedal was pressed, in these instances the automation needed resetting.

2.4.2. Eye-tracker

A Smart Eye Pro eye-tracker was used as a means to record gaze behavior and was mounted on the dashboard of the truck cabin. Four lenses were dispersed within the cockpit along the top of the dashboard to allow a large field of view in which the eye-tracker is capable to record data. The eye-tracking cameras were calibrated in an automated procedure for each participant prior to the drive. Data obtained was utilized for the calculation of TTEoR and for the analysis of visual attention during and immediately after take-over.

2.4.3. Non-driving related tasks (NDRT)

Two different NDRT were analyzed, however, experimental design dictated which NDRT was available. The NDRT was presented in blocks of three trials with each trial within these blocks consisting of a different CAD durations. Rather than choosing NDRT that are abstract and could lead to a loss of attention over time, it was decided to present tasks that were common for everyday use, at best for longer periods of time and would allow comparison between early and late trials. Data analysis showed a constant visual attention over all trials towards the NDRT and confirms the hypothesis, that engagement through realistic tasks was constant.

The experimental block in which the video was shown allows comparability to published studies with this NDRT, e.g. (Carsten, Lai, Barnard, Jamson, & Merat, 2012). Theme of our video was a short documentary about truck driving. This NDRT presented auditory and visual information to participants.

The second NDRT was available during two of the three experimental blocks. This task was an interactive geography quiz also presented on the fixed tablet in the central console. Participants could see a contour map of Germany on the tablet and were asked to locate certain cities. By touch the participants could set a point on the map and received feedback as to the actual location of the city. Feedback consisted of a distance measure between the participants' guess and the actual position. An analysis of the game was not conducted. The only measures drawn were the number of taps on the tablet and their coordinates. The task was chosen due to the hypothesis that professional truck drivers have obtained a high geographical knowledge and are motivated to perform well in such a quiz. Notably, if no NDRT is available or the task is mundane, disengagement can lead to a state of passive fatigue (Marberger et al., 2017). Demonstrating typical tasks that could be allowed by legislation and offer comparability to other studies in which video NDRT was examined (Naujoks & Neukum, 2017). Visual attention and motoric manipulation was necessary to successfully execute this task, as no acoustic interaction was required. All text was written in German.

2.4.4. Take-over situations

Nine independent non-reoccurring take-over situations in the environment caused the event of nine RtI. Prior to RtI the automation function followed a leading convoy of five trucks with a distance of 55 m to the last truck in the convoy.

Participants were instructed to set the speed of the automation function to 80 km/h during Level 3. Due to the requirement of manual driving after RtI and natural speed variation through drivers, the convoy could drive away, however, speed of the convoy decreased in these instances to assure that participants would catch up again.

Time to collision (TTC) represents the time drivers are given between the RtI and a collision with the object causing take-over. Take-over situations were presented within a near proximity of the ego-vehicle to guarantee quick take-over behavior. If situations were to be presented with TTC of larger magnitude, participants' reactions would not need to be as quick. Full observation of the environment and updating of situation awareness could occur before an input at the steering wheel or accelerator/brake pedals would be necessary. This is also a prime aspect that makes comparability of publications with different TTC difficult. The take-over situations all required a rapid reaction from the driver in order to overt an accident by either holding the lane, changing the lane or holding the lane and decelerating, see Fig. 2 and Appendix B. The TTC varied due to the fluctuation of surrounding traffic needed to initiate take-over, see Fig. 2, especially in take-over situations in which moving objects caused take-over (i.e. Cut in Right, Cut in Left). For take-over situation *Fog & Warning Triangle* TTC was less than four seconds, however, the warning triangle was not situated in the path of the truck's trajectory. Furthermore, for take-over *Lane Narrowing* TTC is calculated to the beginning of the cones, actual lane narrowing commenced after an additional two seconds. Further conditions in Appendix B.

2.5. Procedure

Participants were welcomed by an experimenter and were asked to fill out consent forms, bank details to receive their financial compensation and a demographic questionnaire. A written description of the experiment was presented in which the automation function was explained. The occurrence and explicit investigation of take-over behavior was not mentioned, however, take-over functionality for RtI was explained during the detailed explanation of the automation function. The experimenter additionally explained the activation and display of the Level 3 ADAS after participants had read instructions. This was to ensure that all participants had understood basic functionality, as some participants were not fluent in German. It was never mentioned to the participants, that take-overs were the focus of the study, rather user feedback on the new ADAS automation function was being evaluated.

Following the introduction, participants were led to the Actros cabin through a gangway without visually seeing the dome. It was explained, that functionality of the truck was identical to any truck on the road and that all German road traffic rules apply. The automation buttons on the steering wheel were revised and the calibration of the eye-tracker was conducted. An introductory drive including a verbal tutorial of the automation function (activation, overriding of functions, take-over and deactivation) was driven for 10 min after which the experimental drive started by activation of the tablet. The acclimatization phase was conducted to ensure, that participants could control the Level 3 ADAS and were not instructed to react in any way. Participants were asked to follow a leading convoy (five trucks) at 80 km/h during manual and automated drive without overtaking. A total of 38 km were driven within approximately 28.5 min after the acclimatization phase. Additionally, the tutorial was conducted within the first section of highway of approximately 12 km length. Each trial consisted of a phase in which the Level 3 automation function was activated (CAD duration 30, 120 or 240 s) with availability of an NDRT. Following RtI a 50 s long manual drive phased was performed during which a verbal question was presented to participants asking for a subjective rating of the amount of time given for take-over.

3. Results

3.1. Reaction times - Hypotheses 1, 2, 3, 4

A total of 768 take-overs were successfully conducted by the 88 participants. Due to early identification of take-over situations by the participants, 24 take-overs were not regarded. TTFR displayed a positively skewed distribution with a mean of 1.35 s (SD = 0.49 s). Visual representation of mean TTFR for the independent variables *CAD duration* and *NDRT* are displayed in Fig. 4, with the slowest reaction of TTHoS and TTFR at 6.3 s. On average a take-over was conducted sequentially by focusing the eyes on the road followed by taking control of the vehicle and finally showing a first reaction. Due to movement of the driver and possible loss of the cornea, quality of the gaze direction data is not constant. To classify eye-tracking as conclusive a qualitative comparison of TTEoR, TTHoS and TTFR was conducted. If the sequence of reaction times did not match a video of the participants during take-over, eye-tracking was ruled inconclusive for the period of the take-over. Reaction times for all three dependent measures were: TTEoR = 0.34 s (SD = 0.43) (291 trials of TTEoR were disregarded due to inconclusive eye-tracking data during take-over, leading to 477 trials for TTEoR); TTHoS = 1.20 s (SD = 0.48); TTFR = 1.35 s (SD = 0.49). Due to frequent control glances at the road, in 217 of 477 regarded TTEoR trials, drivers were in the midst of a control glance leading to short TTEoR. This aspect will be discussed in more detail in Section 3.2. The formulated hypotheses are investigated through statistical procedure leading to an adjusted alpha-level of $a_{adj} = 0.002$, see Table 2. Three full-factorial within-subject ANOVA were calculated with the four abovementioned independent variables, i.e. *CAD duration*, *NDRT*, *trial* and *situation type* on TTFR (3 [30, 120, 240 s] $\times$ 2 [video, game] $\times$ 9 [1–9] $\times$ 9 [9 different situations, see Fig. 2]). Due to the experimental design, some combinations are non-existent and factor-levels are not considered. A full-factorial ANOVA

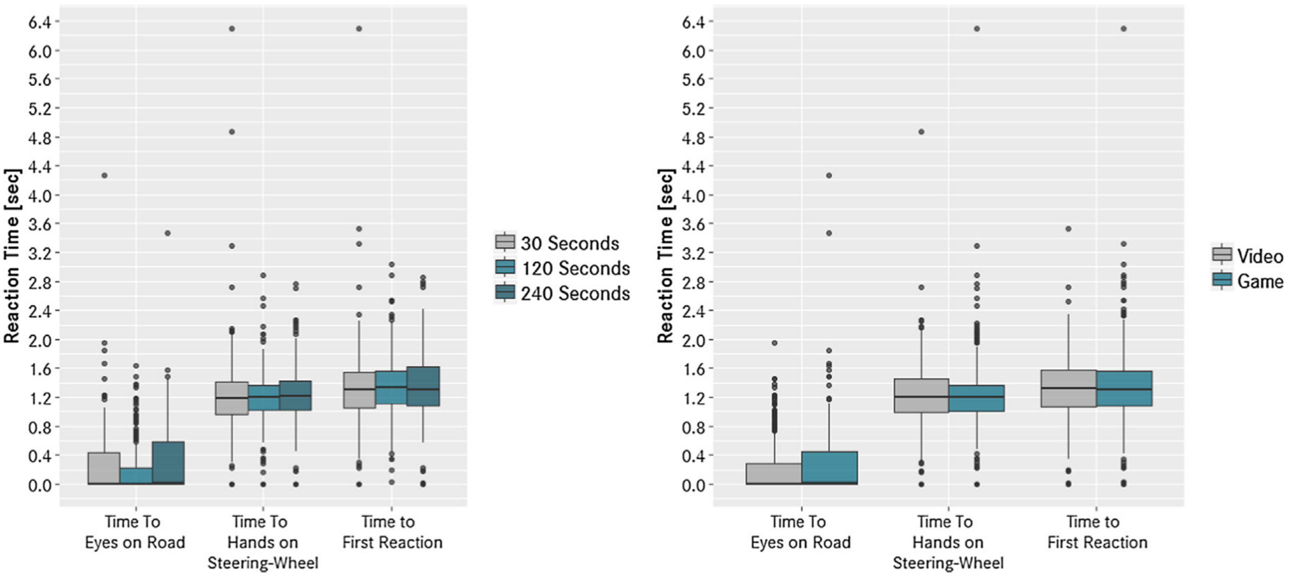

Fig. 4. Reaction times Time To Eyes on Road (TTEoR), Time To Hands on Steering (TTHoS) and Time To First Reaction (TTFR) for independent variables CAD duration (left) and secondary task (right).

Table 2
ANOVA of dependent variable TTFR and TTHoS.

Dependent Variable	Measure	Sum of Squares	df	Mean Square	F	Pr (>F)	Partial η^2
Time To First Reaction	NDRT	0.09	1	0.091	0.52	0.471	0.000732
	Trial	**31.23**	**8**	**3.904**	**22.381**	**<2e-16**	**0.201000**
	Situation Type	**16.62**	**8**	**2.077**	**11.910**	**4.44e-16**	**0.118000**
	CAD Duration	0.73	2	0.367	2.103	0.1229	0.005890
	NDRT * Trial	1.57	8	0.196	1.124	0.3449	0.012500
	NDRT * Situation Type	2.18	8	0.273	1.562	0.1325	0.017300
	Trial * Situation Type	3.41	11	0.310	1.778	0.0540	0.026800
	NDRT * CAD Duration	0.84	2	0.418	2.396	0.0918	0.006700
	Trial * CAD Duration	0.96	6	0.161	0.921	0.4788	0.007730
	Situation Type * CAD Duration	0.75	1	0.752	4.313	0.0382	0.006040
	NDRT * Trial * Situation Type	0.25	2	0.126	0.723	0.4859	0.002030
Time To Hands on Steering	NDRT	0.07	1	0.0678	0.368	0.5444	0.000518
	Trial	**23.12**	**8**	**2.8894**	**15.67**	**<2e-16**	**0.150000**
	Situation Type	**8.26**	**8**	**1.0325**	**5.600**	**6.7e-07**	**0.059400**
	CAD Duration	0.76	2	0.3794	2.058	0.1285	0.005760
	NDRT * Trial	3.03	8	0.3788	2.054	0.0381	0.022600
	NDRT * Situation Type	4.12	8	0.5149	2.793	0.0048	0.030500
	Trial * Situation Type	2.93	11	0.2664	1.445	0.1481	0.021900
	NDRT * CAD Duration	0.77	2	0.3871	2.099	0.1233	0.005880
	Trial * CAD Duration	1.67	6	0.2784	1.510	0.1721	0.012600
	Situation Type * CAD Duration	1.47	1	1.4745	7.997	0.0048	0.011100
	NDRT * Trial * Situation Type	0.30	2	0.1486	0.806	0.4472	0.00227
Time To Eyes on Road	NDRT	1.06	1	1.0591	6.611	0.0105	0.015500
	Trial	**6.93**	**8**	**0.8668**	**5.411**	**1.69e-06**	**0.093600**
	Situation Type	**4.92**	**8**	**0.6153**	**3.841**	**0.0002**	**0.068300**
	CAD Duration	0.22	2	0.1120	0.699	0.4977	0.003320
	NDRT * Trial	1.48	8	0.1851	1.155	0.3251	0.021600
	NDRT * Situation Type	2.29	8	0.2858	1.784	0.0783	0.033000
	Trial * Situation Type	2.38	11	0.2165	1.351	0.1937	0.034300
	NDRT * CAD Duration	0.09	2	0.0427	0.267	0.7660	0.001270
	Trial * CAD Duration	0.77	6	0.1280	0.799	0.5713	0.011300
	Situation Type * CAD Duration	0.10	1	0.0992	0.619	0.4318	0.001480
	NDRT * Trial * Situation Type	0.06	2	0.0311	0.194	0.8237	0.000926

Bold signifies that the adjusted alpha-levels were below the threshold.

was chosen for statistical analysis, as the calculation is mathematically identical to multiple one-way ANOVA and additionally offers the possibility of analyzing interactions.

Two main effects of *trial* $F(8,710) = 22.381$, $p \ll 0.001$, see Fig. 5, and of *situation type* $F(8,710) = 11.91$, $p \ll 0.001$, see Fig. 6, were identified, while no interaction was determined for TTFR. Therefore, our results are in accordance with

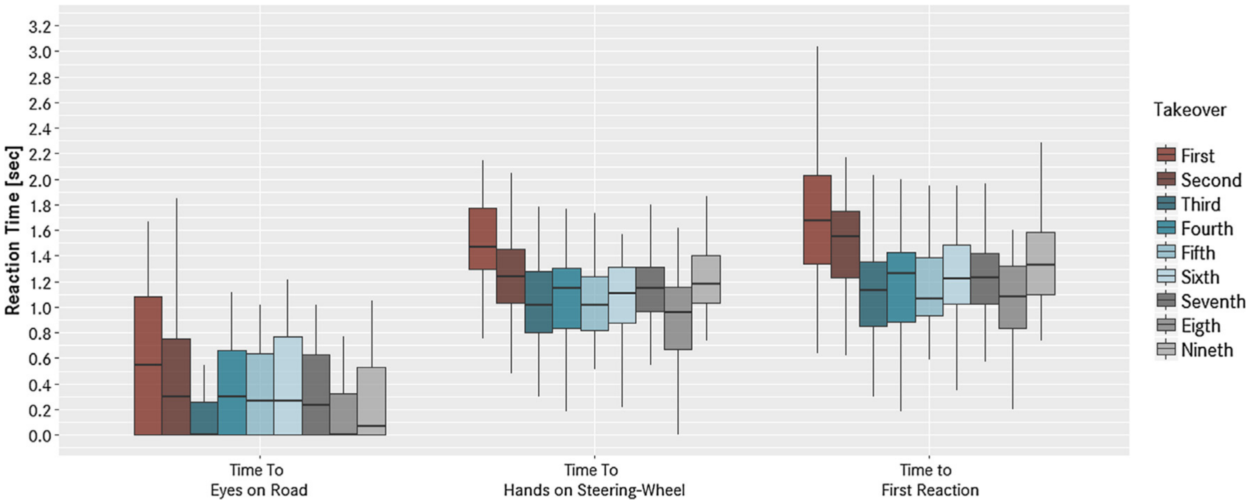

Fig. 5. Chronological visualization of reaction times Time To Eyes on Road (TTEoR), Time To Hands on Steering (TTHoS) and Time To First Reaction (TTFR). For each of the abovementioned reaction times all trials are displayed from left to right (first to last).

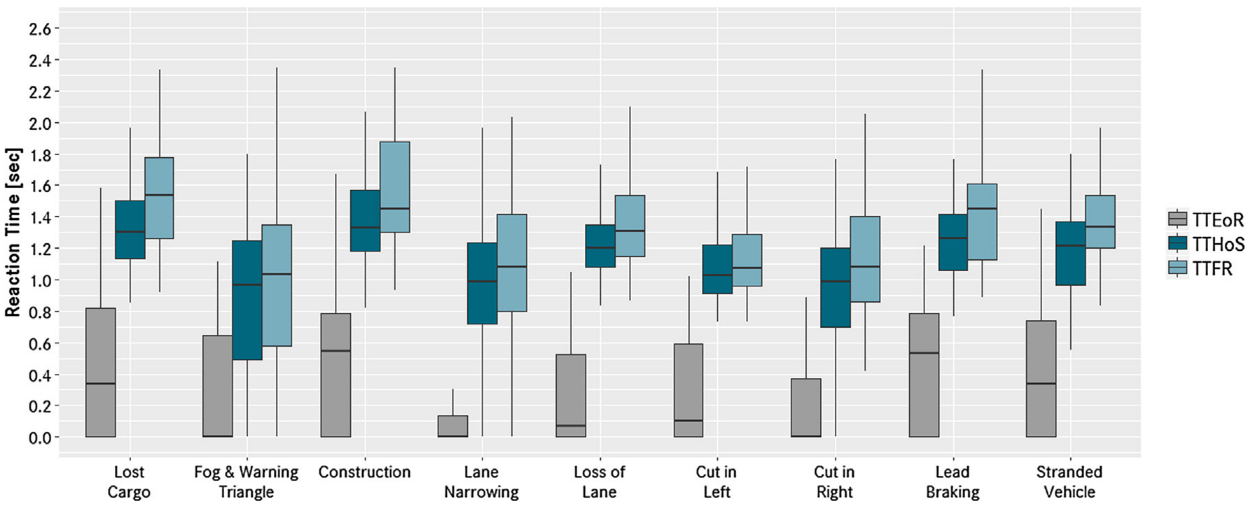

Fig. 6. Reaction times Time To Eyes on Road (TTEoR), Time To Hands on Steering (TTHoS) and Time To First Reaction (TTFR) for all take-over situations as displayed in Fig. 2.

Hypothesis 1 and Hypothesis 4. The classification of take-over situations according to Fig. 2 was not aligned with results of TTFR. All take-over situations exhibit learning behavior, even if different average reaction times were observed as seen in Fig. 5. The results of the ANOVA also leads to a rejection of Hypothesis 2, $F(1,710) = 0.520$, $p = 0.471$ and a rejection of Hypothesis 3, $F(2,710) = 2.103$, $p = 0.123$.

The second full-factorial within-subject ANOVA is calculated on TTHoS. Again two main effects are identified for *trial* $F(8,710) = 15.67$, $p \ll 0.001$ and of *situation type* $F(8,710) = 5.60$, $p \ll 0.001$. The results also present a high correlation between the two dependent measures of TTFR and TTHoS ($r = 0.817$). While statistical significant differences were observed for the independent variable *situation types* which shows accordance of the results with Hypothesis 4, Section 3.2. will display reasoning why this conclusion is false.

A full-factorial within-subject ANOVA on TTEoR, see Table 2, for the reduced dataset was calculated and two main effects for *trial* $F(8,419) = 5.41$, $p \ll 0.001$ and of *situation type* $F(8,419) = 3.841$, $p = 0.0002$ were identified. In combination with the results of the full-factorial ANOVAs for TTHoS and TTFR, this leads to a rejection of Hypothesis 3.

3.2. Take-over situation comparison – Hypothesis 4

To have a more accurate look at different behavioral reaction times, the dataset of 477 trials with conclusive eye-tracking was divided into two parts, see Table 3. The first subset in Table 3 contains only those 217 trials in which the eyes were fixated at the road at RtI. The primary cause for these fixations were control glances. Naturally, the TTEoR equaled 0 s while

Table 3
Comparison of reaction times for all situations.

	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Subset 6
Situation types	All	All	Fog & Warning Triangle Lane Narrowing Cut in Right	Fog & Warning Triangle Lane Narrowing Cut in Right	Lost Cargo Construction Loss of Lane Cut in Left Lead Braking Stranded Vehicle	Lost Cargo Construction Loss of Lane Cut in Left Lead Braking Stranded Vehicle
Number of take-overs	217	260	114	65	103	195
Sum take-overs	∑477		∑477			
Eyes on Road at RtI?	Yes	No	Yes	No	Yes	No
Time To Eyes on Road [sec]	0	0.62 (SD = 0.40)	0	0.59 (SD = 0.36)	0	0.63 (SD = 0.42)
Time To Hands on Steering [sec]	0.94 (SD = 0.51)	1.29 (SD = 0.34)	0.71 (SD = 0.39)	1.29 (SD = 0.36)	1.19 (SD = 0.50)	1.30 (SD = 0.33)
Time To First Reaction [sec]	1.10 (SD = 0.52)	1.44 (SD = 0.37)	0.85 (SD = 0.40)	1.43 (SD = 0.32)	1.38 (SD = 0.50)	1.45 (SD = 0.39)

TTHoS = 0.94 s (SD = 0.51) and TTFR = 1.10 s (SD = 0.52). The second subset contains all other trials in which the eyes were not fixated on the road at RtI. The values of Subset 2 were calculated as TTEoR = 0.62 s (SD = 0.40), TTHoS = 1.29 s (SD = 0.34) and TTFR = 1.44 s (SD = 0.37). If the temporal difference between TTFR and TTHoS are compared within the two subsets, both values equal 160 ms and 150 ms. Therefore, almost no difference is made between the motoric response once hands are place on the steering wheel in dependence to visual information at RtI. Similarly, when comparing TTHoS and TTFR between the subsets, both values are almost identical with 350 ms and 340 ms, confirming the findings of Zeeb et al. (2015) that motoric readiness is isolated from visual information. This means, that visual information of the take-over cause at RtI generates a benefit in motoric response prior to hands touching the steering wheel, without altering the temporal action sequence once hands or feet are placed on controls for the first reaction.

Minimal reaction times (lower whiskers) are an indication for premature visibility of take-over situations, if participants chose to observe the driving situation in these instances through control glances. As seen in Fig. 6 three take-over situations (*Fog & Warning Triangle*, *Lane Narrowing* and *Cut in Right*) are mainly responsible for TTEoR = 0 s. Subset 1 and 2 are therefore subdivided further to differentiate between situation types that display the possibility of premature visibility. Premature visibility is defined in this context as the possibility of observing objects in the environment that can be identified as causing a take-over prior to any warning. When comparing the second subset (all trials where no eyes were on the road at RtI) with the fourth and sixth subset (partition of Subset 2 with regard to different *situation types*), similarity of the mean times of the data is apparent. As the eyes were not on the road at RtI for these three subsets, behavioral procedure at take-overs seems to be identical. However, the third dataset representing take-over situations with the three distinct situations (*Fog & Warning Triangle*, *Lane Narrowing* and *Cut in Right*) and TTEoR = 0 s, displays much quicker TTHoS and TTFR. This benefit can be attributed to the timely perception of the take-over situation. Trials of Subset 3 made it possible for the driver to perceive and plan a reaction prior to RtI, generating the identical benefit as long TTC would, with the difference that immediate action was required in the situations presented in this study. The fifth subset entails the take-over situations (*Lost Cargo, Construction, Loss of Lane, Cut in Left, Lead Braking and Stranded Vehicle*) at which TTEoR = 0 s, TTHoS = 1.19 s (SD = 0.50) and TTFR = 1.38 s (SD = 0.50) were calculated. A comparison of Subset 3 and Subset 5, under the assumption that situations were perceived equally beginning at TTEoR = 0 s, displays the benefit premature perception generates. The benefit of 530 ms for TTFR was observed in the data. The results of the ANOVAs calculated in Section 3.1. regarding the significant results for *situation type* are not confirmed by the results in Table 3. As the comparison by means of a one-way ANOVA of Subset 4 with Subset 5 and 6 displays, the different situations do not cause differences in TTHoS $F(1,361) = 0.347$, $p = 0.556$, partial $\eta^2 = 0.00006$ or TTFR $F(1,361) = 0.021$, $p = 0.885$, partial $\eta^2 = 0.001$ as found by the ANOVAs in Section 3.1. Significant results obtained in Section 3.1. are due to the fact that Subset 3 is included in the data. The conclusions drawn with regard to Hypothesis 4 are contradictory. Taking all aspects into account, Hypothesis 4 is rejected.

