User and System within the Context of Use –

How Users Cope with Partial Automation

Von der Philosophischen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades einer Doktorin der Philosophie genehmigte Dissertation

vorgelegt von

Johanna Josten

Berichter:
Univ.-Prof. Dr. phil. Martina Ziefle
Univ.-Prof. Dr.-Ing. Lutz Eckstein

Tag der mündlichen Prüfung: 20. November 2020

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Vorwort

Mein Dank gilt Univ.-Prof. Dr. phil. Martina Ziefle für die Betreuung dieser Arbeit, insbesondere für die motivierende Diskussion zur Einordnung und Strukturierung der Inhalte. Danken möchte ich auch Univ.-Prof. Dr.-Ing. Lutz Eckstein. Seine Überzeugung vom Wert interdisziplinärer Zusammenarbeit am Institut für Kraftfahrzeuge (ika) hat diese Arbeit erst ermöglicht.

Dem Teilnehmerkreis aus FAT AK2 und BASt danke ich für die wertvollen Diskussionen im Rahmen des von beiden Institutionen gemeinsam geförderten Projektes zur Untersuchung der Bewältigungsleistung des Fahrers von kurzfristig auftretenden Wiederübernahmesituationen nach teilautomatisiertem, freihändigen Fahren, in dem ich einen Teil der benötigten Daten erheben durfte. Gleichfalls danken möchte ich dem Profilbereich ICT der RWTH Aachen, dessen Förderung mir eine Vertiefung und Erweiterung der Untersuchungsfragen ermöglicht hat.

Mein besonderer Dank gilt Felix Fahrenkrog und Magali Kreutzfeldt, nicht nur für ihre wertvollen Anmerkungen zu früheren Versionen dieser Arbeit, sondern auch für ein stets offenes Ohr über diese Arbeit hinaus und ihre uneingeschränkte Bereitschaft, Ideen und Herausforderungen aus technischer oder psychologischer Sichtweise zu diskutieren.

Meinen aktuellen und ehemaligen Kollegen der Bereiche Fahrzeugintelligenz und Automatisiertes Fahren an ika und fka danke ich für den einzigartigen, unersetzlichen B6-Teamgeist auch über den Arbeitsalltag hinaus – hold on to that feeling! Besondere Erwähnung gebührt an dieser Stelle, nicht nur dafür, Philipp Seewald, Mirko Kugelmeier und Alexandra Wulf. Adrian Zlocki danke ich für seine Unterstützung und seinen Glauben an ein Gelingen dieser Arbeit. Mein Dank gilt darüber hinaus dem Team Fahrsimulation, sowie Christoph Klas, Kevin Schulte und allen meinen studentischen Hilfskräften für die zeit- und geduldsintensive Unterstützung bei der Umsetzung und Durchführung der Probandenversuche. Herzlichen Dank an dieser Stelle auch all denjenigen Kollegen, die bei der teils endlos scheinenden Aufbereitung von Fahrzeug-, Video- und Blickbewegungsdaten als Grundlage meiner Analysen beratend und unterstützend zur Stelle waren, insbesondere Amogh Sakpal und Andreas Pütz.

Zuletzt möchte ich all jenen danken, die mich durch die Promotionszeit begleitet und für den manchmal nötigen Abstand zu dieser Arbeit gesorgt haben: Michael, meinen Freunden, meinen Eltern und Geschwistern.

Aachen, im April 2021

Abstract

Driving is a complex and dynamic task, associated with highly variable demands depending on the current context of driving. Vehicle automation, by relieving the driver of the active conduction of the driving task and reducing human error, is considered a means to increase both driver comfort and the safety of driving. In partial automation, as a currently available level of vehicle automation, the system takes over the active execution of the driving task in defined contexts of use, but relies on the driver to supervise both the environment and the system's behaviour within the current driving situation. Because the user is required to stay involved in the driving task, concepts explaining driver performance in unassisted driving, such as situation awareness, remain important for the use of partial automation. Moreover, various characteristics of user, system and context of use have been considered as influential for interaction quality. Sufficient monitoring behaviour and adequate mental models are prerequisites for a safe interaction with partial automation. Performance in transitions from automated to manual driving, either after a system-initiated takeover request or without prior notification of the need to intervene, is frequently considered as an indicator for the relevance of influencing factors.