3.3. Manipulation check

A manipulation check was not conducted during the experimental procedure to validate engagement in the NDRT. However, for the *game* the amount of tablet taps were monitored during the experiment. Of the 517 trials in which the game was presented, 459 participants had touched the tablet at least once within 10 s prior to RtI. On average participants tapped the tablet M = 2.85 times within 10 s, M = 1.40 times within 5 s and M = 0.54 times prior to RtI. A one-way ANOVA was conducted on all trials in which the game was presented. The boolean independent variable whether a tap occurred within 10 s prior to RtI was investigated for the dependent variable TTFR, no significant effect was found $F(1,515) = 0.29$, $p = 0.59$, partial $\eta^2 = 0.00056$. Investigation of engagement with the video NDRT is more difficult as no motoric response was necessary. Control glances directed at the driving task were measured during the complete experiment with respect to the frequency and duration of gazes within the windscreen. The average frequency of control glances occurred every 2.46 s (SD = 3.44), having a duration of 0.43 s (SD = 0.20) and last occurrence prior RtI being 4.63 s (SD = 11.99). Control glances at the road can be compared between the two NDRT with three one-way ANOVAs. A total of 23 observations were disregarded due to no control glances being performed and set to the CAD duration in these instances. No significant differences were found for the time since the occurrence of the last control glance $F(2,451) = 1.83$, $p = 0.161$, partial $\eta^2 = 0.008$ average control glance duration $F(2,451) = 0.252$, $p = 0.778$, partial $\eta^2 = 0.0011$ or average control glance frequency $F(2,451) = 3.886$, $p = 0.021$, partial $\eta^2 = 0.0169$.

3.4. Quality of take-over

A central question regarding take-over from a highly automated vehicle is whether drivers had enough time to conduct the take-over safely. This is based on the introduced discussion how warning times have to be designed, fixed or variable, to allow for adequate take-over times see Section 1. Taking back control from an automated vehicle always brings with it an aspect of safety criticality, due to the temporal facet of the vehicle being in motion. As warning times during Level 3 should never provoke drivers to safety critical behavior, different measures can be consulted to investigate the quality of take-over, such as accelerations (Petermann-Stock, 2013). Therefore, a full-factorial ANOVA was conducted for the maximum acceleration in longitudinal and lateral direction in the 10 s after RtI, to identify if safety critical behavior was promoted through certain factors of the independent variables. No significant main or interaction effects were found, see Appendix C. With regard to Hypothesis 4 no significant effects of lateral and longitudinal accelerations were found for the independent variable *situation type*. Therefore, even if different trajectories were required to circumvent obstacles, the situation itself didn't cause any difference in swerving as neither did any other independent variable.

As a further measure of take-over quality the subjective temporal criticality for each take-over situation was also recorded by presenting a verbal question to the participants. Subjective temporal criticality decreased over the course of the experiment (from first to ninth trial). Analogously to the reaction times, subjective temporal criticality was calculated with a full factorial ANOVA displaying a main effect for *trial* $F(8,710) = 7.48$, $p \ll 0.001$, see Table 4. This shows that participants learned how to react in critical situations and possibly became more familiar with the human–machine-interaction (HMI) over the course of the experiment.

As a final measure of take-over quality we analyzed the *Percentage Gaze* into the left rear-view mirror immediately after RtI. de Winter, Happee, Martens, and Stanton (2014) argue, that "for the shorter take-over request time of 5 s, the drivers were less likely to gaze into the mirrors and over the shoulders [...]". The environment behind the ego-vehicle is generally addressed with less visual attention in take-over situations, as obstacles are typically presented in front of participants. Fig. 7 displays the gaze percentage towards the left mirror within a timeframe of one, two and five seconds after RtI. While keeping in mind that average take-over times (TTFR) was calculated as 1.35 s, we can see only a few instances in which gaze behavior towards the rear was recorded in this timeframe. This behavior is presented even though a lane switch was required for certain situations and participants reacted within 1.35 s on average, oncoming traffic on the left lane was only considered between 2 and 5 s. Additionally, merely take-over situations requiring a lane change and orientation towards the left lane, as seen in Fig. 2, present elongated visual attention towards the rear within 5 s. The overall analysis of quality measures shows that the independent variables caused no significant different steering or braking behavior. However, gaze allocation towards the rear was delayed when compared to the reaction times. Drivers did not assign visual attention towards the rear view mirror on average, even if a prompt analysis of the oncoming traffic was relevant for certain situations. A conclusion drawn from this data could be, that due to temporal criticality, the drivers had not enough time to safely manage the situation even if reaction times were very quick. Self-reported subjective temporal criticality decreased over the course of the drive significantly, substantiating the results on reaction times of Hypothesis 1

Table 4
ANOVA of dependent variable Subjective Criticality.

Dependent Variable	Measure	Sum of Squares	df	Mean Square	F	Pr (>F)	Partial η^2
Subjective Criticality	NDRT	0.30	1	0.304	0.696	0.405	0.00098
	Trial	**26.14**	**8**	**3.268**	**7.486**	**1.29e-09**	**0.07785**
	Situation Type	3.60	8	0.450	1.031	0.4107	0.0115
	CAD Duration	1.32	2	0.661	1.514	0.2208	0.0043
	NDRT * Trial	2.16	8	0.270	0.617	0.7637	0.0069
	NDRT * Situation Type	8.03	8	1.004	2.300	0.0195	0.0252
	Trial * Situation Type	7.39	11	0.672	1.538	0.1131	0.0233
	NDRT * CAD Duration	0.08	2	0.039	0.090	0.9143	0.0002
	Trial * CAD Duration	3.14	6	0.523	1.199	0.3047	0.0100
	Situation Type * CAD Duration	0.01	1	0.006	0.013	0.9096	2e-05
	NDRT * Trial * Situation Type	0.33	2	0.166	0.380	0.6839	0.0011

Bold signifies that the adjusted alpha-levels were below the threshold.

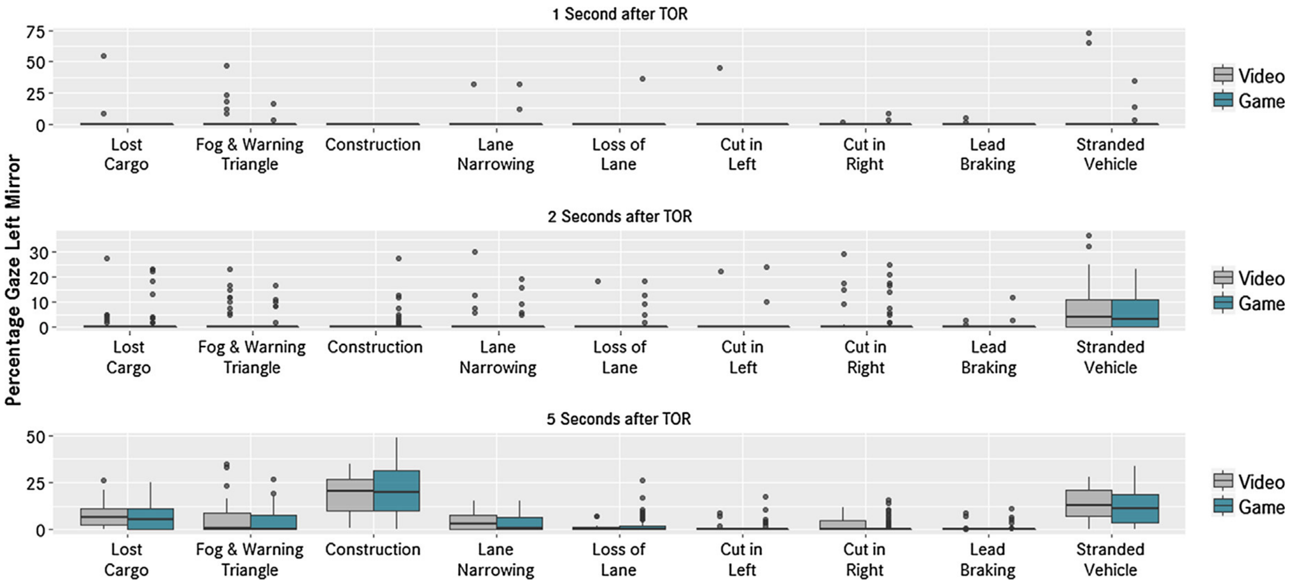

Fig. 7. Percentage of Gaze towards left mirror for each take-over situation and NDRT. (Top) Within a timeframe of 1 s. (Middle) Within a timeframe of 2 s. (Bottom) Within a timeframe of 5 s.

4. Discussion

Focus of this simulator study was to investigate the impact of *NDRT*, *CAD duration*, *situation types* and effect of *trials* (multiple take-overs) on the take-over behavior of professional truck drivers. A generic automation function was implemented in a commercially available semi-truck cabin, allowing participants to activate Level 3 under controlled conditions of a moving-base simulator. Insights were gained for our hypotheses:

(1) Learning behavior will be observed due to first time use of the automation function.
  - Our results are in accordance with this hypothesis, with learning behavior found for the first three trials and a significant main effect. Reaction times converged after the third trial.

(2) NDRT *game* will prolong take-over times compared to *video*.
  - Our results did not support this hypothesis, as no significant differences in reaction times were found.

(3) Longer *CAD durations* will prolong take-over times.
  - Our results did not support this hypothesis, as no significant differences in reaction times were found.

(4) Due to the complexity of the surrounding different take-over times are expected.
  - Our results did not support this hypothesis, due to the results in Section 3.2. for situation type.

(5) Average reaction times will be lower than those reported of passenger car drivers.
  - Without the respect of independent variables, the reaction times of this study are lower than those of reported passenger car publications.

4.1. Learning – Hypothesis 1

Learning behavior is observed in our data with reaction times approximately 0.5 s slower in the first trial compared to the quickest trial (eight), see Fig. 5. Petermann-Stock (2013) also find that reaction time reduces between the first and second trial in their results. Due to the fact that nine take-overs took place in the present study, convergence is detected. It is likely, that this time reduction is caused by the effect of practice, leading to the proceduralization of knowledge (Anderson, 1982). Which reaction is desired when take-over signals are presented and the reduction of scan-paths, can lead to significant time savings that account for the experienced reduction of reaction times. Due to the large number of take-over situations a complete balanced iteration at each chronological trial was not possible. Each take-over situation was, however, presented at least at two different positions in the trial sequence. General learning behavior is observed for each take-over *situation type* within the first three trials. Reaction times seemingly increase for certain trials, specifically the ninth trial, this is however a misconception as the take-over situation 'Loss of Lane' was presented in half of all instances in the ninth trial, resulting in longer take-over times.

4.2. NDRT – Hypothesis 2

The null hypothesis of Hypothesis 2 is rejected. This finding of no NDRT effect on TTFR correlates with those of Radlmayr et al. (2014) and Gold et al. (2016) raising the question if this is a universal finding for the car and truck context or whether certain NDRT on non-nomadic devices show reactional differences as reported by Petermann-Stock (2013). A possibility for the different findings is that the take-over scenario was implemented during a traffic congestion at 35 km/h. A plethora of NDRT could have been investigated in the present study. While telephoning, social media interaction and similar NDRT might provide a realistic test case for professional drivers, we considered these NDRT not engaging enough over multiple trials or assumed difficulties in controllability (social media). Additionally, visual tasks were required to initiate visual inattention towards the driving task.

The visual attention towards the NDRT as well as the frequency, last occurrence and duration of control glances remained constant throughout the experiment, showing that visual behavior was not adapted and was not influenced by the learning effect. Furthermore, the duration of inattention towards the road did not affect the reaction time (TTFR).

4.3. CAD duration – Hypothesis 3

The data displayed in Section 3 *Results* does not confirm a correlation between TTFR and *CAD duration*. The null hypothesis is rejected. The results in Fig. 4 (left) exhibit that duration effects of automation do not affect our results in this specific experimental setting, this finding confirms the finding of Feldhütter et al. (2016) in the truck context. A possible reason that no effect was found can relate to the fact that *CAD duration* did not exceed four minutes. While Feldhütter et al. (2016) did not find a significant effect for reaction times of short and longer drives (20 min), other research suggests that fatigue influences take-over significantly. Vogelpohl and Vollrath (2017) find that Level 3 should not exceed 15 min without NDRT, as fatigue influences drivers' take-over as if drivers were heavily distracted. The excitement of driving in a moving-based simulator and being monitored during driving may have influenced our participant sample with regard to concentration and involvement. Again, internal driver factors as discussed in the introduction, cannot be monitored with sufficient meaningfulness and are a prerequisite for many drowsiness experiments and herein a limitation (c.f. Goncalves, Happee, & Bengler, 2016). When comparing the participant samples on experience, truck drivers may have a higher tolerance for task induced drowsiness compared to the non-professional drivers. Investigating short timeframes is by far not trivial, as it is not possible to guarantee that a Level 3 system will run for multiple minutes before generating a RtI. Unexpected environmental situations, which the technical automation function is not design for can arise at any instance. Another possible reason for missing variation in reaction times due to CAD duration could be that the decay of situation awareness had no effect on take-over, as situations were sudden and surrounding vehicles were not in direct periphery. Gold (2013) also draw this conclusion stating, "[w]ith shorter TOR-time, the subjects come to a decision more quickly, reacting faster, but the quality is generally worse". Under this assumption, the influence of reduced situation awareness could be investigated if correct reactions of participants would be more confined and the possibility of crashing were higher. Within this proposed approach drivers with a higher situation awareness would not be prone to crashing as much as those whose situation awareness has decayed more rapidly.

4.4. Situation types – Hypothesis 4

The null hypothesis of Hypothesis 4 is rejected, see Section 3.2. Different *situation types* generated significantly different TTFR although traffic density was held constant. We have concluded that the combination of situational complexity and premature visibility during control glances seems to affect take-over times massively. As the analysis in Section 3.2. shows, the significant effects on TTFR and TTHoS are due to situations in which drivers observed the environment during RtI leading to quicker take-overs. A quicker reaction time was only achieved for three *situation types*, while no effect was established for the other six situations. The influence of these three situations generated significant results of ANOVAs in Section 3.1. This finding is beneficial, as some environments with complexities similar to those investigated here might allow us to disregard factors such as CAD duration and NDRT to some extent. Validity of this conclusion needs investigation in future work. An environmental model of the truck driver's surrounding would potentially generate added value to the problem, due to the possibility of being able to model visual attention to relevant objects. Such a model expressing the continuous complexity of a situation could allow drivers to engage in more stimulating tasks if low environmental complexities would preside, without the need of systems to monitor cognitive states. This would simplify the problem of guaranteeing adequate take-over times. Schneider (2009) addresses the question that no clear definition of traffic situations exists. First attempts in this regard have been initiated, however, relying on subjective criteria for the description of situations and generic object frequencies (Gold, Naujoks, Radlmayr, Bellem, & Jarosch, 2017; Ohn-Bar & Trivedi, 2016; von Benda, Hoyos, & Schaible-Rapp, 1983).

4.5. Comparing truck to passenger car drivers – Hypothesis 5

A quantitative comparison to literature on reaction times in the passenger car context is conducted for Hypothesis 5. On average the time between RtI and time until eyes fixated within the windscreen (TTEoR) was 0.34 s. This time was, as stated above, dependent on situations. When compared to similar published experiments, the reaction times published by Damböck and Bengler (2012), i.e. approx. 0.8–1.2 s, are far slower than our TTEoR. If only those situations are taken into account, in which participants were not in the midst of a control glance, our TTEoR increases to approximately 0.6 s, see Table 3. This value is far closer to the reported times in the literature. Quickness in response could have been influenced by professional experience of our participants, when comparing values to the previously reported 0.8 s (Damböck, 2013). Our reaction times of TTHoS and TTFR were also at the lower end of the scale when compared to literature values analyzed

(Eriksson & Stanton, 2017). The quickest reported TTFR by Zeeb et al. (2015), Radlmayr, Gold, Lorenz, Farid and Bengler (2014) and Gold, Lorenz and Bengler (2014) range from 1.14 s to 1.67 s and are comparable to our findings. Parallel to other studies, either the experimental setup or participant sample seems to have an influence on reaction times, as the average reaction time in our study undercut the upper scale of reported reaction times of Level 3 scenarios of 10 s (Merat et al., 2014). Gold et al. (2017) state, "when interested in the maximum driver performance within take-over situations, demanding scenarios are needed". Hence, some studies, e.g. Merat et al. (2014), in which uncritical take-over situations are not comparable to our data. Our findings also show, that the visual attention towards the road at RtI allowed drivers to react (TTHoS and TTFR) circa 0.3 s quicker. The procedure of motoric readiness was not affected, confirming previous findings (Zeeb et al., 2015). Two factors limit the applicability of these findings. Firstly, simulator studies are limited in their comparability to real world driving due to underestimation of speed in virtual settings (Bellem et al., 2017). Secondly, drivers knew that they were in a simulation setup and monitored, possibly intrinsically motivating participants to peak performance. However, based on the limited comparison, Hypothesis 5 is failed to be rejected, as truck drivers seem to present quicker reaction times compared to passenger car drivers. Future research will focus on a study in which truck drivers and passenger car drivers are compared directly.

4.6. Generalization of reactions

A statistical evaluation of our reaction times showed that when calculating a 3σ-interval for the TTFR, which depicts 99.7% of the data, we receive a reaction time interval of 0 s to 2.82 s. Reaction times of 0 s (n = 5) were achieved if participants directly engaged in take-over when RtI was presented. A total of six take-overs are not included in the 3σ interval. Five of these six take-overs occurred within the first trial, while one was the *Loss of Lane* situation, requiring no immediate reaction as no obstacle needs to be obverted. It is likely that prolonged reaction times can be attributed to inexperience with the warning cascade and not knowing how to react. The slowest reaction time out of all trials was 6.3 s, a first take-over. When investigating TTFR for Level 3 on a general level, currently it is not possible to define a universal time reserve. This study depicts a small subset of take-over scenarios for truck drivers on the multilane highway. While higher automation functions are in discussion for urban driving areas, research on Level 3 is mainly limited to highways. Without taking the human driver into account, modern radar systems allow for an environmental perception of up to 250 m on a straight road (Continental, 2017). When travelling at 80 km/h this provides information 11.25 s before a possible collision, without taking computation and sensor fusion times into account. Therefore, from a technical standpoint the current maximal time available for a driver to take back control of a vehicle is quite limited. Restricted to the take-over situations and NDRT on non-nomadic devices in this study, this leaves enough time for drivers to regain control of the vehicle. However, inferring the cognitive state of the drivers as well as stress was not investigated.

A quick take-over does not guarantee a safe take-over, as seen in Fig. 7. No effects of any independent variable on lateral or longitudinal accelerations were found. We assume that the driving trajectory and evasion of an obstacle is based on a large part on the strategy and experience of a driver. For future studies, a higher number of surrounding vehicles should be considered around the ego-vehicle. The estimation of distances with long vehicles is difficult through rearview mirrors, especially during take-over situations that require the perception of dynamic objects (Nilsson et al., 2017). Vehicles in direct periphery would limit the amount of correct reactions, possibly leading to a higher collision rate, but also clarifying behavioral differences between successful and unsuccessful take-overs. The data shows that within the first two seconds of take-over drivers are not capable of reacting to the RtI, perceiving the forward environment as well as the rear environment, even if this is required for a safe and adequate take-over, see Fig. 7. Although reaction times suggest that take-overs can be conducted quickly and with a high consistency under three seconds, sufficient visual attention is not observed, leading us to conclude that take-over was not safely managed on average.

4.7. Quality/validity of the experiment

All possible influencing factors cannot be mentioned or observed in such studies. It should be clear, however, that our experiment presented highly critical take-over situations, TTC between 0.92 s and 5 s (M = 4.10 s), which left little time to avoid obstacles. The results presented in this paper describe minimal reaction times to take-over situations for heavy duty trucks in combination with NDRT, similarly to the declared objective of Damböck and Bengler (2012). Comparable published studies of Level 3 within the truck domain do not exist to the best of our knowledge, leaving only room to speculate as to how these participants would have reacted in a passenger car and if significant differences in reaction times would present themselves. Lower speeds and accelerations, different fields of view and driver experiences are factors that can only be considered if a comparison study were to be conducted. When extrapolating the findings to the real world in which the professional driver is highly familiar with all vehicle factors, we hypothesize that similar TTFR would occur. However, no empirical data is currently available of such critical take-overs in real world scenarios.

With regard to our study aim of identifying factors that enable a variable warning time for Level 3, the results show high variance. Primary differences in TTFR occurred due to visibility of objects in the environment prior to RtI. Based on the results collected the intra- and inter-individual differences in reaction times could not be pinpointed to controlled independent variables. It is likely that the variance of reaction times would increase outside of the controlled experimental settings, as drivers would gain more trust in the system and be immersed within their known environment. However, based on the data collected, it is not possible to allow for variable warning times.

5. Conclusions

A large empirical simulator study with professional truck drivers investigating Level 3 take-over behavior has not been published thus far, as most studies focus on passenger car behavior or manual driving aspects of truck driving.

Results obtained lead us to conclude that multiple findings of influencing environmental factors in the passenger car context are transferable to the truck context. This includes that different NDRT and CAD duration did not alter reaction times significantly. It could be possible that the examined CAD durations were not long enough to induce significant differences in behavior and that long-term effects of driving related inactivity are experienced after extensive periods of time, as suggested in some publications (Brandenburg & Skottke, 2014). However, Level 3 systems in current development might require take-overs more often than not, due to technical limitations, forcing drivers to regain control of the vehicle in short intervals. Additionally, we observed clear learning behavior during our experiment leading to quicker reaction times. This is either due to the fact that participants learned how to interact with the novel ADAS, or that they became more alert throughout the drives. However, as the visual attention towards the NDRT did not reduce, it seems that the first explanation is more likely. The in-depth analysis of the situation types presented that premature visibility of take-over situations allows the drivers to react significantly quicker, see Table 3. An important factor that could have caused a major influence on the results of this work is the effect of warning times, hence the TTC. Walch (2015) did not find any effect in these times, however, this was not controlled for in this study and could have taken an influence.

An important finding of the present study is that take-over quality can only be investigated if take-over situations are more confining. Allowing drivers to respond with a variety of reactions with minimal risk of crashing, as long as the obstacle causing take-over is obverted, does not produce enough precise information on take-over quality. We found that although drivers obverted critical obstacles with low TTCs, the gaze behavior analysis offered indications that the take-overs were not of sufficient quality. In the future it could be possible to restrict the amount of correct responses to reduce the amount of crash avoidance. This possibility would also allow for in depth analysis of gaze behavior and the reconstruction of situation awareness at take-over.

Quick take-overs do not automatically guarantee a safe take-over, for which we found indicators such as missing mirror gazes. This could cause severe accidents. It is possible that different strategies are in place if TTC are low, either leading drivers to react quickly without respect for lane markings or safety, with higher percentage of braking maneuvers. Distinguishing between strategies is difficult without subjective input from participants and has to be considered in future studies. In general however, truck drivers seem to present the capability of showing quicker reaction times when compared to published results on passenger car drivers. Based on these findings, a time in which the driver can safely regain control of a vehicle could not be determined reliably. This leads us to question whether Level 3 can or should be made available for drivers. Daimler Trucks recently announced that the benefits of Level 3 do not outweigh the costs of such a system and will skip development of this level (Daimler, 2019).

The investigation of influencing factors on take-over behavior has shown that especially the external environment determines a large part on how quick a driver will react. As this environment can be monitored with sensors, our underlying question regarding individual warning times based on influencing factors seems to be possible to some extent. Further research will consider external traffic models and monitoring to implement a first prototypic varying warning time.

Acknowledgements

We thank our colleagues for continuous technical support, for assistance in carrying out the empirical study and the complete team of the Mercedes-Benz driving simulator.

Appendix A

Experimental Factor	Value
Number of lanes	2 Lanes
Minimal radius of curves	500 m
Length of drive	Approx. 50 km including tutorial
Width of lanes	3.5 m
Color of lane markings	White
	Take-over situation *Lane narrowing* - yellow on right boundary
Maximum speed of traffic	180 km/h
Minimal speed of traffic	65 km/h
Speed of ego-vehicle during Level 3	80 km/h

Appendix A (*continued*)

Experimental Factor	Value
Distance to lead vehicle	55 m
Number of vehicles in environment	14 vehicles/km randomly dispersed and appearing 500 m behind ego-vehicle
Time to Collision calculation	From take-over request to collision object (including acceleration/ deceleration of objects e.g. Cut in Left)

Appendix B

Take-over Situation	Description
Fog & Warning Triangle	Weather gradually worsened starting 1 min before take-over until visibility was reduced to 60 m at which a warning triangle was situated at the side of the road within ego lane. The warning triangle was located on the right lane marking
Stranded Vehicle	A stranded vehicle situated half on ego lane and half on the hard shoulder which was circumnavigated by the lead truck abruptly only then making it visible. The lead truck was traveling at 80 km/h and was 55 m ahead of the ego-vehicle
Lost Cargo	Lost wooden box situated on the side of ego lane which was circumnavigated by lead truck abruptly only then making it visible
Lane Narrowing	Pylons and concrete barriers narrowing lane. Warning was display approx 1 s before first pylon
Construction	Lead truck switched lane abruptly to give way to a construction site. Ongoing traffic gave way to drivers who needed to perform a lane change
Stabilization	Washed out lane markings in a curve representing a system boundary. TTC was calculated as the time to which the vehicle would hit crash barriers if trajectory were kept. Note, the hard shoulder was not present in this condition to reduce TTC
Cut in Left	A vehicle cutting in from the left and decelerating immediately in front of ego vehicle
Cut in Right	A vehicle cutting accelerating from the hard shoulder immediately in front of ego vehicle. Distance to lead truck was increased gradually before take-over to 85 m
Lead Braking	Abrupt deceleration of lead truck from 80 km/h to 45 km/h within 2 s

Appendix C

Dependent variable	Measure	Sum of squares	df	Mean square	F	Pr (>F)
Lateral Acceleration	NDRT	1.10	1	1.077	0.492	0.483
	Trial	15.2	8	1.898	0.866	0.545
	Situation Type	11.7	8	1.460	0.666	0.721
	CAD Duration	1.90	2	0.967	0.442	0.643
	NDRT * Trial	13.3	8	1.657	0.756	0.642
	NDRT * Situation Type	19.5	8	2.440	1.114	0.351
	Trial * Situation Type	22.8	11	2.075	0.947	0.494
	NDRT * CAD Duration	9.20	2	4.586	2.093	0.124
	Trial * CAD Duration	19.9	6	3.309	1.510	0.172
	Situation Type * CAD Duration	0.00	1	0.025	0.011	0.915
	NDRT * Trial * Situation Type	0.80	2	0.398	0.182	0.834
Longitudinal Acceleration	NDRT	0.33	1	0.326	1.103	0.294
	Trial	1.57	8	0.196	0.662	0.726
	Situation Type	0.66	8	0.083	0.280	0.972
	CAD Duration	0.62	2	0.309	1.045	0.352
	NDRT * Trial	1.54	8	0.192	0.650	0.735
	NDRT * Situation Type	2.95	8	0.369	1.248	0.268

(*continued on next page*)

Appendix C (*continued*)

Dependent variable	Measure	Sum of squares	df	Mean square	F	Pr (>F)
	Trial * Situation Type	2.36	11	0.214	0.724	0.716
	NDRT * CAD Duration	0.20	2	0.102	0.346	0.708
	Trial * CAD Duration	2.51	6	0.419	1.417	0.205
	Situation Type * CAD Duration	0.12	1	0.118	0.399	0.528
	NDRT * Trial * Situation Type	0.18	2	0.091	0.307	0.736

Appendix D. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2019.06.008.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). *Shifting intentional set: Exploring the dynamic control of tasks*. Cambridge: MIT Press.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(1982), 369–406.

Baltzer, M. C., Rudolph, C., López, D., & Flemisch, F. (2016). Cooperative guidance and control in highly automated convoys - StrAsRob. In Neville Stanton (Ed.), *Advances in human aspects of transportation*. Florida, USA: Springer Verlag.

Bellem, H., Klüver, M., Schrauf, M., Schöner, H. P., Hecht, H., & Krems, J. F. (2017). Can we study autonomous driving comfort in moving-base driving simulators? A validation study. *Human Factors*.

Brandenburg, S. & Skottke, E. M. (2014). Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving? In IEEE 17th International Conference on Intelligent Transportation Systems (ITSC). p.2978–2983.

Bueringer, H. (2007). *Entwicklung des Straßenverkehrs in Baden-Wuerttemberg - Jahresfahrleisungen mit Kraftfahrzeugen* (pp. 37–41). Stuttgart: Statistisches Monatsheft Baden-Wuerttemberg.

Carsten, O., Lai, F. C., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: does it matter what aspects are automated? *Human Factors, 54*(5), 747–761.

Continental (2017). Continental Automotive. Advanced Driver Assistance Systems. [Cited: 08 01, 2018] https://www.continental-automotive.com/en-gl/Passenger-Cars/Chassis-Safety/Advanced-Driver-Assistance-Systems/Radars/Long-Range-Radar/Advanced-Radar-Sensor-ARS441.

Daimler A.G. (2019). Daimler. Daimler Trucks invests half a billion Euros in highly automated trucks. January 07, 2019. [Cited: April 15, 2019] https://media.daimler.com/marsMediaSite/en/instance/ko.xhtml?oid=42188247.

Damböck, D. (2013). *Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme*. München: Technische Universität München, Lehrstuhl für Ergonomie. Dissertation.

Damböck, D., & Bengler, K. (2012). Übernahmezeiten beim hochautomatisierten Fahren. *Tagung Fahrerassistenz, 15*(2012), 16.

de Winter, Joost C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F, 27*(2014), 196–217.

Department for Transport (2015). The pathway to driverless cars: summary report and action plan.

Deutscher Bundestag (2016). Vermeidung von LKW-Auffahrunfällen auf Bundesautobahnen durch Notbremssysteme. WD 5 - 3000 - 069/16.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors Journal, 37*(1), 32–64.

Eriksson, A., & Stanton, N. A. (2017). Take-over time in highly automated vehicles: non-critical transitions to and from manual control. *Human Factors*.

Feldhütter, A., Gold, C., & Bengler, K. (2016). How the duration of automated driving influences take-over performance and gaze behavior. Kongress der Gesellschaft für Arbeitswissenschaft.

Gasser, T. M., Arzt, C., Ayoubi, M., Bartels, A., Bürkle, L. ... Vogt, W. (2012). Rechtsfolgen zunehmender Fahrzeugautomatisierung. Berlin: Berichte der Bundesanstalt für Straßenwesen. Unterreihe Fahrzeugtechnik (83).

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 Vol. I, Vol. 57 (pp. 1938–1942).

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.

Gold, C., Naujoks, F., Radlmayr, J., Bellem, H. & Jarosch, O. (2017). Testing scenarios for human factors research in level 3 automated vehicles. In Advances in Human Aspects of Transportation: Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, July 17-21 (ed. Neville A. Stanton). The Westin Bonaventure Hotel, Los Angeles, California, USA: Springer.

Gold, C., Lorenz, L., Bengler, K. (2014). Influence of automated brake application on take-over situations in highly automated driving scenarios.Paper presented at the Proceedings of the FISITA 2014 World Automotive Congress.

Goncalves, J., Happee, R., & Bengler, K. J. (2016). Drowsiness in conditional automation: Proneness, diagnosis and driving performance effects. In Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems, ITSC 2016 (pp. 873–878). http://doi.org/10.1109/ITSC.2016.7795658.

Hjälmdahl, M., Krupenia, S., & Thorslund, B. (2017). Driver behaviour and driver experience of partial and fully automated truck platooning - a simulator study. *European Transport Research Review, 9*, 1.

Janson, A., & Lindqvist, C. (2017). Working in a Highly Automated Truck - Who, What and Why? Stockholm: KTH Industrial Engineering and Management Machine Design. MMK 2017:112 MCE 361.

Jarosch, O., Kuhnt, M., Paradies, S., & Bengler, K. (2017). It's Out of Our Hands Now! Effects of Non-Driving Related Tasks During Highly Automated Driving on Drivers' Fatigue. In Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design (pp. 319–325).

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*, 4.

Kraftverkehrsstatistik (2016). Kraftverkehr- Güterverkehr und Fahrleistung: Verkehr in Kilometer. Flensburg: Kraftfahrt-Bundesamt.

Lotz, A., & Weissenberger, S. (2019). Predicting take-over times of truck drivers in conditional autonomous driving. *Advances in Intelligent Systems and Computing. Advances in Human Aspects of Transportation, 2019*, 329–338.

Louw, T., Merat, N., & Jamson, H. (2015). Engaging with Automation: To be or not to be in the loop? In Proceedings of the 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Marberger, C., Mielenz, H., Naujoks, F., Radlmayr, J., Bengler, K., & Wandtner, B. (2017). Understanding and Applying the Concepf of "Driver Availability" in Automated Driving. In Advances in Human Aspects of Transportation: Proceedings of the AHFE 2018 International Conference on Human Factors in Transportation July 17-21 (ed. Neville A. Stantion). The Westin Bonaventure Hotel, Los Angeles, California, USA: Springer.

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Reasearch Part F: Traffic Psychology and Behaviour, 27*(2014), 274–282.
Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Advances in Human Aspects of Transportation part I, 431*, 7.
Naujoks, F., & Neukum, A. (2017). A review of non-driving-related tasks used in studies on automated driving. *International Conference on Applied Human Factors and Ergonomics, 597*, 525–537.
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 F, 38*(2016), 67–82.
Neubauer, C., & Matthews, G. (2012). The effects of cell phone use and automation on driver performance and subjective state in simulated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*, 2012.
Nilsson, P., Laine, L., & Jacobson, B. (2017). A simulator study comparing characteristics of manual and automated driving during lane changes of long combination vehicles. *Transactions of Intelligent Transportation Systems, 9*(18), 2514–2524.
Ohn-Bar, E., & Trivedi, M. M. (2016). What makes an on-road object important? *International Conference on Pattern Recognition, 23*(2016), 3392–3397.
Ohn-Bar, E., Martin, S., Tawari, A., & Trivedi, M.M. (2014). Head, Eye, and Hand Patterns for Driver Activity Recognition. In 22nd IEEE International Conference on Pattern Recognition.
Petermann-Stock, I., Hackenberg, L., Muhr, T., & Mergl, C. (2013). Wie lange braucht der Fahrer - eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren.
Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*(1), 2063–2067.
Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*(I), 2063–2067.
Richardson, N., Flohr, L., & Michel, B. (2018). Takeover Requests in Highly Automated Truck Driving: How Do the Amount and Type of Additional Information Influence the Driver-Automation Interaction? *Multimodal Technologies Interact, 2*(4), 68. https://doi.org/10.3390/mti2040068.
Royal Automobile Club Foundation for Motoring (2017). Road Traffic Estimates: Great Britain 2016. Gillingham: Department of Transport.
Schermers, G., Malone, K. M., & Van Arem, B. (2004). *Dutch Evaluation of Chauffeur Assistant (DECA): Traffic flow effects of implementation in the heavy goods vehicles sector* (p. 2004). Nagoya, Aichi: ITS World Congress.
Schmidt, J., Braunagel, C., Stolzmann, W., & Karrer-Gauß, K. (2016). *Driver drowsiness and behavior detection in prolonged conditionally automated drives* (p. 2016). Gothenburg, Sweden: IEEE Intelligent Vehicles Symposium.
Schneider, J. H. (2009). *Modellierung und Erkennung von Fahrsituationen und Fahrmanövern für sicherheitsrelevante Fahrerassistenzsysteme* (p. 2009). Fakultät für Elektrotechnik und Informationstechnik: Technische Universität Chemnitz, Chemniz.
Standard, SAE International (2018). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. J3016 2018.
Statista (2017). Das Statistik-Portal. Höhen des Frauenanteils bei den sozialversicherungspflichtig beschäftigten Berufskraftfahrern in Deutschland in den Jahren 2012 bis 2016. [Cited: 12 12, 2017.] https://de.statista.com/statistik/daten/studie/202938/umfrage/hoehe-des-frauenanteils-bei-den-kraftfahrzeugfuehrern-in-deutschland/.
Statistisches Bundesamt (2013). Verkehr auf einen Blick. Wiesbaden, Germany: s.n. 0080006-13900-1.
Streit, T., Chlond, B., Weiß, C., & Vortisch, P. (2014). Deutsches Mobilitätpanel (MOP) - Wissenschaftliche Begleitung und Auswertung Bericht 2012/2013: Alltagsmobilität und Fahrleistungen. Karlsruhe, Germany: Karlsruher Institut für Technologie (KIT) Institut für Verkehrswesen, 2014. Forschungsprojekt FE-NR. 70.0864/2011.
U.S. Department of Transportation Federal Highway Administration (2014). Annual Vehicle Distance Traveled in Miles and Related Data. Washington D.C., USA : Policy and Governmental Affairs Office of Highway Policy Information, 2014. VM-1.
Vogelpohl, T., & Vollrath, M. (2017). Tiredness and level 3 - automated driving. Berlin : GDV The German Insurers, 2017. Compact accident research 70.
Vogelpohl, T., Vollrath, M., Kühn, M., Hummel, T., & Gehlert, T. (2016). *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung* (p. 2016). Berlin: Gesamtverband der Deutschen Versicherungswirtschaft e.V.
von Benda, H., Hoyos, C., & Schaible-Rapp, A. (1983). *Klassifikation und Gefährlichkeit von Strassenverkehrssituationen* (p. 1983). Germany: Bundesanstalt für Strassenwesen, Bereich Unfallforschung.
Walch, M., Lange, K., Baumann, M., & Weber, M. (2015). Autonomous Driving: Investigating the Feasibility of Car-driver Handover Assistance. In Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. Automotive UI '15 (pp. 11–18).
Weinbeer, V., Muhr, T., Bengler, K., Baur, C., Radlmayr, J., & Bill, J. (2017). Highly automated driving: How to get the driver drowsy and how does drowsiness influence various take-over-aspects? In Proceedings of the Tagung Fahrerassistenz (ed. M. Lienkamp). München: Volume 8.
Weinbeer, V., Muhr, T., & Bengler, K. (2019). Automated Driving: The Potential of Non-driving-Related Tasks to Manage Driver Drowsiness. In Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018). Advances in Intelligent Systems and Comuting (ed. S. Bagnara, et al.). Florence: Springer, 2019, pp. 179-188.
Wulf, F., Zeeb, K., Rimini-Döring, M., Arnon, M. & Gauterin, F. (2013). Effects of Human-Machine Interaction Mechanisms on Situation Awareness in Partly Automated Driving. In 16th International IEEE Conference on Intelligent Transportation Systems.
Young, M. S., & Stanton, N. A. (2007). Back to the future: brake reaction times for manual and automated vehicles. *Ergonomics, 1*(50), 46–58.
Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis & Prevention, 78*(2015), 212–221.
Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention., 2016*, 230–239.
Zhang, B., Wilschut, E., Willemsen, D., & Martens, M. (2017). Driver Response Times when Resuming Manual Control from Highly Automated Driving in Truck Platooning Scenarios. Conference: RSS2017 - Road Safety & Simulation international conference. 2017, The Hague, Netherlands.

3 Paper No. 2

Lotz, A. and Weissenberger, S. (2019). Predicting take-over times of truck drivers in conditional autonomous driving. In: Stanton N. (Ed.) Advances in Human Aspects of Transportation. AHFE 2018. *Advances in Intelligent Systems and Computing, 786*, pp. 329-338. doi:10.1007/978-3-319-93885-1_30

Predicting Take-Over Times of Truck Drivers in Conditional Autonomous Driving

Alexander Lotz(✉) and Sarah Weissenberger

Daimler AG, Daimler Trucks Advanced Engineering,
HPC T332, 70546 Stuttgart, Germany
{rene_alexander.lotz, sarah.weissenberger}@daimler.com

Abstract. Conditional autonomous driving requires the description of sufficient time reserves for drivers in take-over situations. The definition of this time reserve has not been addressed for the truck context thus far. Through the observation of physiological measures, the possibility of estimating reaction times is considered. Driver data is collected with a remote eye-tracker and body posture camera. Empirical data from a simulator study is utilized to train and compare four machine learning algorithms and generate driver features. The estimation of take-over times is defined as a classification problem with four reaction time classes, leading to a misclassification rate of a linear support vector machine (SVM) of 38.7%. Utility of driver features for reaction time estimation are discussed.

Keywords: Conditional autonomous driving · Take-over prediction
Human factors · Machine learning · Control transition

1 Introduction

One of the prime interests of recent automotive studies is how attention, communication and responsibilities are allocated, when humans and machines share control of a vehicle. While this surge prevails predominantly for passenger cars, the truck context has not been regarded as intensely. Currently, the definition according to SAE J3016 [1] classifies six stages of autonomy for vehicles. Conditional autonomous driving (CAD), which classifies as Level 3, and higher levels allow drivers to disengage from the driving task and focus their attention on non-driving-related tasks (NDRT). Thereby, possibly, the most critical stage is Level 3, in which drivers are required to act as a fallback if system limits are met and accordingly only receive restricted time to take back control. Following the definition of CAD, drivers are no longer required to constantly supervise the vehicle when the automation function is activated. Adapting the responsibility of constant monitoring of the truck would revolutionize logistics and leave drivers to focus on different aspects of their work (e.g. paper work) and other NDRT. At take-over, when the automation function can no longer handle the driving situation, drivers are expected to perceive, address and solve given situations in limited time. Due to the switch of control between driver and vehicle, human factors such as the ironies of automation [2], have to be taken into account. Generally, the truck context offers a more confined set of use cases with less variation in the type of driving

N. Stanton (Ed.): AHFE 2018, AISC 786, pp. 329–338, 2019.
https://doi.org/10.1007/978-3-319-93885-1_30

situations that have to be managed by the advanced driving assistant system providing autonomous driving. Long distance hauls by heavy duty trucks on motorways comprise 21% of freight traffic in Germany and professional drivers' duties [3]. The implementation of a Level 3 system could entail the possibility for drivers to engage in NDRT, improve average traffic flow rates and cause a change in legislation, currently restricting operating times for drivers in some countries, e.g. Germany [4]. As take-over situations vary and driver attention to NDRT is not an inter- or intra-individual constant, the system has to allow for all drivers to transition back to the driving task. Time reserves for an adequate transition have to either be defined globally, with a fixed time reserve sufficient for all drivers, or individually. For individual time reserve definition, ideally, reactions by the driver would be predictable through objective measures, allowing the system to calculate take-over times. This definition of time reserves is far more complex, as all or the most influential factors need monitoring and consideration. Vogelpohl et al. [5] identify four groups of influencing factors on reaction times that will be regarded in this study. (1) *environmental* parameters (e.g. traffic density, weather, complexity of take-over situation), (2) *human machine interaction* (HMI) parameters (e.g. warning modalities, interfaces), (3) *driver* parameters (e.g. visual attention, demographics, individual driving style) and (4) *vehicle* parameters (e.g. field of vision, advanced driver assistance systems).

1.1 Take-Over Times

A variety of take-over times are relevant in the take-over process and defined by Damböck [6] at the occurrence of a "Request to Intervene" (RtI) in CAD. These times include the time until visual attention is shifted back to road (time to eyes on road; TTEoR), time until hands are placed on the steering wheel (time to hands on steering-wheel; TTHoS) and time until a first reaction is given at the steering wheel or accelerator pedals (time to first reaction; TTFR). The measurement of said times allow the evaluation of the take-over process, consisting of physiological attention allocation and cognitive processing [7].

1.2 Prediction of Driver Reactions

Although driver assistance systems should reduce stress and accidents, an ill chosen time reserve at take-over can paradoxically cause the opposite [8]. It is paramount to guarantee this adequate time reserve. Wulf et al. [9] identify the cognitive and motoric availability as essential for take-over readiness. Additionally, [10, 11] show that situation awareness [12] is decisive for take-over times and the quality. Therefore, the implementation of a driver observation could offer relevant data to guarantee sufficient take-over times, as one of the four influencing factors on reaction times. It is the goal of this study to investigate the feasibility of training machine learning algorithms to predict take-over times.

In order to predict take-over times, this study focuses on the recording of behavioral driving data during CAD in heavy duty trucks and investigates the features generated from these recordings. In the driving context of CAD, Braunagel et al. [13] show that machine learning can be applied to identify four different NDRT, based on eye-tracking, with an accuracy of 82%. Apart from eye-tracking, Ohn-Bar et al. [14] additionally investigates hand and head patterns to detect where drivers are interacting. In an extension of their work, Braunagel et al. [15] present an architecture classifying the driver readiness and considering traffic situations for CAD, reaching an overall accuracy of 79%. None of these known architectures have been tested for trucks or the prediction of reaction times.

2 Estimation Concept for Reaction Times

A concept to estimate reaction times at RtI during CAD with NDRT is be presented. All necessary ground truth data for such a concept was collected during a simulator study in the Mercedes-Benz Moving-Base-Simulator with professional truck drivers (n = 88). The experimental setup was designed to generate multiple take-overs per participant, producing data for our estimation approach. When considering the abovementioned four groups of influencing factors on reaction times by [5], the experimental setup was configured to have differentiating values for *environmental* and *driver* parameters, while keeping *HMI* and *vehicle* parameters constant. During CAD, participants were encouraged to interact with two NDRT, to allow for realistic distraction. The available NDRT were either a geography game or a video.

Eye-tracking is a common method for collection of behavioral data, especially important during the dynamic driving task [16]. In order to accumulate data distraction free, i.e. non-invasive, a remote eye-tracker (Smart Eye Pro) collected gazes and head position. Additional interaction by the driver with the environment was collected through a fixed tablet (non-nomadic device) on which NDRT are presented. A Microsoft Kinect gathered RGB-D pictures. These pictures were utilized to calculate participants' body posture during their CAD drive. The extraction of body posture is an adaptation of [17], in which a pre-trained convolutional neural network is retrained for the truck context. The second group of influencing factors, *environmental* parameters [7] are taken into account through the simulator. Examples of these parameters are the take-over situation and the time to collision (TTC), i.e. the lead time to obstacles at RtI.

2.1 Online and Offline Datasets

Two different datasets are generated from the eye-tracking, body posture and simulator data. When considering 755 take-over procedures, each trial is divided into a time series before the RtI in which the driver is not required to steer the vehicle and a time series when manual control is necessary by the driver after the RtI. After an RtI is issued, certain additional features can be calculated that depend on the occurrence of an RtI and the time series in which the driver maintains manual control. This includes all the reaction times such as time to eyes on road (TTEoR) and time to hands on steering wheel (TTHoS) apart from TTFR, which is the desired estimation output. These features may hold high contents of information that generate higher estimation results due to temporal interrelation. However, such features are not applicable for the online estimation of reaction times, as they can only be considered after the presentation of an RtI. Therefore, two separate datasets are configured:

- *Online dataset*: with features computable at any instance during CAD
- *Offline dataset*: all computable features for each trial.

2.2 Feature Generation

To allow the estimation of TTFR, features need to be generated holding preferably high contents of information about driver and/or environment. Prior to training machine learning algorithms with manually configured features, a selection according to correlation is performed. All features are either static singular values or a time series for each of the 755 take-over situations (trials). The time shortly before an RtI probably holds the most information with regard to the take-over capability of a driver and will be considered. To bypass the possibility of manually defining a time interval that may not hold any relevant information, four different time intervals are considered; the complete CAD trial duration, 10 s, five seconds and two seconds before an RtI. Singular values, such as average, standard deviation, minimum and maximum of the abovementioned time intervals are calculated. This methodology is applied for the feature generation of eye-tracking, body posture and simulator data.

The feature generation of eye-tracking and simulator data yields a total of 238. In addition, the calculation of features from extracted body posture of 12 body joints yields 135 further features.

2.3 Feature Evaluation and Selection

Reaction time TTFR is subdivided into four time classes for the application of supervised machine learning classification algorithms. Figure 1 displays a graphic representation of a frequency distribution of all 755 TTFR. The boundaries of the four

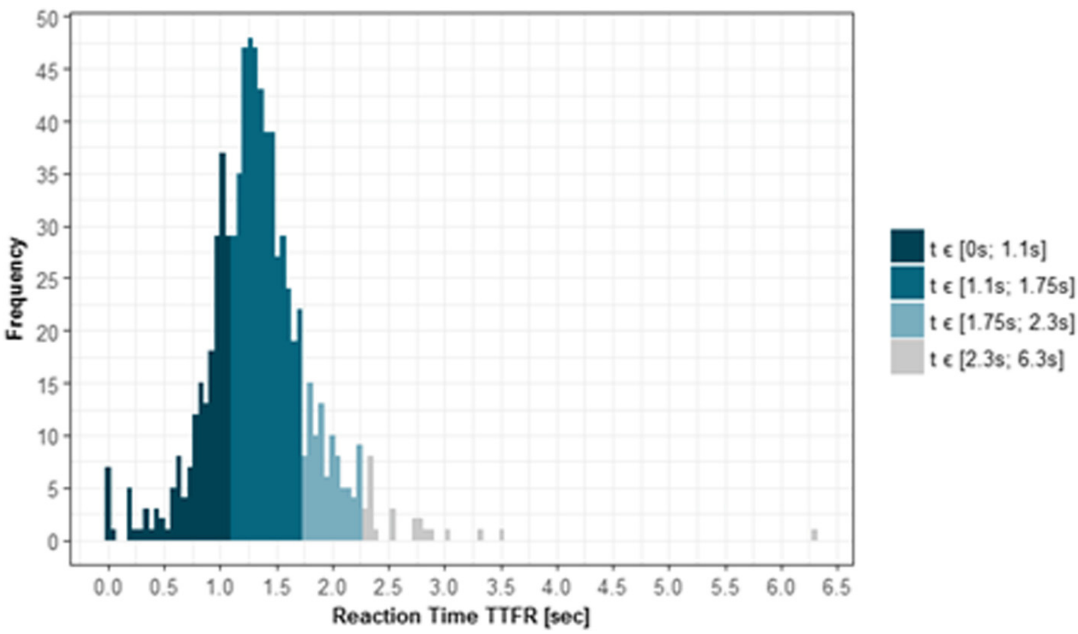

Fig. 1. Frequency distribution of time to first reaction (TTFR). Reaction times are categorized according to the definition found in Table 1 by application of the clustering algorithm k-means (k = 4).

TTFR classes are calculated by application of the k-means algorithm. The classes are partitioned as presented in Table 1.

Table 1. The dataset of all 755 take-overs are subdivided into four classes through a k-means algorithm.