Currently, the driver has to supervise automation with the hands on the steering wheel. While haptic feedback from the system has been considered beneficial for staying in the loop during automation use, hands-on supervision of automation has also been shown to reduce the comfort of use. This discrepancy between an assumed increase in the safety of use and a satisfying system design from a user perspective gave rise to investigations verifying the benefit of continuous hands-on supervision during automation use.

A comprehensive analysis of users' interaction with partial automation from initial contact to the repeated handling of control transitions was pursued in this work, with a focus on the role of haptic feedback. An analysis of users' attitudes towards systems asking for different levels of haptic involvement during use was included to assess changes in the intent to use or monitor. By variation of contexts of use, levels of user experience, abstraction levels of the driving task and in separate analyses for different interaction phases with automation, the relevance of haptic feedback and the stability of its influence were assessed. In three driving simulator studies, the duration of use, the availability of secondary tasks and the complexity of the traffic situation were considered as variations of the context of use with an assumed link to establishing or maintaining situation awareness. In addition to performance after takeover requests, the impact of haptic feedback on planned control transitions was assessed in a fourth simulator study. The investigation of post-automation driving performance, more precisely, of the end of corrections after control transitions, constituted an exploratory focus in this work. Results from the a priori assessment of automation in a survey and automation use in a driving simulator were contrasted with a posteriori assessments and automation use in a test track study.

Prior user experience influenced not only the a priori assessment of automation, but also the change in attitudes after use of specific automated systems. Overall, hands-off monitoring was perceived as more comfortable. A beneficial effect of haptic feedback on takeover performance was found for timing and takeover success as well as in the amplitude of driver interventions. Overall, immediate responses to takeover requests were observed, but hands-on monitoring resulted predominantly in faster and more controlled transitions. The differences between feedback conditions were however small. No main effects of context variations on intervention times were found. Planned control transitions generally enhanced the quality of transitions, but drivers in the hands-off condition waited longer to intervene, even when the need for action should have been clear based on the mental model established before use. Haptic feedback further benefitted the supervision of automation, with a higher amount of visual attention being attributed to the forward road. The supervision of automation varied between driving contexts, especially with the availability of secondary tasks. Post-automation driving performance reached manual performance levels shortly after control transitions, but haptic feedback proved again beneficial for the decay of transition effects. Differences between conditions on the test track were smaller and performance in transitions was more controlled than in the simulator studies.

The stability of the hands-off disadvantage over different contexts of use suggests its attribution to additional movement time. Hands-off monitoring accounted for most of the few unsuccessful takeover attempts, but more than 95 % of all control transitions were successful. In general, the effects of automation in comparison to manual driving outweighed the influence of haptic feedback during automation use for subjective, gaze and performance metrics. In summary, user characteristics, variations in context and haptic feedback proved to be of minor relevance for the safety of interactions in case of a correctly processed takeover request.

Future work should focus on control transitions without takeover requests to quantify effects of lesser monitoring associated with hands-off supervision. The experimentally controlled approach of this work benefitted the clear assignment of cause to the effects found, but might be questioned regarding the validity for user behaviour observed in daily traffic. The voluntary choice for hands-off supervision and for secondary task engagement under consideration of users' mental models might provide further insights into the relevance of haptic feedback during use of partial automation.

Content

1 Introduction and Motivation

Driving a vehicle, although presumably considered a fairly ordinary task by most, is characterised by rapidly changing demands, highly complex at times (e.g., Vollrath & Krems, 2011; Ward, 2000) and challengingly monotonous at others (Desmond & Matthews, 1997; Thiffault & Bergeron, 2003). Driver assistance as a means to help the driver in conducting the driving task more safely and comfortably (Engeln & Vratil, 2008; Stanton & Marsden, 1996) has therefore been an early focus in automotive research (Vollrath & Krems, 2011). Automating parts of human task performance, although undoubtedly connected with various benefits, is however not without risks (Manzey, 2008). Predicting adaptations in driver behaviour to changes introduced into the driving context is a relevant aspect regarding road safety (Saad, 2007). Following a common description framework of the driving task (see, e.g., Lorenz & Hergeth, 2015; Vollrath & Krems, 2011), characteristics of the driver, system and environment likely need to be considered for an accurate prediction of the driver's interaction with automated systems in traffic. To predict the influence of automation on human-vehicle interaction, "knowledge of the driving task and of the psychological processes (cognitive and motivational) that govern drivers' activity" (Saad, 2007, p. 148) is needed. Of importance, under the assumption of technologically reliant systems, are also those aspects influencing if and when drivers choose to use new technology (Brookhuis, van Driel, Hof, van Arem, & Hoedemaeker, 2009).