Class	Definition	Number of Take-Overs
1	t ϵ [0 s; 1.1 s]	212
2	t ϵ [1.1 s; 1.75 s]	430
3	t ϵ [1.75 s; 2.3 s]	91
4	t ϵ [2.3 s; 6.3 s]	22

A filter approach is applied for the evaluation of all previously generated features. A multivariate analysis of variances (MANOVA) is implemented to identify those features that show the highest correlation between input and output. A total of 30 features with the highest congruence for TTFR are selected for the application of the four different machine learning algorithms. Table 2 holds these features for the *online dataset*. Notably, by analysis of the offline dataset only one feature is replaced in this list. Time to hands on steering-wheel (TTHoS) holds the most information for the estimation of TTFR in the offline dataset and is ranked first.

Table 2. The 30 most important features selected from the Online dataset through application of MANOVA.

Rank	Feature	Rank	Feature
1	Trial Number (# of take-overs)	16	Number of Blinks 10 s before RtI
2	Time to Collision	17	Avg. Head Heading
3	Blink Duration	18	Max Gaze Direction Change 5 s before RtI
4	Avg. Head Heading 2 s before RtI	19	Head Roll 0.5 s before RtI
5	Take-over Situation type	20	Avg. Head Roll
6	Head Roll Average 2 s before RtI	21	Number of Blinks 5 s before RtI
7	Min Head Pos. Change X-Y-Z	22	Blink Duration 5 s before RtI
8	Head Pos. Change X 0.5 s before RtI	23	Max Head Pitch 2 s before RtI
9	Head Pos. Change Y 3 s before RtI	24	Head Pitch 1 s before RtI
10	Min Head Pos. Change X-Y-Z 10 s before RtI	25	Min Head Pos. X-Y-Z 2 s before RtI
11	Pupil Diameter 0.5 s before RtI	26	Min Head Pos. Change 2 s before RtI
12	Gaze Pitch St. Dev. 5 s before RtI	27	Head Pitch 1.5 s before RtI
13	Sum of tablet taps 5 s before RtI	28	Min Head Pos. X-Y-Z 10 s before RtI
14	Avg. Head Pos. X-Y-Z 2 s before RtI	29	Max Head Pitch 5 s before RtI
15	Gaze Pitch St. Dev Total	30	Avg. Gaze Direction X

2.4 Machine Learning Algorithm

Four different machine learning methods will be analyzed for this applied problem. For the estimation of TTFR, which is defined as a classification problem, the following machine learning algorithms are tested:

- k-Nearest Neighbors algorithm
- Bayes classifier
- Support Vector Machine
- MLP Neural Network

In terms of a machine learning problem, the dataset collected through this simulator study with 755 instances is rather small. Complex algorithms with a high share of hyper parameters are not suited for this problem as there is no abundance of data. An interaction between many of the abovementioned different features is expected for the estimation of the output variable, i.e. TTFR. Therefore, a trade-off between the complexity of the algorithm and the size of the dataset has to be accepted. Two simple and two more complex machine learning algorithms are chosen.

All tested algorithms are validated by means of a stratified n-fold cross-validation. As proposed by [18], a 10-fold cross-validation is chosen. The complete dataset is divided into a training and test dataset, with a distribution of 80% to 20% respectively.

3 Results

As a quality measure for the classification by all machine learning algorithms, misclassification rate is chosen. The misclassification rate hereby describes the proportion of misclassified observations that were predicted into one of the four TTFR classes, see Table 1 and Fig. 1.

As the estimation of the online dataset bears the most relevance for the CAD context, detailed results are presented. A comparison between additional information from body postures and to the offline dataset are produced.

3.1 Online Estimation

In the following, all four machine learning algorithms will be compared by means of misclassification rates for the online dataset. Especially the third and fourth class are not represented evenly in the data (C3: 12%; C4: 3%), see Table 1. While these times are more seldom due to the Gaussian distribution of reaction times for this study, especially in the context of take-over times these instances are the most relevant as they represent critical take-overs due to slow reaction times. To compensate for the unbalanced classes of the dataset, a cost matrix is introduced. This facilitates weighted penalties for misclassifications.

k-Nearest Neighbours: For the definition of the optimal k-NN the hyper parameters k and the distance function are iterated. The Mahalanobis distance function exceeded the results of the Euclidean distance function. Best results were received for $k = 8$, leading to a misclassification rate of 40.8%.

Bayes classifier: Again, the Mahalanobis distance function provided the best results. A feature scaling did not improve results. Minimal misclassification was achieved at 39.6%.

Support Vector Machine: Three different SVM types are tested, i.e. linear, Gaussian and polynomial. Hereby, the linear and Gaussian SVM achieved a misclassification rate of 38.7%, while the polynomial SVM performed 39.2%.

Neural Network: The optimal misclassification rate for the neural network of 40.0% was achieved with one layer and five nodes.

Overall, the best results are obtained for the linear SVM with a misclassification rate of 38.7%. Table 3 displays a comparison of the classification results of the tested algorithms.

Table 3. Comparison of the best classification results of four tested machine learning algorithms on online dataset

Algorithm	Misclassification rate [%]
k-NN	40.8
Bayes	39.6
SVM	38.7
NN	40.0

3.2 Offline Estimation

The following estimation on the offline dataset is obtained with TTHoS ranked first in the list of 30 most relevant features, see Table 2. Similarly to Sect. 3.1, the SVM obtains the lowest misclassification rate at 22.5% (Table 4).

Table 4. Comparison of the best classification results of four tested machine learning algorithms on offline dataset

Algorithm	Misclassification rate [%]
k-NN	38.3
Bayes	32.6
SVM	22.5
NN	26.4

3.3 Online Estimation with Body Posture

As the SVM delivered the lowest misclassification rate for online estimation without body posture features, no comparison of machine learning algorithms will be provided. Similar methodology as in Sect. 3.1 is applied. After application of a cost matrix the linear SVM misclassification rate presented a value of 37.7%. The consideration of the calculated 30 features shows that 18 new features (body posture features) are replaced in the list presented in Table 2. The highest of these new features ranking in fourth position and representing the maximal displacement of the right knee in the 10 s before RtI.

4 Discussion

Comparing the results of the online and offline classifications without body posture, see Sects. 3.1 and 3.2, a large difference of 16% between both linear SVM exists. As first ranked TTHoS is the only feature presented in the offline dataset not being available during online estimation, it seems to hold a lot of relevant information. This is not surprising, as TTFR partially depends on the motoric hand availability for adequate reaction times. Overall 76% of TTFR are induced due to a motoric reaction at the steering wheel in the data, while only 24% of first reactions were cause by braking or accelerating. While the discrepancy of the algorithm comparison for the online dataset were within 2.1%, this can be attributed to the absence of highly relevant features. In other words, the features were similarly meaningful. TTHoS however does not hold 16% more information in the offline dataset, the algorithms SVM and NN can attribute higher relevance to certain features by redistributing weights.

Based on the features calculated and due to the fact that the reaction time classes were unequally distributed, the quality of the results obtained is not sufficient to safely and reliably estimate reaction times of truck drivers. By placing a new observed reaction time into the second class, 43% of these instances the estimation would misclassify, (430 trials were in this class, see Table 1). Both the online and offline classification achieved a lower misclassification rate, showing that some information is drawn from the features calculated. There are multiple reasons why only a small reduction in misclassifications was achieved. Firstly, the quality of the data may not be sufficient enough. Either the static features calculated by hand or the variance in the data can cause a reduction of the signal quality. Secondly, the problem may be too complex for an analysis by the evaluation methods of this study. Thirdly, a major problem when addressing such a problem is the generation of data. Production of sufficient amounts of data needs time. Especially for problems in which the most relevant data, i.e. slow reaction times, is sparse, calling for extensive data collection. Lastly, human behavior may not be classifiable and features bear no consisted information on reaction times.

To the best of the authors' knowledge, no studies for CAD are published that required nine or more take-overs within 30 min. The data shows, that the trial number was categorized as a feature with high correlation to TTFR. When looking at this variable separately, learning behavior is identified. Participants might have therefore been cautious after the first couple of take-overs resulting in such quick and dense reaction times (TTFR) as seen in Fig. 1.

The additional consideration of body postures, represented by the localization of 12 body joints showed only a minimal improvement of 1% in the misclassification rate of a trained linear SVM for the online dataset. Although more than half the elements in the features list were represented by body posture values, this new remote observation of physiological signals showed minimal information gain. This could mean that body postures before a take-over hold no relevant information. Another possible explanation would be that the experimental setup did not encourage a large variety of body postures. According to *Fitts' Law* [19] movement is a function of target distance and target size. Consequently, the motoric variation seems to have been minimal in this experimental setup and did not cause for large time discrepancies. Finally, the body joint estimation method could entail flaws and supply false coordinates leading to lower signal quality.

Overall the results of this study were surprising, as times to first reactions were far below those reported in current CAD publications. Neither of the independent variables, CAD duration and NDRT, presented significant differences and 99,7% (3σ-interval) of all take-overs were completed within 2.82 s. Our work shows that although information can be drawn from static features from eye-tracking and body posture analysis, no clear correlation between multiple features and reaction times can be drawn. The improvement of measurement technology and feature calculation could reduce the misclassification rate further; however, it is unlikely that full classification can be achieve. Variance in drivers' actions seems too high for a current mathematical prediction method.

4.1 Future Work

This publication is limited to the analysis of take-over time through the implementation of classification algorithms. In future work, the problem could be considered as a regression problem, possibly leading to a better estimation. Secondly, a new classification problem with balance reaction time classes and probabilities will be computed. This, however, presents the drawback that classes with higher standard deviation from the expected Gaussian value will become large. Thirdly, non-static features will be considered for analysis. Distinct correlations between features that were not defined in the generated feature list of this work could possibly provide a lower misclassification rate.

References

1. SAE On-Road Automated Vehicle Standards Committee. Information Report J3016: Taxonomy and Definitions (2014). https://www.sae.org/misc/pdfs/automated_driving.pdf
2. Bainbridge, L.: Ironies of automation. Automatica **19**, 775–779 (1983)
3. Statistisches Bundesamt. Verkehr auf einen Blick, Wiesbaden, Germany (2013)
4. VO (EG) Nr. 561/2006 of the European Parliament (2006)
5. Vogelpohl, T., Vollrath, M., Kuehn, M.: Uebergabe von hochautomatisiertem Fahren zu manueller Steuerung. Forschungsbericht **47** (2017)
6. Damböck, D.: Automationseffekte im Fahrzeug – von der Reaktion zur Übernahme. München, Technische Universität München, Dissertation (2013)
7. Petermeijer, S.M., Hornberger, P., Ganotis, I., de Winter, J.C.F, Bengler, K.J.: The Design of vibrotactile seat for conveying take-over requests in automated driving. In: Stanton, N.A. (ed.) Advances in Human Aspects of Transportation. Advances in Intelligent Systems and Computing, vol. 597, pp. 618–630. Springer, Heidelberg (2018)
8. Regan, M.A., Lee, J.D., Young, K.: Driver Distraction: Theory, Effects, and Mitigation, pp. 85–105. CRC Press, Boca Raton (2008)

9. Wulf, F., Zeeb, K.: Effects of human-machine interaction mechanisms on situation awareness in partly automated driving. In: 16th International IEEE Annual Conference on Intelligent Transportation Systems, pp. 2012–2019 (2013)
10. Rudin-Brown, C.M., Parker, H.A.: Behavioural adaptation to adaptive cruise control (ACC): implications for preventive strategies. Transport. Res. F Traffic Psychol. Behav. **7,** 59 (2004)
11. Radlmayr, J., Gold, C., Lorenz, L., Farid, M., Bengler, K.: How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In: Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, pp. 2063–2067 (2014)
12. Endsley, M.R.: Toward a general theory of situation awareness in dynamic systems. Hum. Factors **37**(1), 32–36 (1995)
13. Braunagel, C., Rosenstiel, W., Kasneci, E.: Ready for take-over? A new driver assistance system for an automated classification of driver take-over readiness. In: IEEE Intelligent Transportation Systems, pp. 10–22 (2017)
14. Ohn-Bar, E., Martin, S., Tawari, A., Trivedi, M.M.: Head, eye, and hand patterns for driver activity recognition. In: 22nd International Conference on Pattern Recognition, pp. 660–665 (2014)
15. Braunagel, C., Rosenstiel, W., Kasneci, E.: Online Recognition of Driver-Activity based on Visual Scanpath Classification. In: IEEE Intelligent Transportaion Systems, pp. 23–36 (2017)
16. Gold, C., Damböck, D., Bengler, K., Lorenz, L.: Partially automated driving as a fallback level of high automation. In: Tagung Fahrerassistenz, vol. 6 (2013)
17. Insafutdino, E., Pishchulin, L, Andres, B., Andriluka, M., Schiele, B.: Deeper cut: a deeper, stronger, and faster multi-person pose estimation model. In: European Conference on Computer Vision, pp. 34–50 (2017)
18. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. Int. Joint Conf. Artif. Intell. **2**, 1137–1143 (1995)
19. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. **47**(6), 381–391 (1954)

4 Paper No. 3

Lotz, A., Russwinkel, N. & Wohlfarth, E. (2020). Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks. Cognition*, Technology and Work*. doi:10.1007/s10111-020-00626-z (licensed under CC BY 4.0)

Cognition, Technology & Work
https://doi.org/10.1007/s10111-020-00626-z

ORIGINAL ARTICLE

Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks

Alexander Lotz[1,2] · Nele Russwinkel[2] · Enrico Wohlfarth[1]

Received: 27 November 2019 / Accepted: 21 February 2020

Abstract

With the introduction of advanced driving assistance systems managing longitudinal and lateral control, conditional automated driving is seemingly in near future of series vehicles. While take-over behavior in the passenger car context has been investigated intensively in recent years, publications on semi-trucks with professional drivers are sparse. The effects influencing expert drivers during take-overs in this context lack thorough investigation and are required to design systems that facilitate safe take-overs. While multiple findings seem to cohere in passenger cars and semi-trucks, these findings rely on simulated studies without taking environments as found in the real world into account. A test track study was conducted, simulating highway driving with 27 professional non-affiliated truck drivers. The participants drove an automated Level 3 semi-truck while a non-driving-related task was available. Multiple time critical take-over situations were initiated during the drives to investigate four main objectives regarding driver behavior. (1) With these results, comparison of reaction times and behavior can be drawn to previous simulator studies. The effect of situation criticality (2) and training (3) of take-over situations is investigated. (4) The influence of warning expectation on driver behavior is explored. Results obtained displayed very quick time to hands on steering and time to first reaction all under 2.4 s. Highly critical situations generate very quick reaction times M = 0.81 s, while the manipulation of expectancy yielded no significant variation in reaction times. These reaction times serve as a reference of what can be expected from drivers under optimal take-over conditions, with quick reactions at high speed in critical situations.

Keywords Conditional automated driving · Truck drivers · Driver take-over · Test-track study · Warning expectancy

1 Introduction

The term 'automated driving' has become a global discussion point in recent years, with political and social interests rising to develop vehicles enabling automation. While the progress has been significant within the past decade, complete automation of the driving task has not yet been achieved. Of late, new assistance systems are entering the market to support drivers in restricted scenarios such as Audi's traffic jam pilot (Audi 2017), Tesla's Autopilot (Tesla 2019) and the Mercedes Active Drive Assist for trucks (Daimler 2018). These systems rank as Level 2 partial automation and Level 3 conditional automated driving (SAE J3016 2018). Especially for Level 3, the difficulty of transitions between machine and driver is an ongoing research field, in which the system reaches limits and the driver needs to regain control of the vehicle in short timeframes (Young et al. 2007). Research within the field of automation transitions has shown that human capabilities can deteriorate due to task switching (Bainbridge 1983; Wylie et al. 2000) and a lack of situation awareness (Endsley 1995; Young et al. 2002). These effects also apply to advanced assistance systems in vehicles (Brookhuis et al. 2001; de Winter et al. 2014). If these theoretical constructs apply to the vehicles with advanced driver assistance systems, the question that is at the core of ongoing research is: how long do drivers require to regain control of vehicles safely (Gold et al. 2013)?

✉ Alexander Lotz
rene_alexander.lotz@daimler.com

Nele Russwinkel
nele.russwinkel@tu-berlin.de

1 Daimler Truck AG, Stuttgart, Germany

2 Technische Universität Berlin, Marchstr. 23, Sekr. MAR 3-2, 10587 Berlin, Germany

Published online: 04 March 2020

The truck context does not contain the amount of publications as the passenger car context does for Level 3, as studies of take-over are often limited to instances without high time–pressure (Zhang et al. 2017). The effect of learning take-over situations (Lotz et al. 2019b), the influence of accustoming to a new automation function in experimental contexts and reacting to a new warning are influcences that have been investigated sporadicly in recent years (Kantowitz et al. 2009) and will be the main topic of the present study. Recent research by the ADAS&ME project also addressed the specific issue of developing new driver monitoring systems to adjust human–machine interface content in the truck on the basis of driver needs (Axelson et al. 2018). This is important as most participants engage with the different automation functions for the first time in published research and in most studies only a handful of take-over situations are presented. Distinguishing underlying human capabilities that can or cannot adjust over time is paramount.

The aims of the present study are to investigate the behavior of semi-truck drivers in a prototypic vehicle in quasi real-world Level 3 automation on a test track. This will allow insight into driver behavior outside of simulated environments and possibly generate other results as previously determined. Different take-over situations will be investigated that vary in criticality to determine possible differences in driver behavior at take-over when a non-driving-related task (NDRT) is performed. The occasion of these take-overs will also be controlled, to investigate the effect of take-over expectancy. The following introductory sections focus on the current state of research of reaction times in critical take-over situation (Sect. 1.1), the effect of expected warnings (Sect. 1.2), the resulting aims of the study (Sect. 1.3), the definition of reaction times (Sect. 1.4) and hypotheses (Sect. 1.5).

1.1 Reaction times in critical take-over situations

Multiple studies have investigated reaction times in recent years, primarily in fixed and moving-base simulators (McDonald et al. 2019). Results indicate a large variation in reaction times ranging between 3 to 8 s according to Vogelpohl et al. (2016) and with an average reaction time until control is regained of M = 2.96 s (SD = 1.96 s) (Eriksson et al. 2017). As pointed out rightly by Zeeb et al. (2015), it does not seem as though reaction times are comparable due to the large variance in influencing factors that have been categorized into clusters by Vogelpohl et al. (2016): driver, environmental, vehicle and human–machine interaction factors. All of these studies mentioned were conducted within the passenger car context. A study of time critical take-overs during Level 3 driving in the truck context provided distinctly quicker reaction times with 99.7% (3σ-Interval) of 755 take-overs ranging between 0 s and 2.82 s (Lotz et al. 2019b). A possible explanation for these quick reaction times could be that within the truck context drivers are professionals, with far more experience controlling a vehicle, higher mileage, and daily repetition. Surprising critical take-over situations seem to reduce the reaction times also in the car context, as Diederichs et al. (2015) determine a reduction of 200 ms compared to expected take-overs. Previous work also investigated the possibility of predicting reaction times based on vehicle sensor data and a driver monitoring system (Lotz et al. 2019a). Limited predictability was achieved due to restricted data and low variance between prediction classes. A comparison study between the passenger car context and truck context investigating the difference in reactions has not been conducted to the best of the authors' knowledge. The detailed analysis of current publications also presents the difficulty of comparability of results, as experimental design and measures vary throughout all studies.

1.2 Expected warnings and impact of trained transitions

Research on advanced driver assistance systems of Level 2 (SAE J3016 2018) has shown that reaction times to start a motoric response is influenced by the expectation of an event occurring (Ruscio et al. 2015). However, while an unexpected situation arose in the abovementioned experimental real-life study, differences in criticality were not investigated. In addition, repetition of unforeseen events has thus far not been conducted, eliminating the possibility of observing learning of such situations (Kantowitz et al. 2009). Melcher et al. (2015) investigate an identical take-over situation while varying integration of a NDRT and automated braking. Prior learning of requests to intervene (RtI) has displayed positive effects on take-over performance. Additionally, the trust in automation seems to increase after an adequate stage in which drivers can accustom to a new function (Hergeth et al. 2016). Further research on false warnings has also shown that false warnings can affect subsequent reactions negatively (Lees et al. 2007). Therefore, an unanswered research question that arises is 'how learning of expected situations, not warnings, and the criticality of these take-overs may affect reaction times'? We thereby differentiate training from expectation. While we define the expectancy towards a take-over situation as the current state of the driver anticipating a take-over situation will arise within a short timeframe. This assumption by the driver is formed based on the surrounding of the vehicle, explicitly the approaching towards a steep curve in this study as described in Sect. 2.2. The opposing construct of training is the effect of improving take-over procedure through learning. While this study entails aspects of task switching and the current identification of situation awareness, these constructs will not be measured explicitly.

1.3 Scope of study

The scope of the present study can be categorized into two main aspects.

1. A test-track study in a prototype vehicle providing the possibility of Level 3 automation with critical take-overs is investigated. By allowing participants to drive on a high-speed oval test-track with traffic, highly realistic highway scenarios are investigated with a prototypic vehicle rather than utilizing a simulator. This allows a comparison to previous simulator results regarding driver behavior for critical take-over situations for truck drivers. Additionally, possible shortcomings of simulator studies such as motion sickness and driving comfort as mentioned by Bellem et al. (2016) can be largely disregarded for the results of this study. Finally, traffic on the test track increases comparability to real-world driving on highways, as driver's are inclined to observe their surroundings during automation due to real rather than simulated vehicles. This generates important behavioral data regarding real-world take-over reactions. Due to safety regulations on the test track, no obstacles could be tossed in front of the vehicles at such high velocities with surrounding vehicles. Criticality, therefore, was induced through fictive malfunction of the automation function, i.e., by prompting a lateral steering swerve maneuver.
2. The test scenario intends to display two different take-over warnings throughout the drive resulting in three different take-over situation types. While all warning types are introduced in a preparatory tutorial phase, only the take-over type with the lowest criticality, in which these warnings will be initiated, is trained. This will lead to a discrepancy between learned behavior for different warnings and the current take-over situation witnessed. It is to be expected that this discrepancy causes differences in reaction times as explained in Sect. 2.2.5 and to adjust due to learning throughout the drives.

1.4 Reaction time definition

The reaction times collected within this study are defined in accordance with the work of Damböck (2013). Common take-over quality measures such as lateral variance in trajectory after take-over will not be reported, as precise localization was not possible in the prototype vehicle.

Time to eyes on road (TTEoR): the time from the beginning of the auditory and visual warning until the first fixation is started within the *windscreen* area of interest. This serves as a measure for cognitive processing and visual reaction.

Time to hands on steering (TTHoS): the time is defined from the beginning of the auditory and visual warning until hands touch the steering wheel for the first time. The measure expresses the motoric reaction time needed for the driver to regain the possibility of intervening in the lateral control of the vehicle.

Time to first reaction (TTFR): two thresholds are monitored to define the minimal reaction time measure. The minimum of the time required for pressing the ac-/decelerator pedal more than 5% or applying more than 1 Nm at the steering wheel defines the TTFR. The metric allows for the interpretation of the time required by a driver to form a situation awareness and plan a route to circumvent possible obstacles in the trajectory of a vehicle.

1.5 Hypotheses

Based on the aims of the study in Sect.1.3, the following hypotheses are generated. While the data entail far more information, the study aims are reduced to the following four hypotheses.

1. Take-over situations of higher criticality will promote quicker first reactions (TTFR), see Sect. 1.4 for definition. This does not include the time to hands on steering (TTHoS) which will be constant for all take-over situations.
2. A reaction to expected situations will occur quicker than to unexpected situations, so long the criticality is similar. This is because reactions towards expected situations can be planned prior to the warning itself. Expected and, therefore, anticipated take-over situations will yield no difference in reaction times.
3. Reaction times will be within a similar scope to the results obtained in the previous simulator study (Lotz et al. 2019b). The magnitude of the average time to first reaction will be under 2 s for highly critical take-over situations with a NDRT.
4. The engagement in the NDRT will cause the TTFR to be slower if the NDRT is not addressed.

2 Methods

2.1 Participants

A total of 27 participants were recruited through direct inquiry at 60 hauler and logistic firms. Participation was voluntary; all participants were professional truck drivers, receiving a financial incentive. Seven participants were excluded from the data analysis due to technical synchronization failures of the eye-tracker with the vehicle data and one participant mentioning discomfort driving with the automated function. The experiment could be terminated at any point in time and all participants gave an informed consent to the experiment. It was explained in detail that the truck was fitted with a prototypic function and warnings, which could arise at any moment in time. Physical wellbeing was addressed and queried prior to the start of the experiment and monitored during procedure. In addition, the experimental design underwent an internal audit by an ethical steering committee.

The remaining 20 participants were all male with a mean age of M = 47.1 years (SD = 12.0 years) and reached an average annual mileage of 63,770 km/year (SD = 38,732 km/year). A self-estimated 32.3% of the annual mileage was spent on highways, 42.0% were driven on overland roads, and the remaining 25.7% was driving in urban settings. In the sample of this study, one participant was not familiar with speedometers or adaptive cruise control, while over 90% used these systems daily.

2.2 Apparatus

2.2.1 Vehicle

A Mercedes-Benz Actros 1845 including a prototypic automation function enabling Level 3 on highways was provided for all drives. Controls of the automation function were identical to that of the series Active Drive Assist (Daimler 2018), on the right-hand side of the steering wheel. The automation function was engaged through the 'Set' button to the current speed and could be toggled to the desired speed through plus or minus controls. The second option of engaging the automation function was initiated through the 'Resume' button, reactivating the previously set speed, see Fig. 1. Automation could be terminated at any moment in time through the 'OFF'-button.

Once Level 3 driving was activated, the vehicle controlled the lateral and longitudinal maneuvering, kept the vehicle within lane markings, and decelerated if obstacles appeared ahead. Lateral control required the input of lane markings to a camera and a radar provided objects in the surroundings, to brake accordingly. Braking was not necessary within the experimental setting, as the truck was the slowest vehicle on the test track. All surrounding vehicles were steered by drivers with safety training.

2.2.2 Track

All participants were invited to the automotive testing ground in Papenburg, Germany (ATP Automotive Testing Papenburg GmbH 2018) and instructed on all safety regulations. The testing ground includes a high-speed oval with a length of 12.3 km and five lanes for simulated highway driving. Due to regulations on the track, the minimal driving speed was 90 km/h, which all drivers followed during the drives. The Level 3 driving was only initialized on the 4-km-long straights, while the steep curves were driven manually on the lowest lane. The study was conducted in December, which included quickly changing weather including sunshine, heavy rain, strong winds, and overcast skies.

2.2.3 Eye-tracker

Gaze behavior was tracked through a Smart Eye Pro eye-tracker. Mounted on the dashboard, four lenses were dispersed around the driver to monitor gazes into the windscreen, onto the tablet, into the instrument cluster and towards both rear-view mirrors. The system was utilized to record the visual attention towards a non-driving-related task during procedure. The sampling rate was 60 Hz.

Fig. 1 (Left) Cockpit overview with Smart Eye Pro eye-tracker and mounted tablet. Right: details of automation controls on the steering wheel and the tablet providing the non-driving-related task (NDRT)

2.2.4 Non-driving-related tasks (NDRT)

During activation of the automation function, interaction with NDRT was available. Participants were not encouraged to attend the NDRT, to allow them to allocate attention freely. This resulted in some drivers not attending the NDRT throughout Level 3 automation. The affected trails were handled separately during data analysis.

The NDRT was based on a previously developed interactive geography quiz (Lotz et al. 2019b). A contoured map of Germany was displayed on a mounted tablet in the central console, see Fig. 1. The task consisted of locating cities on the map by setting coordinates through touch. Chosen specifically to engage participants for longer periods and allow non-fluent speakers to interact in a quiz, this task proved as highly engaging in previous studies. To interact with the NDRT, visual and motoric attention was required. Through the new environment, the possibility of driving a highly automated vehicle, the necessity to drive manually in the curves and the novel test track environment for each participant, we are confident that passive fatigue did not occur, as discussed by Marberger et al. (2018). However, this was not controlled and will not be addressed further.

2.2.5 Take-over situations and warnings

Three different take-over situations were investigated in the present study with varying time criticality. In each situation type, one of two optical–acoustic warnings was presented. One possible combination was not tested; this resulted in three take-over situations: (1) yellow warning with low situation criticality, (2) red warning with low situation criticality, and (3) red warning with high situation criticality. A yellow warning was only presented at the end of a straight and red warnings could appear at any moment due to system limits being breached.

1. Yellow take-over situation: at the end of each straight, a yellow take control warning with a 10-s countdown was displayed in the instrument cluster. The warning consisted of a periodical tone with 200 bpm at 1200 Hz with a yellow icon in the instrument cluster, see (4) in Fig. 2. As Level 3 was not available in the track's steep turns, this automatic prompt was always displayed. To let the drivers accustom to the vehicle and surroundings, this warning was not displayed on Straight #1. The reason for the yellow warning due to the steep curves was provided to all participants. The yellow warning was, therefore, repetitive and could be expected by participants. The criticality was low, as there were more than 10 s until the steep curve began. Participants were instructed to turn off the automation function and continue driving manually when the yellow warning was displayed throughout the complete curve and reactivate automation on the following straight.
   - Yellow warning: occurring at the end of straights as an indication to regain manual control with at least 10 s headway. The vehicle continued on its trajectory with no perceivable increased criticality.
2. Red take-over situation: consisted of a red take control warning presented when the vehicle reached system limits and a continuous tone at 1200 Hz occurred, see (5) in Fig. 2. Participants were instructed to take back control of the vehicle as soon as possible when this signal appeared, as the system reached its limit. The red warning was presented multiple times during Level 3 activation, see Fig. 3. The automation function registered the take-over at the steering wheel upon which the warning terminated instantly. Participants did not need to deactivate the function with the 'OFF'-button. Additionally, the final yellow warning on Straight #9 was substituted with this red take-over warning, to investigate the response towards switched warning types and,

Fig. 2 Automation system warnings in the instrument cluster between speedometer and rev counter. Left to right: (1) automation available and ready for activation. (2) Automation function activated as symbolized through blue 'HP' symbol. (3) Activation function passive due to oversteering by participants. (4) Yellow warning at the end of all straights with a 10-s counter. (5) Red warning advising immediate take control (color figure online)

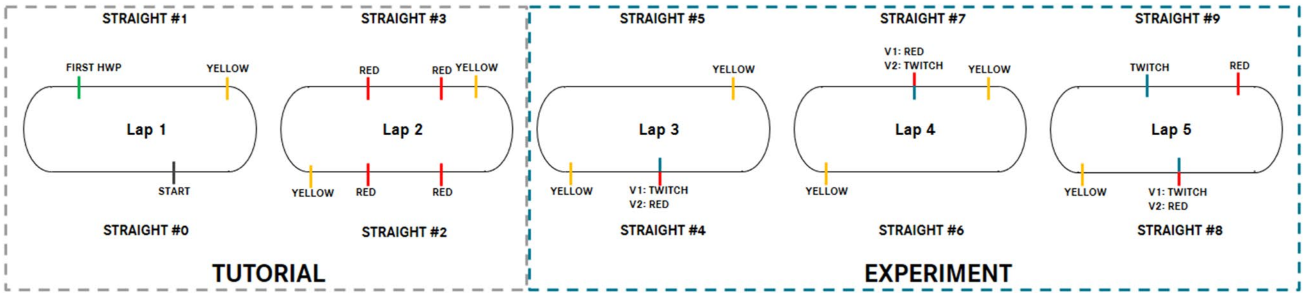

Fig. 3 Overview of the procedure of participants with the chronology of take-over types during Level 3. Three different take-over situations were presented: yellow, red and twitch (lane twitch) scenarios. The procedure was completely identical except for the take-over situations on Straight #4, #7 and #8 in which each driver drove either variant V1 or V2 (color figure online)

therefore, an unexpected warning signal with identical criticality

- Red warning: occurring anywhere on the straights with the instruction to regain control as quickly as possible. The vehicle continued on its trajectory with no perceivable increased criticality.

3. Lane twitch take-over situation: instead of warning without seeming environmental reason and low criticality, a rapid trajectory change was induced to generate a take-over situation with high criticality and the red warning was displayed. This warning was only induced on the straights without the need to regain manual control immediately in front of steep curve, due to safety reasons. This take-over type is referred to as a *lane twitch* hereon.

- Red warning and lane twitch: occurring during automation anywhere on the straights with the instruction to regain control as quickly as possible. The vehicle trajectory changed rapidly to increase criticality.

The trajectory change induced resulted in a steering wheel angle change (M = 18.68°, SD = 2.81°) to the right, a track offset (M = 0.21 m, SD = 0.12 m) and lateral acceleration (M = 0.60 m/s^2, SD = 0.25 m/s^2). A safety engineer seated on the rear seat behind the participants observed traffic in the rear-view mirrors and did not start the take-over if vehicles were within the proximity.

2.3 Procedure

After completing all safety introductions participants filled out a socio-demographic questionnaire, the automation function and NDRT were described and one lap on the test track as a passenger was completed. During all drives, one safety engineer was present in the vehicle and an additional experimenter guided the participants through the tutorial and drive. During the introductory lap, the automation function was not shown; however, the controls were clarified at the steering wheel and function-related questions were answered. The participants had no specific instructions for the drive other than to experience a prototypic automation function novices but with a professional background and high driving experience.

Participants were in control of the vehicle for five laps (Straight #0–#9), of which the first two laps consisted of a guided tutorial to accustom to the automation function. The automation function was initiated on Straight #1 for the first time, see Fig. 3. Drivers could overrule the function at any time by pressing the ac-/decelerator pedals or guiding the vehicle laterally with the steering wheel without turning the automation function off, see (3) in Fig. 2. While the automation function was being overruled, the function was passive and reinitiated itself upon terminating the overruling. During the first four straights, the participants experienced a total of seven take-overs (three *yellow*, four *red*), see Fig. 3. The first *red* take-over warning on each straight during the tutorial phase was always announced prior to the take-over. On Straight #3, the NDRT was activated the first time and the red warning symbol was displayed on the tablet during the request to intervene. The *yellow* take-over situation was announced in the tutorial phase to accustom participants to turn off the Level 3 automation when approaching the steep curves.

The actual experiment phase began with the approach onto Straight #4. The participants drove one of two scenarios in the experiment phase in which the *red* and *lane twitch* take-overs on Straight #4, #7 and #8 were reversed, see Fig. 3. The reversal of the two scenarios was aimed at investigating the influence of the tutorial phase on known take-over scenarios, *red* take-over, and unknown critical take-overs, *lane twitch*, directly after the tutorial phase. The order of the time critical take-overs is depicted in Fig. 3 for both scenarios (V1 and V2). The participants activated the automation function at the beginning of each straight and the NDRT was unlocked. The five laps were completed within approximately 45 min. Due to the fact that sunset was at 16:15 (4:15 pm); it was possible to collect data in the early phases of night as two drives started after 16:30 (4:30 pm) each day.

3 Results

The study design is a within-subject design with repeated measures resulting in 356 take-overs. The vehicle data, eye-tracking and video recordings were synchronised and analyzed with respect to the formulated hypotheses of Sect. 1.5. Due to the fact that the prototypic vehicle was not fitted with a capacitive steering wheel, the time to hands on steering (TTHoS) were evaluated from a synchronised video of the drives. For the time to first reaction (TTFR), the threshold of 1 Nm at the steering wheel or 5% ac- or decelerator pedal position change was defined.

Figure 4 depicts TTFR for both phases of the drive (tutorial/experimental) for each take-over situation type, when drivers had their hands off the steering wheel. For the experimental phase, all reaction times are documented in Table 1. On average during the experiment phase, this resulted in a TTFR (M = 1.087 s, SD = 0.417 s), TTHoS (M = 0.727 s, SD = 0.253 s), and TTEoR (M = 0.364 s, SD = 0.425 s) for a total of 162 take-overs in which the hands were off the steering. An analysis of all take-overs through a full factorial ANOVA with factor phase (tutorial, experiment), situation type (yellow, red, lane twitch), hands on steering (true, false), and daytime (day, night) was conducted for TTFR and TTHoS. The results of both ANOVA are presented in Table 2. An overview of the quickest and slowest reaction times is also given in Table 1. For some considered take-over situations of situation type *yellow* and *red*, TTEoR and TTFR of 0 s were recorded. This displays take-over situations in which the driver had their eyes on the road and was not engaging in the secondary task. For the *yellow* situation types this TTEoR was 0 s in 33 take-overs, while in the *red* situation types this only occurred once, displaying anticipation of a take-over situation. Even the maximal outliers display very quick reaction times for all take-over situations. Notably, the maximum TTEoR are slower than TTFR and TTHoS. This can be attributed to drivers focusing on the instrument cluster with the hands already being placed on the steering. The analysis of both the TTFR and TTHoS portrays that there is a significant effect of drive phase in

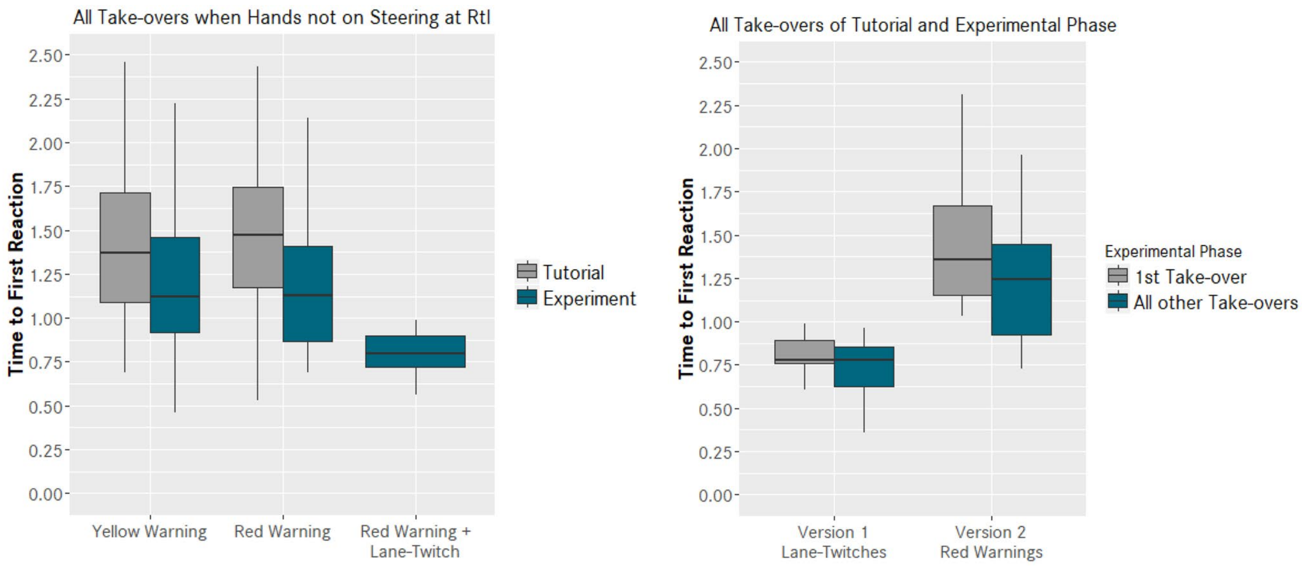

Fig. 4 Left: time to first reaction (TTFR) in the tutorial and experimental phase when hands were not guiding the steering wheel for each different warning type (i.e., yellow, red and red, and lane-twitch). Right: time to first reaction (TTFR) for the two different versions with respect to the first take-over compared to all other take-overs of the same situation type (color figure online)

Table 1 Time to first reaction, time to hands on steering and eyes on road in experiment phase

Situation type	# of take-overs	Time to first reaction (s)			Time to hands on steering (s)			Time to eyes on road (s)		
		Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
Yellow	67	0	2.31	M = 1.222 (SD = 0.403)	0.06	1.65	M = 0.705 (SD = 0.300)	0	2.28	M = 0.324 (SD = 0.449)
Red	53	0	2.31	M = 1.109 (SD = 0.469)	0.36	1.56	M = 0.772 (SD = 0.244)	0	2.01	M = 0.496 (SD = 0.469)
Lane Twitch	42	0.36	1.58	M = 0.841 (SD = 0.233)	0.36	1.56	M = 0.705 (SD = 0.173)	0	0.78	M = 0.262 (SD = 0.267)

Minimum, maximum, and average times are reported

Table 2 Results of full-factorial ANOVA for TTFR and TTHoS

Dependent variable	Measure	Sum of squares	*df*	Mean square	*F*	Pr (> F)	Partial η^2
Time to first reaction	***Phase***	***12.59***	***1***	***12.59***	***43.43***	***1.69e-10***	***0.114167***
	Expectancy	0.93	1	0.93	3.203	0.07441	0.009414
	Criticality	***2.60***	***1***	***2.60***	***8.974***	***0.00294***	***0.025939***
	Hands on steering	***34.91***	***1***	***34.91***	***120.5***	**< 2e16**	***0.263357***
	Daytime	0.00	1	0.00	0.016	0.90028	0.000047
	Phase * expectancy	0.08	1	0.08	0.280	0.59711	0.000830
	Phase * hands on steering	***3.15***	***1***	***3.15***	***10.87***	***0.00108***	***0.031260***
	Expectancy * hands on steering	***2.16***	***1***	***2.16***	***7.441***	***0.00671***	***0.021603***
	Criticality * hands on steering	0.03	1	0.03	0.106	0.74553	0.000313
	Phase * daytime	0.29	1	0.29	0.998	0.31843	0.002954
	Expectancy * Daytime	0.22	1	0.22	0.752	0.386601	0.002252
	Criticality * Daytime	0.00	1	0.00	0.000	0.98561	9.67e-07
	Hands on steering * daytime	***3.84***	***1***	***3.84***	***13.25***	***0.00032***	***0.037824***
	Phase * expectancy * hands on steering	0.04	1	0.04	0.135	0.71391	0.000399
	Phase * expectancy * daytime	0.00	1	0.00	0.006	0.93690	0.000019
	Phase * hands on steering * daytime	0.36	1	0.36	1.243	0.26563	0.003676
	Expectancy * hands on steering * daytime	0.21	1	0.21	0.716	0.39822	0.002119
	Phase * expectancy * hands on S.* daytime	0.00	1	0.00	0.000	0.99852	1.02e-08
Time to hands on steering	***Phase***	***1.288***	***1***	***1.288***	***19.985***	***1.07e-05***	***0.055982***
	Expectancy	***4.509***	***1***	***4.509***	***69.968***	***1.62e-15***	***0.171925***
	Criticality	0.013	1	0.013	0.204	0.6517	0.000605
	Hands on steering	***27.29***	***1***	***27.29***	***423.51***	**< 2e-16**	***0.556880***
	Daytime	0.339	1	0.339	5.260	0.0224	0.015368
	Phase * expectancy	0.015	1	0.015	0.226	0.6347	0.000671
	Phase * hands on steering	0.156	1	0.156	2.414	0.1212	0.007113
	Expectancy * hands on steering	0.028	1	0.019	0.295	0.5132	0.001270
	Criticality * hands on steering	0.010	1	0.010	0.162	0.6877	0.000480
	Phase * daytime	0.003	1	0.003	0.040	0.8418	0.000118
	Expectancy * daytime	0.009	1	0.009	0.143	0.7052	0.000425
	Criticality * daytime	0.000	1	0.000	0.005	0.9463	0.000014
	Hands on steering * daytime	0.062	1	0.062	0.956	0.3289	0.002828
	Phase * expectancy * hands on steering	0.000	1	0.000	0.000	0.9988	7.04e-09
	Phase * expectancy * daytime	0.028	1	0.028	0.439	0.5082	0.001300
	Phase * hands on steering * daytime	0.001	1	0.001	0.012	0.9144	0.000034
	Expectancy * hands on steering * daytime	0.001	1	0.001	0.011	0.9153	0.000034
	Phase * expectancy * hands on S.* Daytime	0.004	1	0.001	0.056	0.8137	0.000165

the data, this is to be expected as drivers were instructed and novices in the tutorial phase. If the hands of the driver were on the steering, this resulted in a highly significant main effect for both reaction times. A small effect is registered for the factor criticality (yellow/red vs. lane twitch), while the expectancy (yellow vs. red) yielded no effect for TTFR. Three further mild interaction effects were identified for TTFR. These results lead to a rejection of Hypothesis #1, as no clear indication of expectancy and criticality is identified for TTFR. Contrarily, a highly significant main effect for expectancy (yellow vs. red warning) was identified for TTHoS, while the data did not display any significant differences for criticality, see Table 2.

As participants drove one of two versions, see Sect. 2.3, a version comparison was conducted for the first take-overs in comparison to all other take-overs in the experimental phase. The comparison aims to identify whether the novel situation lane-twich (M = 0.807, SD = 0.118) caused prolonged TTFR compared to all other lane-twitches (M = 0.725, SD = 0.340), see Fig. 4. As the red warning was learned in the tutorial phase participants were familiar with this situation type. Due to unequal sample size homoscedasticity was tested through a levene test with factors situation type and the binary factor first take-over or all others for the dependent variable TTFR. As homoscedasticity was not achieved $F(3,65)=2.63$, $p=0.058$, a Welch test was performed displaying significance only for situation type $F(1,65)=17.617$, $p< <0.001$. The results display no difference with respect to first appearance of take-over situations, showing that the highly critical lane twitch situation was attended unconditional of novelty. With respect to the *red* situation type, non-significant results strengthen the assumption that learning of take-over procedure in the tutorial phase was effective.

Similarly, on the last straight of the experimental phase (Straight #9), drivers experienced a different expected warning, a red take-over warning was displayed instead of a yellow warning. Figure 5 (left) depicts the TTFR for all yellow warnings on Straight #8 compared to the novel red warning at the end of Straight #9. A two-way ANOVA for TTFR was calculated with the categorical variable whether hands were touching the steering or not and the situation type. This resulted in a significant effect for hands on steering $F(1,36)=18.67$, $p=0.0002$, partial-$\eta^2=0.329$, but not for situation type $F(1,36)=0.349$, $p=0.558$, partial-$\eta^2=0.0091$.

Also in Fig. 5 (right), TTFR is presented for the three different situation types and whether participants engaged in the NDRT within 2 s prior to RtI. Engagement was determined through eye-tracking data and if the tablet was touched within this timeframe. A one-way ANOVA for NDRT engagement was calculated for TTFR with no significant effect, $F(1,162)=0.004$, $p=0.948$, partial-$\eta^2=0.00002$.

4 Discussion

Multiple research questions are addressed in the present study concerning take-over behavior of professional truck drivers in Level 3 automation. Test track drives, simulating highway scenarios, with non-affiliated professional truck drivers were investigated to identify reactions towards different warning types and criticalities resulting in take-over situations. The information gathered is paramount to design future driver assistance systems, enhance usefulness (DIN EN ISO 9241-210:2010 2010) and apply a human-centered design approach (Cooley 1982). A positive byproduct of applying this knowledge to the design process of future systems is to achieve the goal of reducing workload on drivers and therein decrease accidents.

4.1 Take-over criticality

One of the prime focuses of this study was to identify driver behavior for varying critical take-over situations in a real vehicle. The statistical analysis of TTFR displayed a mildly significant effect of the two levels of take-over criticality, while TTHoS was constant for this factor. Criticality was manipulated by inducing a steering impulse in the high condition, causing the steering wheel to swerve to the right when drivers were occupied with a NDRT and hands were not on the steering wheel. However, no visual obstacles caused the variation in criticality to bypass any risk of causing an accident for the participants. Driver interaction depending on criticality seems to cause a quicker first reaction at the controls of the vehicle, see Fig. 4. However,

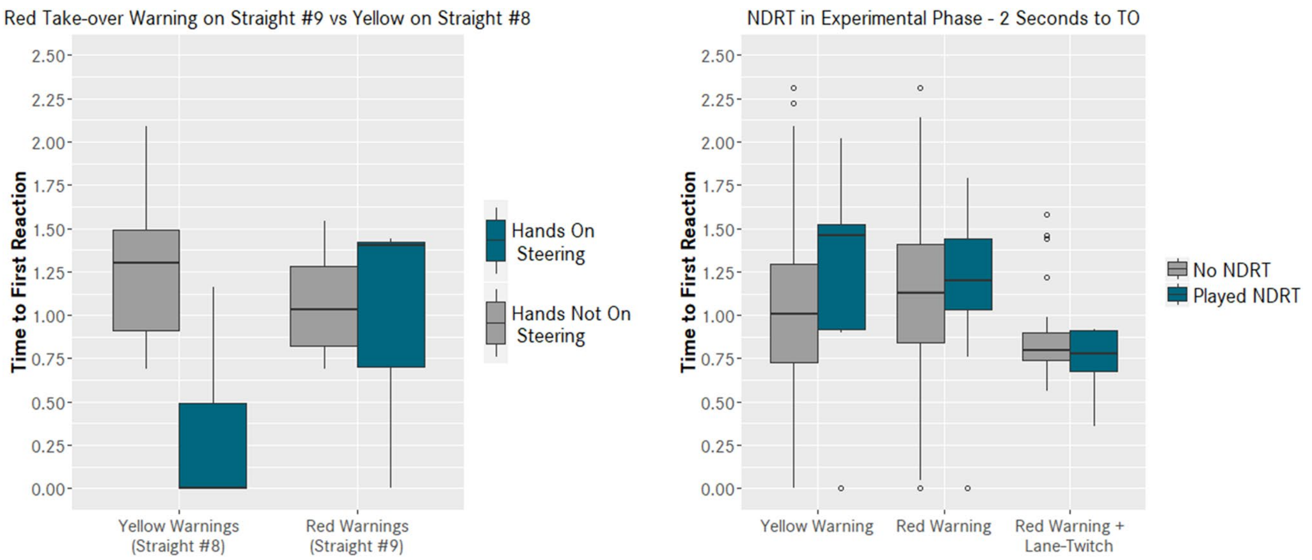

Fig. 5 Left: time to first reaction (TTFR) for all expected yellow warning and expected red warning on Straight #9. Right: comparison of TTFR for all warning types and whether NDRT was played or not within 2 s prior to a take-over request (color figure online)

this does not signify that route planning for the handling of the situation was quicker. The *lane twitch* situation in this study caused an acceleration to be perceived by the driver. Instinctive reaction to the kinaesthesia could have caused the quicker time to first reaction. Hypothesis 1 is therefore failed to be rejected, with necessary future investigations focusing on the causation of quicker reactions.