Current vehicle automation (i.e., Level 1 and 2 automation; SAE, 2018) relieves the driver of the active conduction of driving for a limited period of time, but requires a continued visual and cognitive engagement with the driving task to supervise the actions of the system. Visual and cognitive driver distraction, directly influencing the level of engagement with the driving task, have thus been considered widely in research on vehicle automation (Louw, Kountouriotis, Carsten, & Merat, 2015; Merat, Jamson, Lai, & Carsten, 2012). Another possibly adverse aspect of automation, next to distraction, is the physical disengagement from the driving task (Gasser et al., 2012; Othersen, 2016), as contact to the vehicle's controls is required at least to a lesser degree. As a consequence, less information is potentially received about the system's behaviour in the current situation of use via the vehicle's controls. The relevance of haptic feedback for the quality of interaction with partial automation is, albeit first indications in the sense of adverse effects (Damböck, 2013; Gold, Damböck, Bengler, & Lorenz, 2013; Othersen, 2016), however unclear, due to a lack of data on combined effects of user, system and context characteristics. This work aims to establish a more comprehensive understanding of the interaction with partial automation in this regard from the initial user contact to the repeated exposure to system limits (see Figure 1-1).

The following review of the state of the art (Chapter 2), as the basis for the formulation of research questions (Chapter 3) and the definition of subsequent experimental studies (Chapters 4.1, 5.1, 5.2, 6.1), provides a short characterisation of the driving task without assistance, followed by an overview on behaviour adaptations invoked by automating parts of the driving task and factors influencing said adaptations. Attitudes and perceptions of users are considered before (Chapter 4) and after use of automation (Chapter 7). The interaction with automation is analysed in different contexts and phases of use (Chapter 5). Analyses on the effect of specific influencing factors are supplemented by a methodological approach validating the occurrence of effects over different abstraction levels of the driving task and levels of user experience (Chapter 6). Finally, a comprehensive overview on results aims to derive conclusions on the impact of a physical disengagement from the driving task under consideration of different characteristics of user and context of use (Chapters 8 and 9).