4.2 Warning expectancy

Two different warning types were presented in the three different take-over situations. The two take-over warnings consisted of a *yellow* take control warning issued solely at the end of a straight and a *red* take control warning issued when system limits were met and drivers had to regain control immediately. These two take-over situations did not differ in criticality and were issued on the straights. In these situations, the system did not appear to mishandle the driving task; however, take-over was required (yellow and red). A yellow take-over warning was always presented at the end of a straight, allowing each participant to foresee (expect) this combination of situation type and warning when approaching the steep curve. Solely on the last straight, the yellow warning was replaced with a red warning.

The comparison of the red and yellow warnings displayed no significant difference for TTFR. Interestingly, a highly significant effect for expectancy was identified for TTHoS. As the yellow warning could be expected at the end of the straights, take-over procedure could be anticipated at the end of the straights. This is also present in the data, as at roughly 50% of all *yellow* take-over situations, the eyes were already focused on the road, demonstrating expectancy. The overall tendency of expectancy to have no effect on TTFR was also identified for the first and last take-overs in the experiment phase in which new and switched warnings were presented, respectively. The discrepancy between TTFR and TTHoS regarding expectancy can have multiple reasons. However, primarily only placing the hands at the steering caused the warning to terminate. The effect of expectancy on TTHoS is, therefore, anticipated. As a clear driver interaction based on obverting an obstacle at take-over was not necessary, the threshold for TTFR could have been set too high for the reaction to register similar to TTHoS.

An analysis of the final take-over warning at the end of Straight #9, which was switched from an expected yellow warning to a red warning, produced no significant effect for TTFR. However, reaction times did differ significantly in the yellow warning situation when hands were guiding the steering. A probable explanation is that some participants expected yellow warnings, while others trusted the automation function to warn them adequately possibly also verifying the warning in the instrument cluster. With the appearance of a different auditory–visual, warning was expected; however, the criticality was not and led to a reevaluation of the situation. The motoric response to unforeseeable take-overs did not differ throughout the drives and averaged at 0.727 s. A similar conclusion was drawn when explicitly investigating the hands-off times in partially automated traffic jam scenarios with critical take-over situations (Naujoks et al. 2015). Therefore, based on the information gathered, Hypothesis #2 is failed to be rejected.

4.3 Comparability to simulator studies

Another research focus is to compare previously gathered information in a moving-base simulator to driver behavior in a real vehicle. Generally, current investigations in the Level 3 context are examined exclusively in simulators, to exclude the possibility of technical malfunctions or human error causing accidents. This is mainly because prototypes often are not as sophisticated as series technology and novel or extreme situations are investigated. While the possibilities in simulated environments have excelled in recent years, unexpected influences might take an effect on results when trying to extrapolate results to a real-world context. Such influences could include the intrinsic trust of participants that no direct physical harm will occur in a simulator and the aspect of treating driving simulators as games (Bellem et al. 2016). Overall, results obtained in simulators have to be validated in controlled and safe real-world scenarios to manifest conclusions drawn from the data. As previous experimental results in a truck simulator produced extremely quick TTFR ($M = 1.35$, $SD = 0.49$) for 755 highly time-critical take-overs (Lotz et al. 2019b), it was our motivation to validate these findings.

Results in Sect. 3 show that especially the motoric response time, TTHoS, was constant regardless the expected criticality. Learning was observed from the tutorial to the experimental phase, see Fig. 4. The TTFR reduced significantly between the tutorial and experimental phase generating consistent results of learning behavior as observed previously (Lotz et al. 2019b). These near-constant motoric response times observed in the data cohere with the findings from Zeeb et al. (2015). Second, the different warning types yielded a significant effect for TTHoS in the experimental phase. This contradicts the results for TTFR, as only the most critical take-over situation generated a mildly significant decreased reaction time. Overall, a TTFR of 1.08 s was observed for all take-over situations in which the hands were not guiding the steering wheel. A separate investigation of TTFR for NDRT engagement yielded no significant results.

For the consideration of reaction times in the present study the results of the descriptive and inferential analysis fail to reject Hypothesis 3 with a 3σ-interval of 0–2.343 s for 164 take-overs, depicting 99.7% of the data. One drawback and possible explanation of the difference in TTFR could be the low torque required to take-over the vehicle in the yellow and red warning conditions. Only in the lane-twitch condition was a strong steering wheel torque required. It should be noted that the present study did not present highly time critical situations due to environmental obstacles, but the steering was manipulated in the lane-twitch condition to induce criticality as a safe option to investigate take-over criticality.

4.4 Non-driving-related tasks (NDRT)

Multiple publications investigation the effect of NDRT during Level 3 have been published recently (Merat et al. 2012; Petermann-Stock et al. 2013). In the present study participants could engage in a NDRT during Level 3 automation. As a measure of engagement the visual attention towards the NDRT and tablet interaction were consulted, see Fig. 5. The data reported shows no significant effect of NDRT on TTFR. This is unexpected and leads to a rejection of Hypothesis 4. While the results in published research show no significant effects amongst different NDRT (Radlmayr et al. 2014), a general negative effect of tasks is documented (Vogelpohl et al. 2016). In the present study, participants could choose freely if they wanted to attend the NDRT or not. It is possible that the self-motivated engagement was, therefore, lower than in studies in which direct task instruction was provided.

5 Conclusion

With little experience to lean on, as no studies addressing a standardized procedure of investigating take-over behavior in critical take-over situations exists, this study presents novel insights into the experimental procedure and results of near real-world take-over behavior for Level 3. A total of 292 take-overs in which hands were not guiding the steering wheel and 64 with hands on steering were examined. The information gathered allows for a better understanding of reaction times and behavior during take-overs to expected and unexpected events. Ultimately, this is valuable knowledge to reduce accidents in critical take-over situations in which the machine and humans share responsibility of guiding vehicles safely. For optimal data comparability, we would welcome a definition of take-over situations and measures throughout the community.

Compared to previous results, the reaction times are in the lower spectrum of reported reaction times (Eriksson et al. 2017). As argued previously, professional truck drivers form a unique sample as experience and training are high. It should be clear, however, that no driver had previously driven on this test track, the automation function and vehicle was uncommon and drivers were inclined to act safely. The time to first reaction (M = 1.087 s, SD = 0.417 s), time to hands on steering (M = 0.727 s, SD = 0.253 s) and time to eyes on road (M = 0.364 s, SD = 0.425 s) are most likely as quick as possible for the presented scenarios.

As no obstacles were manipulated in the environment of the vehicles at take-over, the reaction times only display the cognitive and motoric response times required to react towards a take-over signal. Resulting trajectories or displacements within the lane were not investigated. Primarily due to short distances on the test track straights and as participants reactivated the automation function within 10 s, an analysis of take-over quality is absent. In conclusion, the collected reaction times present the quicker spectrum of take-over capability by professional truck drivers. While the limit of take-over reaction times cannot be much quicker, slower reaction times due to long automation phases, higher distraction and repetition are very likely and need thorough investigation in the future. The reaction times here present the minimum time an automation function needs to span to assist driver take-over behavior. This still disregards if the driver feels comfortable with these take-over times and is capable of delivering this high performance over longer periods of time.

Acknowledgments Open Access funding provided by Projekt DEAL.

Compliance with ethical standards

Conflict of interest Alexander Lotz and Enrico Wohlfarth are employed by Daimler AG.

References

ATP Automotive Testing Papenburg GmbH (2018) Automotive Testing Papenburg. Citation: 20.02.2019 https://atp-papenburg.de

Audi (2017) Audi Technology Portal. Citation: 26.02.2019 from Audi A8 - Audi AI traffic jam pilot: https://www.audi-technology-portal.de/en/electrics-electronics/driver-assistant-systems/audi-a8-audi-ai-traffic-jam-pilot

Axelson M, Ahlström C, Krupenia S, Anund A, Leeuwen W, Kecklund G (2018) Is it possible to adjust the driving and resting times when operating highly autonomous trucks? J Sleep Res 27(1):P345

Bainbridge L (1983) Ironies of automation. Automatica 6:775–779

Bellem H, Kluver M, Schrauf M, Schoner H-P, Hecht H, Krems JF (2016) Can we study autonomous driving comfort in moving-base driving simulators? A validation study. Hum Factors 59(3):442–456
Brookhuis K, de Waard D, Janssen W (2001) Behavioural impacts of advanced driver assistance systems - an overview. Eur J Transp Infrastruct Res 1(3):245–253
Cooley M (1982) Achitect or Bee?: Human Price of Technology (Current affairs). Chatto & Windus, UK
Daimler (2018) Mercedes-Benz. Citation: 26.08.2019 from The new Actros 2019: https://www.mercedes-benz.com/en/mercedes-benz/vehicles/trucks/the-new-actros-2019/
Damböck D (2013) Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme. Technische Universität München, Lehrstuhl für Ergonomie, München
de Winter JC, Happee R, Martens MH, Stanton NA (2014) Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. Transp Res Part F 27:196–217
Diederichs F, Bischoff S, Reilhac P (2015) Welchen Einfluss hat das HMI?. Vergleich von Übergabezeiten in Notfallsituationen bei SAE Level 2 und 3 Automatisierung mit neuartiger Smartphone-Integration im direkten Fahrersichtfeld und mit Lenkradbedienung. Berliner Werkstatt Mensch-Maschine-Systeme: Trends in Neuroergonomics 11:72–79
DIN EN ISO 9241-210:2010 (2010) Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems
Endsley MR (1995) Toward a Theory of situation awareness in dynamic systems. Hum Factors J 37(1):32–64
Eriksson A, Stanton NA (2017) Take-over time in highly automated vehicles: non-critical transitions to and from manual control. Hum Factors. https://doi.org/10.1177/0018720816685832
Gold C, Damböck D, Lorenz L, Bengler K (2013) Take over! How long does it take to get the driver back into the loop? Proc Hum Factors Ergono Soc Annu Meet I 57:1938–1942
Hergeth S, Lorenz L, Krems JF (2016) Prior Familiarization with takeover requests affects drivers' takeover performance and automation trust. Hum Factors J Hum Factors Ergonom Soc 59(3):457–470
Kantowitz BH, Roediger HL, Elmes DG (2009) Research techniques: experiments. Wadsworth Cengage Learning, Belmont
Lees MN, Lee JD (2007) The influence of distraction and driving context on driver response to imperfect collision warning systems. Ergonomics 8:1264–1286
Lotz A, Weissenberger S (2019) Predicting take-over times of truck drivers in conditional autonomous driving. In: Stanton N (ed) Advances in human aspects of transportation. AHFE 2018. Advances in intelligent systems and computing, vol 786. Springer, Cham, pp 329–338
Lotz A, Russwinkel N, Wohlfarth E (2019) Response times and gaze behavior of truck drivers in time critical conditional automated driving take-overs. Transp Res Part F 64:532–551. https://doi.org/10.1016/j.trf.2019.06.008
Marberger C, Mielenz H, Naujoks F, Radlmayr J, Bengler K, Wandtner B (2018) Understanding and applying the concepf of "driver availability" in automated driving. In: Stantion NA (ed) Advances in human aspects of transportation. AHFE 2017. Advances in intelligent systems and computing, vol 597. Springer, Cham
McDonald AD, Alambeigi H, Engström J, Markkula G, Vogelpohl T, Dunne J, Yuma N (2019) Towards computational simulations of behavior during automated driving take-overs: a review of the empirical and modeling literatures. Hum Factors 61:642–688
Melcher V, Rauh S, Diederichs F, Widlroither H, Bauer W (2015) Take-over requests for automated driving. 6th international conference on applied human factors and ergonomics. Procedia Manuf 3:2867–2873
Merat N, Jamson H, Lai F, Carsten O (2012) Highly automated driving, secondary task performance, and driver state. Hum Factors J Hum Factors Ergonom Soc 54(4):762–771
Naujoks F, Purucker C, Neukum A, Wolter S, Steiger R (2015) Controllability of partially automated driving functions–does it matter whether drivers are allowed to take their hands off the steering wheel? Transp Res Part F Traffic Psychol Behav 35:185–198
Petermann-Stock I, Hackenberg L, Muhr T, Mergl C (2013) Wie lange braucht der Fahrer–eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. 6. Tagung Fahrerassistenzsysteme. Der Weg zum automatischen Fahren
Radlmayr J, Gold C, Lorenz L, Farid M, Bengler K (2014) How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. Proc Hum Factors Ergonom Soc Ann Meet 58:2063–2067
Ruscio D, Ciceri MR, Biassoni F (2015) How does a collision warning system shape driver's brake response time? The influence of expectancy and automation complacency on real-life emergency braking. Accid Anal Prev 77:72–81
SAE J3016 (2018) from https://www.sae.org/misc/pdfs/automated_driving.pdf
Tesla (2019) Citation: 26.02.2019 from Tesla Autopilot https://www.tesla.com/de_DE/autopilot
Vogelpohl T, Vollrath M, Kühn M, Hummel T, Gehlert T (2016) Übergabe von hochautomatisiertem Fahren zu manueller Steuerung. Gesamtverband der Deutschen Versicherungswirtschaft e.V, Berlin
Wylie G, Allport A (2000) Task switching and the measurement of "switch costs". Psychol Res 3(4):212–233
Young M, Stanton N (2002) Malleable attention resources theory: a new explanation for the effects of mental underload on performance. Hum Factors 3:365–375
Young M, Stanton N (2007) Back to the future: brake reaction times for manual and automated vehicles. Ergonomics 1:46–58
Zeeb K, Buchner A, Schrauf M (2015) What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accid Anal Prev 78:212–221
Zhang B, Wilschut E, Willemsen D, Martens M (2017) Driver response times when resuming manual control from highly automated driving in truck platooning scenarios. Conference: RSS2017 Road Safety & Simulation international conference (The Hague, Netherlands). https://doi.org/10.13140/RG.2.2.28249.01127

5 Paper No. 4

Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E. (2020). An adaptive assistance system for subjective critical driving situation: understanding the relationship between subjective and objective complexity. In: D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference (pp. 97-108).* Nantes, France.
https://www.hfes-europe.org/wp-content/uploads/2019/10/Lotz2019.pdf

An adaptive assistance system for subjective critical driving situations: understanding the relationship between subjective and objective complexity

Alexander Lotz[1], Nele Russwinkel[2], Thomas Wagner[1], & Enrico Wohlfarth[1]
[1]Daimler AG, [2]Technische Universität Berlin
Germany

Abstract

Partial and conditional automated driving allows the driver to transfer responsibility to the vehicle. While assistance systems are designed to deal with aspects of the driving task, currently no assistance systems are available to predict driver behaviour for take-over when the vehicle is handling the driving task. This is important as drivers might interpret driving situations differently than an activated automation function. This can cause self-initiated take-overs leading to a reduction of trust in the system. In theory, if a prediction is robust, an assistance system could also adapt based on this prediction. A new subjective complexity model addressing these situations is introduced. The subjective complexity model learns situations in which individual drivers have previously self-initiated control of the driving task. Based on exemplary sideswipe manoeuvres, the system concept is explained and simulated with a training and test dataset. Upon introducing this system, a discussion is initated on the difference between objective and subjective situation complexity. A distinction is drawn between mathematical descriptions based on vehicular sensor data and human interpretation of the environment. The proposed system also functions as a carrier technology for further investigations between the differences of objective and subjective complexities.

Introduction

The driving task consists of many short-term decisions, e.g. steering to hold vehicle in lane, and long-term decisions, e.g. route navigation. Different factors need to be considered in order for the driver to successfully solve these tasks. Therefore, it is important to understand which aspects are considered complex. This can help to include, enhance or adjust assistance accordingly. At the same time, humans are decisively influenced by the environment, which they encounter and thereby confined in their range of interactions. With various developments in the field of automated driving, possibilities of directing attention away from the driving task will become possible. In Level 3 (SAE J3016, 2018) the driver can focus on non-driving related tasks, but needs to regain control of the vehicle when warned. Recent research regarding take-over behaviour has shown that environmental factors such as traffic density and time-budget (distance to objects) play a crucial role in successful take-over capability for Level 3 (Gold, et al., 2016; Lotz, et al., 2019; Zhang, et al., 2019). There is a large variety of definable environmental factors, making it difficult to

In D. de Waard, A. Toffetti, L. Pietrantoni, T. Franke, J-F. Petiot, C. Dumas, A. Botzer, L. Onnasch, I. Milleville, and F. Mars (2020). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

pinpoint isolated factors to varying driver take-over behaviour. When revisiting traffic density as an exemplary factor, other environmental factors such as time to collision, number of lanes and colour of vehicles can form interaction effects. Multiple isolated environmental factors can merge to form singular driving situations through the relation of several of these factors over time. Arbitrary measures, such as low or high traffic density, also make a comparison difficult. However, the driving environment can be measured with a variety of different sensors mounted on a vehicle. Based on the chosen sensor setup, this creates a representation with a sensor-specific degree of detail of environmental factors or higher-level situations. As a driver also perceives the environment with her/his senses and develops a representation of the situation, influences of environmental factors such as traffic density on the driver can be compared and deduced. If a certain driver reaction is linked to an environmental factor, based on the sensor representation of the environment an assistance system could possibly predict driver behaviour. This information would especially be valuable in the abovementioned take-over situations, in which responsibility shifts from the machine to the human. The mathematical description of the environment could be attributed to subjective complexities, identifying scenarios that cause higher workload and situations in which the driver needs assistance. As there are possibly also inter- and intra-individual differences in perceived subjective complexity, an ideal assistance system would adapt individually and specific to different driving situations. This could lead to a better usability, correct allocation of assistance and acceptance of automated driving function.

Sensors such as cameras, radars and lidars collect data that describe an abstraction of their perceived environment and allow interpretation either through humans or computational algorithms. In a simplistic form, this data collection is similar to the cognitive processing for the first perception phase towards building situation awareness (Endsley, 1995). In what terms does environmental complexity differ mathematically (objective complexity) to an individual perceived situation complexity (subjective complexity) and how can this be measured? The second part of this question will be addressed in this paper and a solution will be developed to enable the investigation of the first part of the question in future work.

It is worth defining our interpretation of these two different versions of complexity, explicitly regarding driving environments. *Objective complexity* is the mathematical describable driving situation in which all objects within a predefined area are continuously referenced to an ego-vehicle. This mathematical description includes metrics such as the distances, velocities (relative and absolute) and the time to trajectory intersections. The mathematical composure of the factors can vary and yield different values of objective complexity depending on the interpretation of the mathematical description. In a practical example, the data would be obtained from singular or combinations of sensors, capturing information of environmental objects. This differs from general global descriptions of complexity such as the number of vehicles in the environment (Gold, et al., 2016), in which no references to an ego-vehicle and driver are drawn. The problem with global descriptions, without reference to the driver in the environment, is that the dispersion of vehicles is not evident from the point of view of the driver. When listing the amount of vehicles surrounding the ego-vehicle, no information is given where all these vehicles are (front, behind, lane etc.). *Subjective complexity* is the perceived complexity of a driving situation from the human's perspective. This includes all stationary and moving objects relevant to the driving task. Abstract cognitive and psychological constructs such as driving situation familiarityaffect this complexity and are not measureable with similar accuracy as the metrics of objective complexity. This is mainly because measurements from designated sensors such as electroencephalography, skin conductance or any other psychophysiological measurement are not unambiguously linked to any of these constructs and quantification of human response is not possible. A scale for the subjective complexity is also arbitrary, relative to the psychological constructs and subject to individual differences.

Previous research has focused on describing environmental factors specifically for the driving environment, such as the time to resume control and the quality of the transition depend on driver-vehicle-environment factors (Gold, et al., 2016). Early work resulted in a classification scheme of driving situations with three million unique situations (von Benda, 1977). This classification scheme was later simplified to incorporate only four major aspects; horizontal course, traffic density, special weather and hazards (Fastenmeier, 1995). Due to the high complexity of factors, different types of models have been introduced to predict driver transition behaviour. The first class of models utilizes mathematical models, e.g. regression models, to extrapolate data based on empirical findings post-hoc and explain correlations in the data (McDonald, et al., 2019; Zhang, et al., 2019). A second class of models provides online prediction based on data obtained through driver and environment monitoring (Nilsson et al., 2015; Braunagel et al., 2017; Lotz & Weissenberger, 2019). However, subjective driver interpretation is missing as input data. The problem with all of these models is that defined factors can interact, e.g. traffic density or driver experience, effects that cannot be investigated in isolation within one study. Therefore, the investigation of differences between objective and subjective complexity, as defined above, has been difficult in the past. The investigation was especially difficult as drivers continuously needed to control the vehicle, always generating a response at steering. This has now changed through automated driving.

Subjective relevance is an important factor to predict individual behaviour. Ohn-Bar and Trivedi (2016) conducted research on the subjective relevance of objects in the driving environment, stating that spatio-temporal reasoning is needed to identify relevance by the driver. Therefore, the context of space and time in the driving situation of any automation level should be regarded when investigating environmental effects on the driver.

Through recent technical advances of automated driving, it is possible for the driver to take their hands off the steering wheel and observe the environment. Automated driving, specifically Level 2 and Level 3, is an ideal enabling technology suited for the investigation of differences between objective and subjective complexity. Therefore, it is possible to gather data on subjectively perceived critical complexity where previously the driver continuously generated responses at the steering and the data were open for interpretation. A distinction of intended interventions was difficult, because drivers constantly had their hands on the steering wheel.

This paper introduces a conceptual advanced driver assistance system. The system is designed to learn situations in which the driver takes back control of the self-driving vehicle, when no request to intervene is issued. The assistance system thereby registers situations based on current objective complexity from the vehicular sensors and associates it with subjective complexity. The moment drivers reclaim control through self-initiated take-over, the objective complexities gathered by vehicle sensors can be identified in which no automated driving is desired. Thereby, the assumption is formed that the drivers consider the environment as subjectively complex. Hence, the automated vehicle can learn when the automation function itself can suggest take-over predictively.

Subjective Complexity Model

The proposed subjective complexity model relies on the fact that the vehicle has a driver assistance system capable of simultaneous lateral and longitudinal control, i.e. without the need of having the hands on the steering wheel. Typically, this approach is only possible with advanced Level 2 or Level 3 systems. The objective of the proposed model is to make predictions when the surrounding driving complexity reaches a point in which the driver feels intervention is necessary. Thereby, a relationship between objective and subjective complexity can be investigated. The hypothesis is followed that the driver subjectively decides that complexity of the driving environment is too complex and external vehicular sensory data is recorded at intervention. Other reasons for self-initiated take-over are also possible, e.g. low satisfaction with vehicle control, and intention cannot be differentiated. It is worth noting, that the trust in the automated vehicle is affected by driving experience (Gold, et al., 2015). To show the functionality of the model, sideswipe manoeuvres were recorded with an advanced Level 2 automated truck and divided into a training and test dataset. These sideswipe manoeuvres were limited to vehicles crossing onto the ego-lane from the fast lane (left).

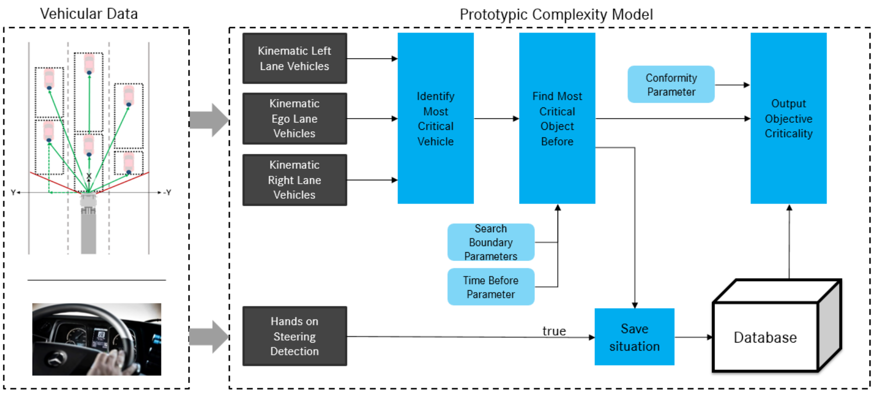

Figure 1. Workflow of adaptive assistance system. Sensor data are split into regions of interest (left, ego, right). Kinematics of all perceived objects in these regions are calculated upon which the most critical objects is identified (see Figure 2). Output criticality is calculated based on previously recorded data.