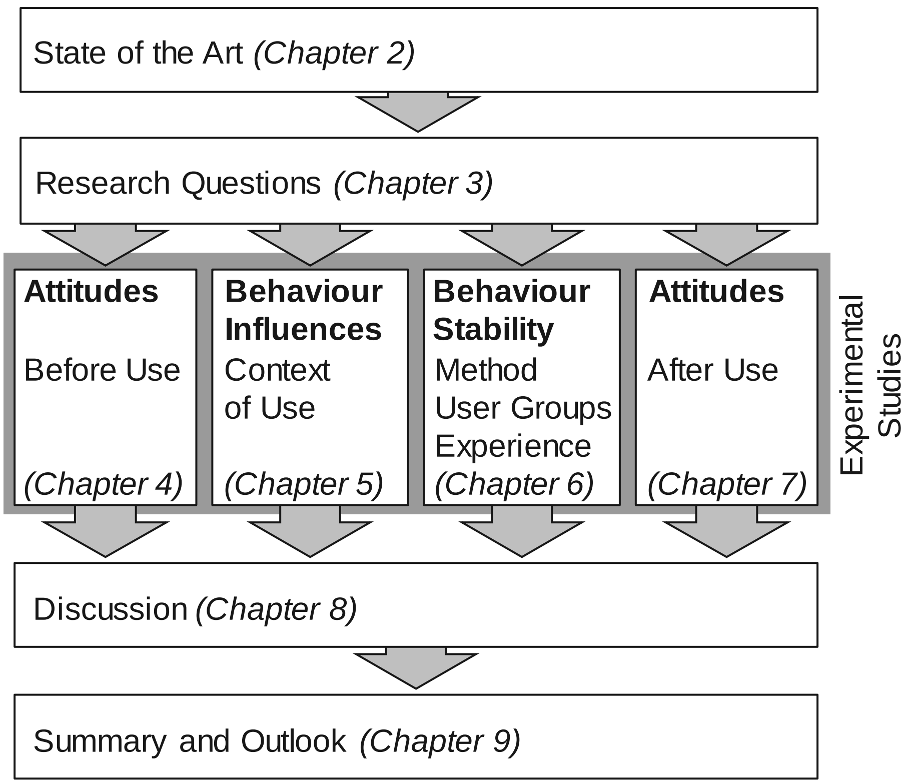

Figure 1-1: Overview on the content of this work

2 Partial Automation – Benefits and Challenges from a User Perspective

Driving is a complex and dynamic task (e.g., Gugerty, 1997; Vollrath & Krems, 2011; Ward, 2000), placing high demands on information perception and information processing. Its physical demand, that is, the translation of processed information into according actions such as operating controls, is comparatively small compared to the cognitive demands (Abendroth & Bruder, 2009). Driving requires thus both "cognitive skills", such as "decision making, prediction, selective attention and fault diagnosis", as well as "perceptual-motor skills" for operating controls (Stanton & Marsden, 1996, p. 39). When driving, "the driver perceives and interprets information, determines and initiates a response, then resumes the process by interpreting the feedback in relation to any discrepancy from a goal state, in order to sustain target performance within a margin of tolerance" (Rumar, as cited by Ward, 2000, p. 396). The driving task is defined by the driver, the vehicle and the current environment (see, e.g., Abendroth & Bruder, 2009; Donges, 2009; Vollrath & Krems, 2011), each element within this system being characterised by specific features. Different models and concepts have been introduced to describe the driving task, focusing on the skills needed and processes involved on the part of the human driver.

2.1 The Driving Task in Unassisted Driving

The driving task is frequently described on three levels, namely navigation, guidance and stabilisation (for examples, see, Donges, 2009; Vollrath & Krems, 2011). The navigation level, including tasks such as route choice and planning, requires primarily knowledge-based behaviour, relevant in new and complex situations, when different options have to be compared and behaviour has to be planned (Donges, 2009). The guidance level includes the selection of safe parameters, amongst others for distance keeping and speed (Vollrath & Krems, 2011). Rule-based behaviour is considered characteristic for this level, including the selection of the most effective strategy based on a set of learned behaviours from encountering similar situations (Donges, 2009). Maintaining a stable vehicle control by steering, braking and accelerating is part of the stabilisation level (Donges, 2009; Vollrath & Krems, 2011; Ward, 2000). Skill-based behaviour, most often being associated with the stabilisation level, is characterised by reflexive, "automatic action patterns" (Ward, 2000, p. 396; see also Donges, 2009), which can be performed with little need for attentional resources (Abendroth & Bruder, 2009). Driving experience in general and the specific situation at hand define which type of behaviour, that is, rule-based or skill-based behaviour, is applied to manage the stabilisation and guidance task (Donges, 2009).

Another important aspect related to driving experience is learning effective visual strategies to gather relevant information (Vollrath & Krems, 2011; Horswill & McKenna, 2004). Visual strategies are considered especially important as the main modality of information for performing the driving task is visual (Vollrath & Krems, 2011). However, relevant information can also be acoustic, haptic or vestibular (Abendroth & Bruder, 2009). As the amount of all information available at a time exceeds human perception and processing capacities, selection mechanisms are necessary, focussing attention on specific information either by bottom-up (e.g., by saliency of cues) or top-down mechanisms (e.g., by expectancy of the driver; for an overview, see, Vollrath & Krems, 2011). Thus, “one of the major preconditions of safe driving is that drivers correctly perceive and interpret the relevant objects and elements of the current traffic situation and that they consider these elements in planning and controlling their behaviour” (Baumann & Krems, 2007, p. 253).

Following Stanton and Marsden (1996), driving can be seen to rely on “shared resources” (p. 39), with different tasks competing for limited attention and processing capacity. Ideally, current demand leaves unused capacity for sudden changes in task demand, whereas degraded performance results from an exceedance of capacity (e.g., Wickens, 2008b; see also Parasuraman, Sheridan, & Wickens, 2008, on workload in driving). Furthermore, the cost of performing multiple tasks, for example when combining additional visual input with the driving task, has to be considered and depends on the resource demand and overlap of the current tasks (multiple resource model; for an overview, see Wickens, 2008b). The other two main principles frequently applied in driver psychology concern the development of skills, with increasing automaticity in performing the driving task reducing the amount of attention necessary to perform it, and the principle of risk optimisation, stating that behaviour is guided by the resulting subjective risk level (Stanton & Marsden, 1996).