Concept

The functionality of the subjective complexity model is presented in Figure 1. The model is split into four major components. First, the perception of vehicles in the periphery of an ego-vehicle are identified, including the calculation of kinematic relationships. Secondly, the most critical object is determined based on the previously calculated kinematics. Thirdly, situations in which the driver intervenes with the vehicle without a take-over warning being displayed, are recorded and saved in a database. Fourthly, the criticality of current driving situations is calculated based on the conformity parameter of current kinematics with saved situation kinematics.

Sensory perception and kinematics

Sensory data of the surrounding vehicles are gathered and split into three possible lane positions. This includes the ego-lane as well as the lanes directly to the left and right. The raw data received from the sensors includes the lateral $Dist_y$ and longitudinal $Dist_x$ position as well as the speed of each object relative to the ego vehicle $RelSpd$ and object width $Width_y$. This allows the calculation of lateral Spd_y and longitudinal Spd_x speed of each object. Additionally, a safety corridor is defined through the width of the variable $Buffer$, see Figure 2. In this version of the proposed model, a maximum of six vehicles could be perceived around the ego-vehicle, with a maximum of two objects per lane. It should be noted that different sensor setups can alter the outcome of the system dramatically. By adding different sensors, e.g. cameras for object classification, additional data can offer subsequent critical object identification. Based on available radar data with the current sensor setup, the following kinematic variables were calculated.

$$TT_{cross\_border} = \left(\frac{|Dist_y| - \left(\frac{1}{2}\left(Width_y\right) + Buffer\right)}{Spd_y} \right) \quad (1)$$

$$TT_{headway} = \frac{Dist_x}{RelSpd} \quad (2)$$

$$TT_{collision} = TT_{headway} - TT_{cross_{border}} \quad (3)$$

$$Dist_{cross\_border} = TT_{cross\_border} * RelSpd \quad (4)$$

In total, ten kinematic variables are taken into account with the available sensors, see Table 1. Every relevant object on any of the three lanes has a separate set of these ten variables. Further variables in following implementation versions could include crossing angles, further crossing times, trajectory predictions.

Identification of most critical object

In the case of a self-initiated driver take-over, i.e. no request to intervene, either the complete constellation of the surrounding vehicles or a single object causing the driver to intervene needs to be identified. Here an assumption needs to be formulated, to differentiate between these two options. The proposed model assumes that one object is the most critical in the environment and it can be defined as the object that would enter the safety corridor first, if all vehicles maintain their trajectory. This assumption corresponds to the smallest $TT_{collision}$, see equation (3), of any of the six surrounding objects. Previous development versions of the adaptive model also incorporated multiple critical objects. However, as there is always one object which is hit prior to all the others, the assumption was made that the driver reacts primarily towards this object. If the constellation of all vehicles were to be recorded, a higher amount of constallations would be possible with less likelihood of reoccuring.

Recording self-initiated take-over situations

If a driver intervenes with the automation function controlling the ego-vehicle, the currently most critical object is recorded to a database. Simultaneously, the model identifies where this most critical object was located for a certain amount of time previously to the take-over. The time is an adjustable parameter as well as the size of the search region, defined by a lateral and longitudinal measure. The two sets of ten kinematic variables, current and delayed, are saved with object lane positions resulting in 22 mathematical variables. Every time the driver regains control of the vehicle, the current situation with its delayed prior position is recorded to the database. As the driving environment can vary dramatically based on the type of road or national restrictions, the data and type of driving culture are completely adaptable. Similarly, the driver's interpretation of situations may vary over time and compared to other drivers. As more and more data is recorded the model adapts over time, this enables learning of personalized self-initiated take-over.

Table 1. Kinematic variables calculated from sensor data.

Variable Name	Definition
$\boldsymbol{Dist_x}$	Longitudinal distance from front of ego-vehicle to rear of object.
$\boldsymbol{Dist_y}$	Lateral distance from front of ego-vehicle to rear of object.
$\boldsymbol{Spd_x}$	Longitudinal speed of object relative to the longitudinal axis of the ego-vehicle.
$\boldsymbol{Spd_y}$	Lateral speed of object relative to the lateral axis of the ego-vehicle.
$\boldsymbol{EgoSpd_x}$	Speed of the ego-vehicle along its longitudinal axis.
$\boldsymbol{RelSpd}$	Difference of longitudinal speed between the object and the ego-vehicle
$\boldsymbol{TT_{cross\_border}}$	The time required for the object to cross into the safety corridor. Only considered if the trajectories of the vehicles cross.
$\boldsymbol{TT_{headway}}$	The time required for the ego-vehicle to bridge the longitudinal distance to the object.
$\boldsymbol{TT_{collision}}$	The time required for the ego-vehicle to reach the point where the object crosses into the safety corridor minus the time required to reach that point. This measure considers the time to collision once the safety border is breached.
$\boldsymbol{Dist_{cross\_border}}$	The longitudinal distance the object is from the ego-vehicle once the safety corridor is breached.

Continuous Criticality Output

Upon identifying a most critical object, the model relies on fuzzy logic (Ross, 2010) to compare current situations with previous unforced take-over situations from the abovementioned database. All kinematic variables are taken into account for the prediction method and a majority voting mechanism determines comparability of saved situations with the current driving environment. It is possible to adapt to this mechanism in the future. The model searches through all previous situations, comparing current kinematic variables to the saved situations. As it is highly unlikely that the exact situation appears twice during an self-initiated take-over, a confidence percentage in form of a conformity parameter is introduced. The definition of this confidence percentage has a profound influence on the precision of the model as discussed in the conclusion.

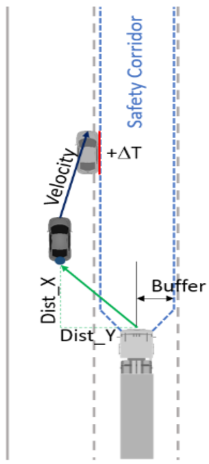

Figure 2. Overview of distances and variables for kinematic calculations. The vehicle on the left lane requires ΔT time, corresponding to TT_{cross_border}, to cross into the safety corridor on the ego-lane in front of the ego-vehicle.

Data Collection

To present the functionality of the subjective complexity model, a small number of manoeuvres were recorded on a German two-lane federal road with speed restrictions. This dataset is too small to investigate the full potential of the system. However, first indications of the functionality can be examined. The sensors were mounted to a prototype Mercedes-Benz Actros with an Active Drive Assist (Daimler AG, 2019). Over the course of two hours, sideswipe manoeuvres from the left lane towards the ego-lane were recorded. This manoeuvre was an exemplary situation, which our fictive driver was uncomfortable in and chose to take-over. It can be expected that real-world traffic situations in which a driver intervenes with the automation function are seldom and would not deliver adequate data. The dataset was divided into a training and test dataset with proportions of approximately 90% to 10% respectively.

This resulted in a total of 105 training sideswipe manoeuvres, see Figure 3, and 13 test manoeuvres.

Results

The model is evaluated based on the self-initiated test manoeuvres that are examined qualitatively. These 13 test manoeuvres are not limited to sideswipes, they consist of take-overs due to a construction site, one sideswipe in a traffic jam at low speeds, five delayed take-overs due to sideswipes and six sideswipes from motorway entry-ramps, i.e. right side. A qualitatively comparison of vehicular signals synchronised with a dashcam video was realized for the model proof of concept. A quantitative analysis was not meaningful, as the data are limited. The complete model was simulated in MATLAB/Simulink, see Figure 4 and Figure 5, which depict the prediction value of the most critical object currently and delayed as well as the hands-on signal when the driver intervened.

Figure 3. Exemplary sideswipe manoeuvres recorded in the training dataset.

Figure 4. Two exemplary self-initiated take-over situations that were not trained in the training set. Predicted sideswipe manoeuvres never reach a confidence greater 80%.

Figure 5 displays the qualitative comparison of the five sideswipe manoeuvres, which the driver initiated with a varying delay. As shown in the graphs portraying the current similarity prediction, delayed similarity prediction of 0.5 sec and when the take-over was initiated (top to bottom), the snapshot of the actual sideswipe was predicted very accurately (vertical blue line). Overall, in four of the five delayed take-overs, the model correctly identifies a sideswipe manoeuvre with 100% confidence. The third depicted sideswipe take-over in Figure 5 with a delayed response shows a low prediction quality. It can also be seen, that sensor dropout appears quite frequently throughout the drives.

The self-initiated take-overs that were not included in Figure 5 and consisted of the six take-overs from sideswipes at motorway entry-ramps, displayed a poor quality of prediction and are not depicted. Overall take-over prediction value by the model in these other eight situations never reached over 80%. Two of these eight exemplary situations are depicted in Figure 4.

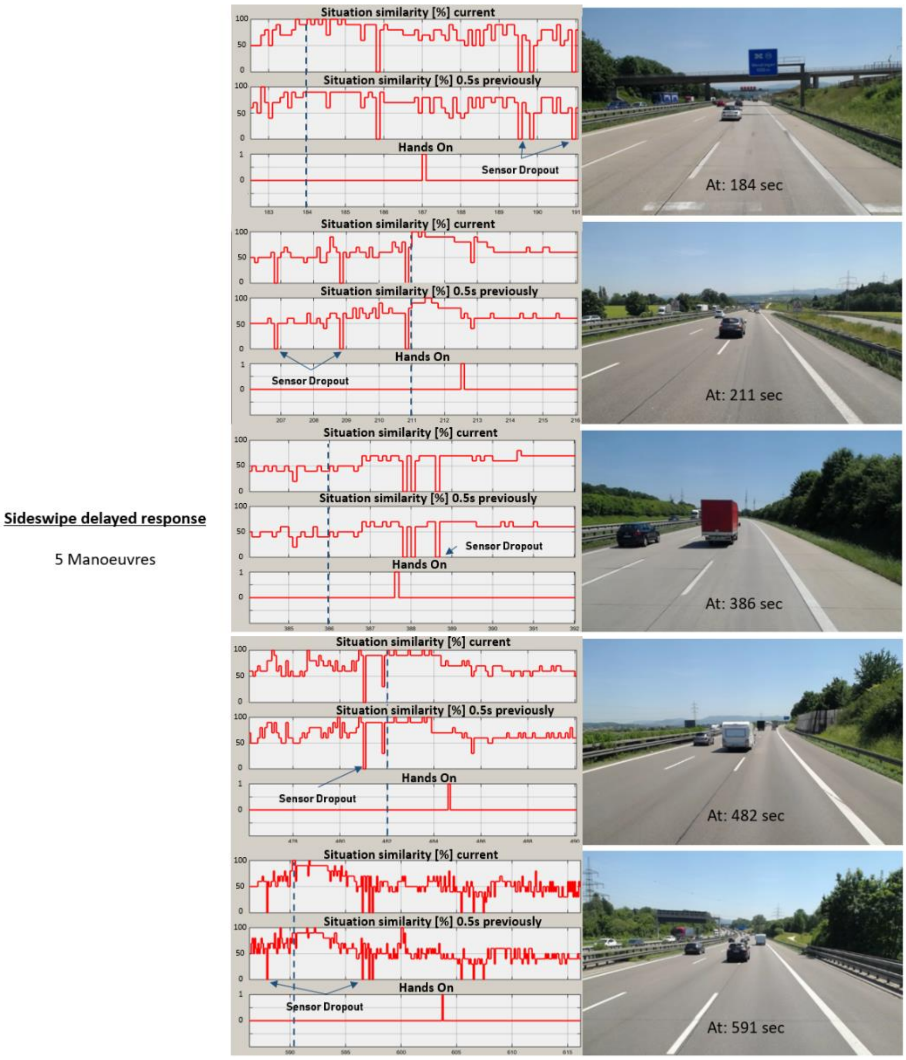

Figure 5. Five sideswipe manoeuvres with delayed driver response. Trace data depicts confidence values for current and delayed most critical objects in the environment. The hands-on signal generated by the driver is also depicted. The blue line indicates the point in time, to which the video-snapshot corresponds.

Conclusion

The introduction of our subjective complexity model is an innovative solution of a learning and adaptive assistance system. By recording driver self-initiated unforced take-overs during automated driving, it is possible to monitor take-overs without interpretation of intent and behaviour. If proven reliable and beneficial, this system can predict preferences in which the driver does not trust the self-driving vehicle or feels the need to manage the driving situation. Trust is an essential component in the human-machine-interaction during automation, as drivers should be able to anticipate system behaviour. If this is not possible, self-initiated take-overs are likely and the model offers assistance. Through continuous learning of relevant situations in which the driver wishes to control the vehicle, the vehicle itself can suggest take-over predictively. This offers different configurations of predictions for different drivers and roads. The functional layout of the model also allows the adjustment of sensors, where the effect on the predictability of take-over can be tested.

Apart from being an adaptive driver assistance system, the model can function as a carrier technology for the investigation of objective and subjective complexity. Thereby, a solution for the second part of our research question is proposed. One of the main obstacles is that sufficient data are difficult to record for this theoretical comparison. On the brink of introducing automated driving to vehicles, previously the driver continuously held control of the vehicle. This made a differentiation difficult between instances, in which the driver considered the environment to be complex. Self-initiated take-overs are valuable for the interpretation of subjective complexity. These situations show that meaning of the temporal and spatial characteristics of surrounding objects from the drivers' perspective was complex enough to motivate a take-over. It should be mentioned that self-initiated take-overs could also occur due to uncritical situations, e.g. terminating automation. In order to truly investigate the differences of subjective and objective complexities, the first part of the research question, a long-term data collection of individual drivers is required that needs documentation of driver intent.

The results of the prediction accuracy of the model shows satisfactory results. While the sideswipe manoeuvres in the test dataset were identified prior to delayed take-over in four of the five instances, one unlearned situation was not identified, see Figure 5. However, there are several reasons and possibilities to improve prediction and the validation of the model. A filtering of the signal is required for subsequent versions to bypass sensor dropout and smooth the prediction value signals. Another shortcoming in our proof of concept are the high number of false positives. It should be investigated whether these false positives occurred, due to the low distinction between a sideswipe manoeuvre being initiated and vehicle continuing in their lane. Furthermore, the point in time in which the driver initiates take-over can vary dramatically, making it difficult for the system to reference the correct critical object to the situation. Reaction times of a driver must possibly be taken into account. Finally, the system can never abstract the data to new situations. Each situation has to have happened similarly in order for the model to predict the situation in the future. However, based on the introduced conformity parameter, see Figure 1, the model can parameterise to achieve different levels of generalisation.

The model shows that this type of assistance system has promising applications in the driving context as well as research. An applied subject complexity assistance would require larger datasets, a higher variance of critical unrequested take-overs as well as seldom occurrences. The model could also be realized with a machine learning approach, however, the presented solution has the added benefit of clearly showing how and why the system functioned with specific predictions. In the future, a large dataset will be utilized as a basis for a parametrization of all variables as well as the expansion of vehicular sensor for further kinematic description of the environment.

References

Braunagel, C., Rosenstiel, W., & Kasneci, E. (2017). Ready for Take-Over? A New Driver Assistance System for an Automated Classification of Driver Take-Over Readiness. *IEEE Intelligent Transportation Systems Magazine*, 9, 10-22.

Daimler AG. (2019, May 02). Daimler Global Media Site. Retrieved July 08, 2019, from The new Actros - all new features in detail: https://media.daimler.com/marsMediaSite/ko/en/43216408

Endsley, M.R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*, 32-64.

Fastenmeier, W. (1995). Autofahrer und Verkehrssituation: neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Strassenverkehrssysteme. Köln, Germany: TÜV Rheinland.

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in Automation - Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle. *Procedia Manufacturing*, *3*, 3025-3032.

Gold, C., Korber, 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,* 642-652.

Lotz, A., & Weissenberger, S. (2019). Predicting take-over times of truck drivers in conditional autonomous driving. In: N. Stanton (Ed.), *Advances in Human Aspects of Transportation. AHFE 2018. Advances in Intelligent Systems and Computing*, *786*, (pp. 329-338).

Lotz, A., Russwinkel, N., & Wohlfarth, E. (2019). Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F*, *64*, 532-551. doi:10.1016/j.trf.2019.06.008

McDonald, A., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward Computational Simulations of Behavior During Automated Driving Takeovers: A Review of the Empirical and Modeling Literatures. *Human Factors*, *61*, 642-688.

Nilsson, J., Falcone, P., & Vinter, J. (2015). Safe Transitions From Automated to Manual Driving Using Driver Controllability Estimation. *IEEE Transactions on Intelligent Transportation Systems*, *16*, 1806-1816. doi:10.1109/TITS.2014.2376877

Ohn-Bar, E., & Trivedi, M.M. (2017). Are all objects equal? Deep spatio-temporal importance prediction in driving videos.*Pattern Recognition, 64,* 425-436.

Ross, T.J. (2010). *Fuzzy Logic with Engineering Applications*. United Kingdom: John Wiley and Sons, Ltd.

SAE J3016. (2018). *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*, June 2018.

von Benda, H. (1977). Die Skalierung der Gefaehrlichkeit von Verkehrssituationen. I. Teil: Ein Klassifikationsschema für Verkehrssituationen aus Fahrersicht. *FP 7320 im Auftrag der Bundesanstalt für Strassenwesen*. München, Technische Universität München.

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*, *64*, 285-307.

Discussion

The following discussion is organized into two main sections. Firstly, a summary of Paper No. 1 – 4 is presented. This includes drawbacks of concepts, methods, results and general weaknesses. Secondly, a comprehensive discussion on the overall findings with regard to the introduced research questions of this dissertation is given. This includes an outlook on future research aspects within the domain of this dissertation.

Summary of papers

Paper No. 1 (Lotz, Russwinkel & Wohlfarth, 2019) presents empirical research of an extensive moving-based simulator study with 88 professional truck drivers. Each participant experiences nine critical take-over situations during Level 3 automated driving. The paper presents the evaluation of reaction times with regard to the independent variables *automation duration*, *type of take-over situation* and *non-driving related task*. Additionally, *learning behaviour* for new interactions with a previously unknown automation function is investigated.

The obtained data of Paper No. 1 is utilized in Paper No. 2 (Lotz & Weissenberger, 2019) to investigate the possibility of predicting reaction times with data-driven computational models for individual assistance. A classification problem was defined, with four reaction time classes of varying width and uneven weighting. Four different machine learning algorithms (k-nearest neighbour, Bayes classifier, support vector machine and a multilayer perceptron neural network) were analysed with regard to their suitability and accuracy. The results show that for limited amounts of data, support vector machines offer the lowest misclassification rate. Overall, eye-tracking data and vehicular sensor data was not sufficient to accurately classify take-overs into the correct reaction time classes with satisfactory accuracy. Rather, no clear correlation was determined between physiological and environmental features to reaction times. This is attributed to the high intra- and inter-individual differences drivers present during Level 3 automation and due to the confined experimental setup.

The focus of the study presented in Paper No. 3 (Lotz, Russwinkel & Wohlfarth, 2020) was to validate the reaction times obtained in the simulator study of Paper No. 1 in real-world conditions. The empirical study of Paper No. 3 was conducted with a prototypic semi-truck, enabling the driver a Level 3 automation function on a test track. The effect of expectancy of take-over situations was investigated. Reaction times calculated from the empirical data presented very quick reaction times compared to the literature, similar to Paper No. 1. Expectancy of take-over situations did not generate significant differences in reaction times by the professional truck drivers.

Paper No. 4 (Lotz, Russwinkel, Wagner & Wohlfarth, 2020) introduces a prototypic advanced driver assistance system. Its aim is to predict environmental driving situations in which the driver feels take-over is necessary and self-initiates control, although no warning is initiated. The paper provides a description of the conceptual function of the system. On the basis of sideswipe manoeuvers, the potential accuracy of the system is discussed. Additionally, the theoretical constructs of objective and subjective driving complexity are introduced. The conceptual assistance system thereby also functions as a carrier technology, to enable the future investigation of differences between objective and subjective complexity.

Limitations/Improvements

Paper No. 1

The extensive moving-base simulator study of Paper No. 1 was designed to investigate professional truck driver behaviour in critical take-over situations. In parallel, the long-term goal of generating usable data for the investigation of computational models for driver predictability was pursued. The experimental methodology of the study in Paper No. 1 had to be adapted, to also serve this secondary objective, see Paper No. 2. Especially the generation of multiple take-overs by each participant was paramount to gather enough data, to train data-driven computational models. Naturally, this causes certain drawbacks, which are discussed in the following.

The study was comprised of a complex test scenario with many different groups and variants. A complete balanced design was not possible, due to the independent variables; *automation duration*, *non-driving related task*, *take-over situation* and *trial*. A complete balanced design would have required nine factorial (9!) variants. In sum, 768 take-overs were successfully recorded.

The experimental design was also arranged to only incorporate take-over situations that were highly critical, with times to collision of 2.4 sec to 5.8 sec. It was reasoned during the design that by presenting short times to collision, the drivers would be more inclined to react quickly. This seems to be the case, with the statistical evaluation of the reaction times showing a 3σ-interval of 0 sec to 2.82 sec. The downside of only presenting take-over situations that force the driver to react quickly, is that a broader variance and therefore a broader picture of the drivers' reactions was not formed. In hindsight, it would have also been valuable to gather data of take-overs with higher times to collision by identical drivers. This would have probably caused a higher variance in reaction times and possibly produced a different predictability of the investigated machine learning algorithms of Paper No. 2. However, due to the previous exclusion of take-over without high time criticality, this was not considered.

A restricting factor of the presented study is the choice of take-over situations or rather when the optical acoustic warning occurred. As discussed in Paper No. 1, some of the take-over situations were identifiable prior to an optical acoustic warning. Through the continuous or sporadic observation of the driving environment, three take-over situations (Fog & Warning Triangle, Lane Narrowing and Cut in Right) were identifiable prematurely. This resulted in significantly quicker reaction times for these take-over situations. If a participant did not identify one of these three situations prior to the warning, the reaction times were statistically identical to those of the other six take-over situations.

After each take-over situation, the drivers were presented with a question regarding the perceived subjective criticality. The investigation revealed that especially the subjective criticality had a negative correlation with trials, yielding lower subjective criticality after more observed take-over situations (trials). Due to language limitations of some of the participating truck drivers, a simple question was formulated, to query the subjective experience of each take-over. "Was the time for take-over sufficient, short or much too short?" The question was asked verbally 20 sec after a request to intervene, through a standardized auditory query. To ensure that all participants could understand the question, the question was referenced towards the available time to react. A standardized Likert-scale (Likert, 1932) was not utilized, to ensure understanding of the participants and the possibility of answering verbally. During questioning, the drivers were continuously driving manually and should not be distracted through complex interaction. In this regard, it should be noted that the subjective criticality construct was not defined adequately as a trade-off for non-native speakers. A large variety of different factors could have influenced drivers to respond to one of the three answers of the verbal question after take-over. Instead of defining and identifying subjective criticality, it should be pointed out that the results could also be summarized as subjective experience of take-over not criticality.

An important point needs to be raised regarding the quality of take-over. Data recording was aimed at recording reaction times, while an identification of take-over quality was determined on the occurrence of a crash. This binary consideration of take-over quality does not portray the complexity of identifying safe and unsafe take-overs. More research on the techniques and approaches of classifying the safety of take-overs is required. In retrospect, take-over quality should incorporate further facets. This could include whether the reaction of the driver would cause other drivers around the ego-vehicle to react, setting off a chain reaction of driver responses. Following vehicles should be excluded from this concept, as a braking automatically requires a reaction of vehicles behind any decelerating object. A high take-over quality thereby would include measures such as the initiation of interaction with other vehicles situated laterally to the ego-vehicle and safety critical behaviour to which the ego-driver subjects themselves.

Based on the premise of gathering a large variety of data on multiple independent variables, the experimental design provided the desired knowledge. As critical take-over situations thus far had not been investigated in the truck context, the experiment closed a vital knowledge gap in the research field. It addressed the research question, focusing on the cognitive and physical capability of truck drivers. As discussed, the data was utilized for Paper No. 2 and the predictability of driver take-over reaction times.

Paper No. 2

The quality of the data decisively influences the quality of the results presented in Paper No. 2. No concrete measure was formed of data quality prior to investigating the classification of reaction times. The complete sensor data available from the study presented in Paper No. 1 was utilized for the classification task in Paper No. 2. Especially eye-trackers are known for their susceptibility to external factors such as lighting and subject positioning (Nyström, et al., 2013). As a main model input, eye-tracking data could have a major influence on the results of the classification task. Although 291 trails were disregarded in the calculation of the time to eyes on road of Paper No. 1, the complete dataset was utilized for the machine learning classification. A distinction of data conclusiveness at take-over, as argued in Paper No. 1 (section 3.1) was not conducted. This had two major reasons. Firstly, data prior to take-over was also considered for the classification task and could render information, although at take-over data was corrupt. Secondly, even with partially inconclusive data, a computational algorithm should be able to conduct a meaningful predictive classification.

The results presented in Paper No. 2 only regard the gaze behaviour and interaction with a secondary task for driver-centre features. The body posture estimation, which was also recorded, was not submitted in the publication. After publication, a calculation for the offline dataset revealed a further decreasing of the misclassification rate for the support vector machine to 32.6%. Specifically, the position of the left and right knees were identified as a major contributing feature for misclassification reduction. Further data that could be introduced to future investigations of this classification problem include personal data such as age, yearly mileage and daily wake-up times.

Regardless of the quality of input data, the main shortcoming of Paper No. 2 is the reduced dataset that was utilized for the reaction time prediction. Based on the fact that the reaction times presented a positively skewed distribution, classes for the classification problem were either unevenly weighted with data tuples or with varying class widths. As wide classes would not yield a precise estimation of reaction times, the former approach was chosen. This meant that a large amount of data was clustered between the mean of the distribution. At random, the misclassification rate for a four class problem would lie at 75%. However, as the second class held the majority of data tuples, by placing all test take-overs into this class, misclassification would be as low as 43%. The results of 38.7% (post-publication with body posture 32.6%) therefore show only a slight improvement in the misclassification rate.

Consequently, driver monitoring did benefit the predictive accuracy of the four class problem. It was still much too high for a system considering safety critical take-overs. Other problem definitions such as designing the prediction method into a regression problem, could potentially improve predictability.

Only six trials were outside of the 3σ-interval. Unfortunately, especially these six take-overs are the most interesting from a human-machine-interaction perspective, as communication between the driver and system was not ideal. The lack of occurring relevant take-overs is common in testing (Winner & Wachenfeld, 2013). An adequate dataset of relevant take-overs over the 3σ-interval, with similar size to the data received of the study in Paper No. 1, would require testing of 585.000 km in the moving-base simulator. This is equivalent to 7315 hours of testing, a highly implausible feat.

It would be ideal to fit multiple vehicles with driver-facing sensors capturing data during conditional automation and thereby generating more data for the introduced computational data-driven models. Regardless of the technical capabilities of such an automation, without clear strategies how to predict driver interaction during critical take-overs, this is not possible. The results seem to be disconcerting and leading research towards a chicken-and-egg problem. Without sufficient data, such a predictive system, analysing the driver and executing informed forecasts on their reaction during Level 3, is not acceptable.

As the data analysis of Paper No. 1 and 2 revealed a correlation between premature visibility of take-over situations and the reaction times, this was analysed further in Paper No. 4. Although driver monitoring did not produce a predictive system capable of higher accuracy based on the available data, a detailed investigation of environmental factors could contribute to a higher accuracy in the prediction of driver behaviour. Based on the available data of Paper No. 1, it seems that the identification of driver availability and take-over predictability in critical take-over situations can improve marginally based on driver monitoring. As pointed out, critical take-over situations only form a subset of all take-over situations and the introduced models can potentially function with higher accuracy in other types of take-overs.

Paper No. 3

The study presented in Paper No. 3 displays the results of a test track experiment with a vehicle enabling Level 3 automation. Due to the restricted driving environment, the results are not fully comparable to public roads. Drivers experienced a steep banked curve after a 4 km straight, which concedes that they were on testing grounds. The amount of traffic was also unusual when compared to public roads, as vehicles regularly reached 250 km/h and the traffic density was quite low. A maximum of 20 vehicles were allowed on the test track. The truck was typically the slowest vehicle on the circuit, with occasionally another slow vehicle in the periphery. Additionally, it was required that the truck drove 90 km/h to meet the track requirements. This speed exceeds the maximum allowed speed on highways in most of Europe, where the limit is 80 km/h. The experimental design also required that one experimenter and one safety engineer was present at all times in the truck. This is uncommon for a semi-truck scenario, as drivers are typically alone in the vehicle.

Contrary to the moving-base simulator study of Paper No. 1, the real-world study had an extensive training phase for safety measures. In addition, no critical objects were presented during the drive, due to track regulations and as a safety measure to avoid accidents. Criticality was induced without a critical object either by an unforeseen warning or due to a rapid impulse on steering wheel. The hypothesis was established that criticality would be perceived equally as high as in the moving-based simulator experiment due to these two events. A validation of this hypothesis could not be conducted as no comparison was possible. The participants were not identical for the studies.

A measure of take-over quality was not generated. In the paper, the argument is made that motivation and strategy plays a major role how a take-over is interpreted. Again, the interpretation of driver responses, especially the take-over quality, is difficult without the clear identification of intentional and unintentional actions. As an example, if a driver chooses to drive closely to another object as opposed to barely missing it through an unexpected reaction, typical take-over quality measures fail to differentiate. This calls for a new definition of quality measures including driver strategy and intention. This identified missing link in the research field would greatly benefit the discussion of take-over behaviour in future research.

Overall, a similar investigation with a vehicle in real-world conditions has not been published previously. Multiple health and safety regulations of the test track narrowed down the investigation possibilities. Due to the weather conditions, a large portion of eye-tracking data did not yield sufficient results or could not be synchronized.

Paper No. 4

Paper No. 4 discusses the possibility of exploring the differences of subjective and objective complexities. While this is a relevant and interesting research field, the paper is limited to focusing on the introduction of a new assistance system. A study investigating the two abstract concepts is not included. Rather, the assistance system of Paper No. 4 can function as a system for the investigation of the differences between objective and subjective complexity. By learning situations in which the driver regains control of the vehicle through self-initiated take-over, the objective data of the situations can be gathered. This allows the statistical analysis of a binary subjective complexity rating to kinematic metrics describing the environment objectively.

The prototypic assistance system introduced in Paper No. 4 currently only takes a limited amount of kinematic metrics into account. This is mainly because only radar data was utilized in the exemplary analysis presented in Paper No. 4. Additionally, it should be investigated whether or not the interpretation of raw camera data benefits the predictive model towards higher accuracy or redundancy of signals. Future development stages will also consider crossing trajectories of all vehicles when exiting the safety corridor, as introduced in the publication. The classification of objects according to attributes such as type is not incorporated in the proposed version and could possibly be beneficial. Such additional attributes unknown to the system currently could be the cause for differences between subjective and objective complexities.

The sideswipe data available for a preliminary investigation of functionality is limited. A total of 118 sideswipe manoeuvres were not enough to train the complexity prediction system adequately. Further investigation should address the reasoning of failed situation recognition. As the system summarizes all kinematic metrics, the failure of singular kinematic metrics still allows high prediction rates. However, the investigated situations often had a high baseline of predicted occurrence. This can either be caused by input data being inaccurate or the investigated sideswipe manoeuvres being indecisive with the chosen kinematic metrics. A weighting function could address both of these effects in future development. In parallel, the application of filters and signal pre-processing could improve prediction accuracy.

A major shortcoming of the introduced system is the fact that it learns based on previous situations. Therefore, it can never truly predict a new situation that is unique from previous situations. The introduced assistance system considers variance in the input signals; however, a completely new situation cannot be identified by such a system. Therefore, a threshold amount of training is required in order to have a functioning system. This training will certainly almost always exclude situations due to the lack of occurrence. Similarly, this highlights the main shortcoming of data-driven computational models and the current approach of realizing automated driving.

Overall experimental limitations

In all studies, the same vehicle was utilized with an automation function that did not change in configuration or layout. The vehicle also never changed in appearance. Therefore, the results obtained are limited to the vehicle and human-machine-interface factors the system displayed, as defined by Vogelpohl et al. (2016).

All data evaluated in the studies introduced in Paper No. 1 and No. 3 included 108 participants. Although all drivers held a semi-truck driver's license, driving experience varied. Especially in the test track study, the variance of experience was higher. Due to the location of the test track in a rural area, the amount of applicants was lower than in the previous simulator study. All data and conclusions drawn are limited to this selection of participants with generalizations drawn on this dataset.

While the *time to first reaction* was measured based on steering angle and steering torque, the other two metrics relied on remote sensors, making the signals not as robust. Especially the *time to eyes on road* relied on an eye-tracker that was prone to external influences, such as lighting, reducing the accuracy. Finally, the *time to hands on steering* was recorded with four capacitive sensors integrated into the steering wheel in the study of Paper No. 1. If drivers kept their hand close to the steering wheel without interacting, or if the hands touched areas that were not covered by the sensors, the time was not recorded accurately. Regions that were especially problematic included the spokes connecting the inner and outer regions of the steering wheel.

Overall

Reaction times in critical take-over situations

This thesis investigated the behavioural reactions of professional drivers in semi-trucks in Level 3 automated driving. The recorded reaction time metrics consisted of the *time to eyes on road*, *time to hands on steering* and the *time to first reaction*, as defined (Damböck, 2013) and presented in in Paper No. 1, No. 2 and No. 3. In the passenger car context, large collections of investigations have been published in recent years. Most of these studies aim at referencing environmental factors to reaction times and inferring psychological effects, e.g. (Ruscio, et al., 2015; Hergeth, et al., 2016; Jarosch, et al., 2017). In their meta-study, Zhang et al. (2019) investigate universal effects found in 129 studies. One of the main findings is the strong effect of time budget (time to collision), with less time budget resulting in lower reaction times. As argued in Paper No. 1, by permitting larger time budgets drivers seem to utilize this time to observe the driving environment. Quick reactions are necessary with lower time budgets, in order to obvert accidents and safely navigate the vehicle through the situation. There seems to be a minimal time budget required to form a situational awareness (Endsley, 1995), where physical constraints of the drivers and misconception of the environment cause higher accidents. This time is extremely difficult to determine, as drivers have the possibility of readjusting their reaction after initial interaction. Hence, in current experimental studies examining situation awareness, situations are dynamic and can affect the initial composition of situation awareness. A continuous updating of situation awareness occurs, possibly also improving reactions. As both studies of this thesis primarily addressed research questions regarding behavioural reactions, a focused investigation of situation awareness was not possible.

A universal comparison of effects between the truck and passenger car context was partially possible. The results of Paper No. 1 and Paper No. 3 indicate that truck drivers can present extremely controlled and quick reactions. Especially with cargos of permissible 44 tons for drivers in many European countries, swerving and confined lateral control is paramount. When comparing the reaction times of the samples of the presented studies in this thesis with published results in the passenger car context, it seems that truck drivers react quicker. However, a valid statement can only be made in a controlled comparison with limiting independent variables and participants of varying experience. Generally, truck drivers do experience travel on the highway more frequently and the critical situations encountered are limited by road restrictions affecting the scope of permissible driving. Therefore, a confined set of situations can arise on the highway that is certainly broader for the less experienced passenger car drivers.

Throughout both the simulator and test-track study, the times to collision were below six seconds. Nevertheless, no accidents occurred. This could signify that the maximum capabilities of the drivers were not demanded. However, in both studies dramatic consequences due to ill-informed actions at take-over had minimal effect. As no vehicles were within the direct periphery of the ego-vehicle, the only objective for the driver was to obvert the object causing take-over. As discussed in Paper No. 1, more confined take-over situations would have certainly generated more accidents. However, this would have filtered the drivers into three classes based on their take-overs. Firstly, those who reacted poorly and caused an accident with near vehicles. Secondly, drivers who had enough situational awareness to lead the vehicle through the take-over situation successfully with intent. Thirdly, drivers choosing the correct route by chance. A differentiation between the second and third group would be difficult. Therefore, this methodology was not adopted, to allow all drivers to complete the experiment and ensure that enough data would be generated for Paper No. 2.

Apart from time budgeting, other influencing factors were investigated. A clear correlation between singular factors and reaction times does not seem to exist. Influences of driver and environmental factors as well as vehicular and HMI configurations interact. Therefore the data-driven approach of Paper No. 2 could possibly work to some extent, if enough data were available. A first successful application was presented in Paper No. 2.

Overall times to first reaction below three seconds seem highly plausible and relevant for a large amount of critical take-over situations. However, this does not consider take-over quality. Depending on the visual information perceived by the driver before or at a request to intervene, this threshold could be far lower. Similarly, the empirical studies also portrayed situations in which the driver capabilities would not yield sufficient safe transitions.

Therefore, it is important to consider take-over quality when analysing reaction times or any metric describing take-overs in automated driving. Typically, different mathematical metrics are consulted when quantifying the take-over quality such as minimal time to collision, lateral acceleration (Gold, et al., 2016) and mean longitudinal accelerations (Radlmayr, et al., 2014). Based on the data gathered in Paper No. 1, these metrics do not paint the picture fully, as all of the abovementioned units disregard driver strategy and intent. Additionally, the intensity of these metrics describing an adequate take-over may vary due to situational demands. As an example, consider a lane change task. Here the ego-vehicle has a stationary object in front of it, dictating the lane change. While there is still a large distance to the object, the other lane is occupied and a free space is apparent after the vehicles pass that are currently occupying the overtaking lane. By continuing to approach the object on the ego-lane, the metric time to collision reduces. However, due to the drivers' strategic decision and experience, this reaction is desired and intended. Depending on the driver preference, the lateral and longitudinal acceleration can also vary when merging into the intended lane. Hence, current metrics for the description of take-over quality are imperfect and should not be considered dogmatically. Ideally, the quality metrics should be complemented by expert ratings, similarly to the way a passenger can interpret driver competence. The combination of expert rating concerning the calmness (strategic planning) and quality metrics could be hugely beneficial in avoiding incorrect causal reasoning of take-over quality and predicting driver capability.

Prediction of driver behaviour with computational models

The algorithms and computational models inferring driver state, based on sensory driver monitoring systems, have to be reliable, robust and accurate. Therefore, recent research has addressed how sensor setups need to be configured to guarantee valuable information of the driver (Ruff, 2017). In the context of this thesis, an analysis of eye-tracking setup was developed (Fuhl, et al., 2019). Misunderstanding of the state can cause the system to initiate communication in ill-fit moments and sharing of the driving task in situations is requested, which are not suitable for a driver. This can lead to situations, which are far worse than if no automation were ever available.

Different types of models and algorithms can serve as a foundation for driver state predictions. This thesis focuses on a data-driven approach, by utilizing machine learning algorithms to predict take-over reaction times in critical situations. A problem that can result from this approach is the unaccounted psychophysiological effect on the human. By only matching results with input data, human decisions generating the resulting take-over are disregarded. It should also be noted that computational models such as those generated by cognitive architectures, also offer a possibility of predicting take-over behaviour that is not data-driven (Prezenski, 2018). This type of computational model relies on behavioural data and the interpretation by an experience modeller. However, data-driven models offer a scalability with regard to input and output data that cognitive architectures due not. Especially in the context of an applied approach for large quantities of vehicles and drivers, this allows data-driven models to adapt more fluidly.

As investigated in this thesis, take-over can either be required by an automation function or self-initiated by a driver. Depending on these two types of take-over variants, assistance needs to vary. For both variants a safe and easy take-over is the optimal outcome. When comparing a system-generated with a self-initiated take-over, one deciding factor differs. For most self-initiated take-overs an internal goal, due to interpretation of the driving environment, causes the driver to launch take-over procedure, while the human can completely be out-of-the-loop during a system-generated request to intervene. It should be noted that trust in the system might not be the sole deciding factor for an unforced or self-initiated take-over. Other factors such as surprise or false interpretation of the driving situation can motivate the driver to take-over. Nevertheless, in both take-over variants the human needs to be assisted in regaining manual control and full sensory perception of the situation. Therefore, the goal of a computational prediction differs slightly for these two cases. For a self-initiated take-over, the goal of a computational prediction aims at identifying the driver's correct sensory perception of the environment. A representation of current situations in the environment is highly relevant. For a system-generated take-over, the goal of a computational prediction focuses on the prediction of the correct driver state. Instead of environmental cues, the current driver interaction with non-driving related tasks holds a higher importance.

Based on the findings of both introduced computational models, a limited prediction of take-over readiness and ability can be made. There is no computational model that can accurately describe driver behaviour, as non-intrusive sensors cannot measure internal thoughts or reasoning. However, tendencies could be identified in the work of this thesis. Herzberger et al. (2018) present an interesting concept that could be constructive in the objective of predicting take-over behaviour. Instead of searching for a possibility of computing models of successful take-over behaviour, missing take-over capability might be easier to detect. An adaptation of this approach is implemented by the assistance system in Paper No. 4, which specifically tries to assist in situations that were identified as critical by the driver.

In the case of self-initiated take-over prediction, it should be possible to statistically analyse the data that is generated by the system post hoc, to cluster situations and identify underlying environmental reasons for take-over. Research has shown interesting correlations between environmental criticality and uncertainty to correct reactions (Stoll, et al., 2019). Thereby, it could be possible to determine kinematic thresholds, which allow the recognition of those situations that enforced a self-initiated take-over. This information would also be vital for the application of cognitive models as cognitive architectures, such as ACT-R (Anderson, et al., 2004), rely on rules to compute human behaviour. By identifying kinematic relationships in perceived criticality during take-over, human-machine-interaction could also be modelled with this approach.

Predictive driver assistance system

Overall, the prediction of driver behaviour through computational models is always prone to error. As most computational models rely on statistical analysis of large data, this makes the adaptation for individuals difficult. This is not only the case for data-driven models. Rather, trade-off of false-positive and true-positive results in model predictions has to be accounted for in the design of future assistance systems. While information might improve human-machine-interaction in take-over scenarios, false prediction can also reduce take-over capability of drivers or cause low acceptance of assistance systems.

In a preliminary study, Wiese, Lotz & Russwinkel (2019) address a possibility of connecting ACT-R with vehicular data. By introducing a new visual module in ACT-R based on the SEEV approach (Wickens, 2015); this bids the possibility of including bottom-up and top-down processes and semantically annotating real-world scenes. The combination of this research with the assistance system of Paper No. 4 should allow the consideration of environmental factors on a cognitive model, simulating driver interaction. Future research should investigate if this combination can reduce the amount of false positives the complexity prediction system of Paper No. 4 exhibits.

It is clear from the research conducted that for a predictive system to function at peak performance, information is required of the driver. However, inter- and intra-individual differences in behaviour are too great for a system to rely on generic models. By making it possible for assistance systems to monitor a driver, false conclusions can be evaded partially. However, complexity of human behaviour will not allow for precise prediction of drivers in near future. Especially in situations in which responsibility is shifted between human and machine in Level 3, the research conducted in this thesis explains some of the underlying influences. In addition, a system is introduced that addresses a problem previously ignored during the operation of an automation function. By predicting self-initiated take-overs, adaptive assistance is enhanced in phases in which previously drivers presented discomfort due to misaligned interpretation of the environment compared to automation. This can further improve the human-machine-interaction during Level 3 automation due to perceived subjective complexity in the environment. Assistance will not diminish shortcomings in human driver behaviour in the future. With the approach of developing assistance systems for automated driving, the goal should be to adaptively and situationally assist drivers. At the same time, as humans and machines switch their roles of observer to controller, prediction will fail in some cases. This has to do with the definition of Level 3 automated driving rather than with application of assistance systems.

Bibliography

Anderson, J. R. et al., 2004. An Integrated Theory of the Mind. *Psychological Review*, 111(4), pp. 1036-1060.

Damböck, D., 2013. *Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme,* München: Technische Universität München, Lehrstuhl für Ergonomie.

Endsley, M. R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors Journal 37(1)*, pp. 32-64.

Fuhl, W. et al., 2019. Ferns for area of interest free scanpath classification. In: *2019 Symposium on Eye Tracking Research and Applications (ETRA '19), June 25-28, 2019.* Denver, CO, USA: ACM, New York, NY, USA.

Gold, C., Korber, 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), pp. 642-652.

Hergeth, S., Lorenz, L. & Krems, J. F., 2016. Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, pp. 457-470.

Herzberger, N., Eckstein, L. & Schwalm, M., 2018. Detection of Missing Takeover Capability by the Orientation Reaction to a Takeover Request. In: *27. Aachen Colloquium Automobile and Engine Technology.* Aachen, Germany: Aachener Kolloquium Fahrzeug- und Motorentechnik GbR, pp. 1231-1240.

Jarosch, O., Kuhnt, M., Paradies, S. & Bengler, K., 2017. It's Out of Our Hands Now! Effects of Non-Driving Related Tasks During Highly Automated Driving on Drivers' Fatigue. *Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, pp. 319-325.

Likert, R., 1932. A Technique for the Measurement of Attitudes. *Archives of Psychology*, Issue 140, pp. 1-55.

Lotz, A., Russwinkel, N., Wagner, T. and Wohlfarth, E., 2020. An adaptive assistance system for subjective critical driving situation: understanding the relationship between subjective and objective complexity. In: D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2019 Annual Conference*. Nantes, France.

Lotz, A., Russwinkel, N. & Wohlfarth, E., 2020. Take-over Expectation and Criticality in Level 3 Automated Driving - a Test Track Study on Take-over Behaviour in Semi-trucks. *Accident Analysis & Prevention*.

Lotz, A., Russwinkel, N. & Wohlfarth, E., 2019. Response Times and Gaze Behavior of Truck Drivers in Time Critical Conditional Automated Driving Take-overs. *Transportation Research Part F*, *64*, pp. 532-551.

Lotz, A. & Weissenberger, S., 2019. Predicting take-over times of truck drivers in conditional autonomous driving. In: N. Stanton, (Ed.) Advances in Human Aspects of Transportation. AHFE 2018. *Advances in Intelligent Systems and Computing, 786*, pp. 329-338.

Nyström, M., Andersson, R., Holmqvist, K. & van de Weijer, J., 2013. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods*, March, 45(1), pp. 272-288.

Prezenski, S., 2018. *Interactive learning of dynamic systems: A cognitive modeling approach to mental model building and updating,* Berlin, Germany: Technische Universität Berlin, Doctorial Thesis.

Radlmayr, J. et al., 2014. How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, I(Vol. 58), pp. 2063-2067.

Ruff, S., 2017. *Driver cognitive workload - A comprehensive multi-measure approach,* Berlin, Germany: Technische Universität Berlin, Doctoral Thesis.

Ruscio, D., Ciceri, M. R. & Biassoni, F., 2015. How does a collision warning system shape driver's brake resonse time? The influence of expectancy and automation complacency on real-life emergency braking. *Accident Analysis and Prevention*, Issue 77, pp. 72-81.

Stoll, T., Müller, F. & Baumann, M., 2019. When cooperation is needed: the effect of spatial and time distance and criticality on willingness to cooperate. *Cognition, Technology & Work,* 21, pp. 21-31.

Vogelpohl, T. et al., 2016. *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung,* Berlin: Gesamtverband der Deutschen Versicherungswirtschaft e.V..

Wickens, C. D., 2015. Noticing events in the visual workplace: The SEEV and NSEEV models. In: R. R. Hoffman, et al. eds. *Part VI - Perception and Domains of Work and Professional Practice.* Cambridge: Cambridge University Press, pp. 749-768.

Wiese, S., Lotz, A. & Russwinkel, N., 2019. SEEV-VM: ACT-R Visual Module based on SEEV theory. *International Conference on Cognitive Modeling (ICCM 2019)*, 19-22 July.

Winner, H. & Wachenfeld, W., 2013. *Technische Universität Darmstadt - Absicherung automatischen Fahrens.* [Online]
Available at: https://www.fzd.tu-darmstadt.de/media/fachgebiet_fzd/publikationen_3/2013/2013_Winner_FAS_Muenchen_Absicherung_automatischen_Fahrens_3_0_Praesentation.pdf
[Accessed 20 August 2019].

Zhang, B. et al., 2019. Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F*, pp. 285-307.

Thank you

Although my name presides this thesis, it would not have been possible without the support from colleagues, co-authors, friends and family, to whom I am very thankful.

For the support, advice, guidance, trust and open reigns over the past four and a half years, I would like to thank Nele Rußwinkel sincerely. Thank you, for allowing me to jump into a world of cognitive modelling I was previously unfamiliar with. Bouncing back ideas and finding applications for technical systems was always enriching, informative and enjoyable. Also, thank you for your time and advice over the past three years of my doctorial phase at Daimler Trucks AG. It meant trusting me with conceptual ideas, as distance inevitably reduced meeting frequency. Finally, thank you for helping me as a co-author on the majority of my academic work and for reviewing this thesis.

Secondly, a much appreciated thank you to Enrico Wohlfarth for supporting and supervising my work over the past three years. You assisted with great ideas, technical support, a network of knowledge and great banter, with which this work was only possible.

Thirdly, thank you Martin Baumann for agreeing to review this thesis and allowing me to present my work over the past years.

Thank you Stefanie Marker for agreeing to chair my defense.

I would like to thank the unknown reviewers for their advice and time in improving all papers.

To all co-authors, thank you for the time, ideas and work you put into our publications. Their standard would be lower without your input.

Thank you, to all my colleagues of the past three years at Daimler AG. A list of names would exceed this page. Each one of you has helped me understand or discuss a matter of which I previously was not as good at or informed.

Special thank you to Thomas Wagner, Markus Gekeler and Stephanie Preuss, who helped with every technical question of the different experimental setups. Thank you for your support and effort Sarah Weissenberger, Hugo Dabet and Simon Schuck.

I would also like to thank all my colleagues from the TU Berlin for the support; Sabine Prezenski, Oliver Klaproth, Stefan Lindner, Lisa Dörr, Marc Hallbrügge, Sebastian Wiese, Alexandra Weidemann, Kai Preuß and Marlene Scharfe.

Thank you, to my family and friends who supported me with indulgence, advice, support or just an open ear.

Finally, thank you to two uniquely irreplaceable people in my life. One knows who she is; the other will grow up to know.