A widely-used concept describing and integrating the processes on which drivers base decisions and perform actions in dynamic traffic is *situation awareness (SA)*. SA, being “a state of knowledge” (Endsley, 1995, p. 36), is the result of an assessment of the current situation. SA, in its most prominent definition, is defined as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1995, p. 36). The perception of elements in the current situation (Level 1 SA), the precondition for achieving higher levels of SA, highlights the need for (visual) attention towards the current environment. The comprehension of the situation, by integration and interpretation of information, forms Level 2 SA. Finally, Level 3 SA comprises the anticipation of future states based on the knowledge gathered and integrated, thereby benefitting decision-making (Endsley, 1995). SA is related to performance in that poor performance is expected “when SA is incomplete or inaccurate, when the correct action for the identified situation is not known or calculated, or when time or some other factor limits a person’s ability to carry out the correct action” (Endsley, 1995, p. 40).

To understand how drivers make decisions in traffic, considering SA on its own is not enough. The role of limited capacity of the visual system (see Vollrath & Krems, 2011) was already pointed out above, necessitating the appliance of strategies to focus attention on specific, relevant information. Additionally, available capacity for processing of information has to be considered, as “a person’s SA is restricted by limited attention and working memory capacity” (Endsley, 1995, p. 49), as well as SA-related constructs such as workload (Endsley, 1995). Against this background, the relevance of not only visual, but also cognitive driver distraction becomes apparent when considering decision making and action selection in traffic (Baumann & Krems, 2007). Visually demanding secondary tasks should impair the perception of elements, whereas cognitively demanding tasks should impair the maintenance and updating of SA (Baumann, Rösler, & Krems, 2007). In a driving simulator study, Baumann, Petzoldt, Groenewoud, Hogema and Krems (2008) tested the prediction that tasks loading on working memory impair the prediction of future states, leading to later reactions to critical events. During a 20 minute drive, participants encountered four obstacles on their lane and were warned beforehand (e.g., by a warning sign) in two encounters. Additionally, participants performed either no secondary task or one of two cognitive tasks loading on working memory functions in the relevant interval. Results showed that the time to collision (TTC) at the point where the driver released the gas pedal was significantly larger for predictable events. Additionally, the effect of predictability was greatest for the no secondary task condition. The same pattern was found for speed data when braking in front of the obstacle. The authors concluded that tasks interfering with working memory and especially tasks relevant for retrieving information from memory have detrimental effects on predicting future states, a key feature of SA.

Characteristics of the driving situation as well as driver characteristics have been connected to variations in SA by means of underlying processes such as visual attention, workload or vigilance. SA has been found to decrease with age and with increased complexity of the environment (Bolstad, 2001). In opposition, driving experience is generally thought to benefit SA, here referred to as the heightened ability for hazard perception (e.g., Pradhan et al., 2005; Wright, Samuel, Borowsky, Zilberstein, & Fisher, 2016), a term sometimes used equivalently with or as a measure of SA (see, e.g., Gugerty, 1997). “Experienced drivers have a much greater chance of having stored in long-term memory the relevant information that identifies the current situation as dangerous because their database of experienced traffic situations is much greater than that of novices” (Baumann & Krems, 2007, p. 256). In their comprehension based model of SA, Baumann & Krems (2007) hypothesised that experience will establish connections between situations encountered and the outcome of actions performed, with elements within the situation serving as trigger conditions for the selection of appropriate actions in future encounters (see also

Endsley, 1995, for an overview on the role of long-term memory for SA in other domains). Experience can thus be expected to influence the selection of appropriate actions. Effects were however found only if the benefit of experience is not diminished by distraction (Horswill & McKenna, 2004).

Against the background that driving is a complex and strenuous task, one of the aims of human factors research in the automotive sector has long since been to enable the driver to perform the driving task more comfortably by means of driver assistance (e.g., Engeln & Vratil, 2008). Driver assistance has further been promoted as a means to reduce human error, one of the main reasons for traffic accidents (for an overview, see, e.g., Stanton & Marsden, 1996; Vollrath & Krems, 2011). Driver assistance systems either assist or replace the driver at least in parts of the driving task (Stanton & Marsden, 1996; Stanton & Young, 1998; SAE, 2018). Relieving the driver of parts of the driving task presumably reduces workload (Stanton & Marsden, 1996) and increases comfort and safety (Ward, 2000). The automation of the driving task clearly offers the potential of considerable benefits, but it also changes the driver's role from performer to supervisor, "leav[ing] the driver with fundamentally different tasks to perform" (Stanton & Young, 1998, p. 1016). In the following chapter, this change in the driver's role as well as benefits and concerns related to this change will be addressed in detail.

2.2 The Driver as Supervisor During Partial Automated Driving

Automation can be applied to each of the stages of the four-stage model of human information processing, namely to information acquisition, information analysis, decision and action selection as well as action implementation (Parasuraman, Sheridan, & Wickens, 2000). Next to classifications along the stages of human information processing, driver assistance systems can alternatively be aligned along the three levels of the driving task, that is, navigation, guidance and control (e.g. Vollrath & Krems, 2011; Ward, 2000). Automation has further been considered in terms of the responsibilities shared between the system and the driver (Parasuraman et al., 2000; Flemisch et al., 2012; see also classifications by SAE, 2018; Gasser et al., 2012). Responsibilities vary in two regards, one being the active execution of the driving task (physical demand), the other referring to the supervision of a safe execution of the driving task within the current environment (cognitive demand).

In a commonly used classification, six levels of automation have been distinguished (SAE, 2018; for a similar classification, see Gasser et al., 2012). These are defined as no driving automation (Level 0, human driver only), driving assistance (ADAS; Level 1), partial driving automation (PAD; Level 2), conditional driving automation (Level 3), high driving automation (HAD; Level 4) and full driving automation (Level 5, system only). The driver is always responsible for executing parts of the driving task up to Level 1 and always responsible for monitoring up to Level 2 automation. Under the proposition that transitions in control between human and machine present a key challenge (Flemisch et al., 2012), primarily those automation levels in which the driver shares responsibilities with a system[1] within a defined driving mode, such as a traffic jam, are of interest (Level 1 to Level 3). The driver's part during shared responsibilities lies either in continuously performing parts of the driving task (Level 1), in continuous monitoring (Level 2) or in providing backup performance in case of a timely takeover request (TOR; Level 3).

The focus of this thesis is on Level 2 automation driving systems, where the system controls lateral (steering) and longitudinal (acceleration / deceleration) movement in defined driving modes with the human driver being responsible for a safe execution of the driving task within the dynamic driving environment (SAE, 2018). Regulations state that the system must allow the driver to override it at any time (UNECE, 2017a). The functioning of Level 2 automation within specific driving modes restricts the system use to specific use cases and thus requires the driver to have an understanding of the system's capabilities and limits, that is, a correct mental model of the technical system. Common use cases for Level 2 automation include driving on highways or similar road types (e.g., Volvo Car Corporation, 2019) with some system functionalities being limited further in their area of application by speed restrictions, such as the Traffic Jam Assist (TJA) as a combination of ACC Stop and Go and lane keeping assistance for low speed applications (ERTRAC, 2017; Gasser et al., 2012). Technical limits of Level 2 automation are due to, e.g., sensor range or sensor functioning, but also to the absence of redundancies. Examples include narrow curves, multiple or faded lane markings or adverse weather conditions (Ford Motor Company, 2019; Volvo Car Corporation, 2019).

For Level 2 automation, as defined by Gasser et al. (2012) for the assessment of legal consequences of partial automation, the driver is therefore expected to be available on short notice, or, more precisely, immediately, as a backup to perform the driving task again. This is in opposition to higher levels of automation, granting 'sufficient' time for takeover (Gasser et al., 2012, p. 31) and therefore allowing the driver to withdraw not only physically, but also cognitively from the driving task during automation use. Thus, for Level 2 automation, the driver has to monitor the environment and system behaviour permanently, as if performing the driving task

[1] The term system, as used in this work, is not necessarily restricted to one advanced driver assistance system (cf. SAE, 2018), but also extends to Level 2 automation. The disengagement of one advanced driver assistance system can be accompanied by a change in automation level, e.g., from Level 2 to Level 1.

himself, to ensure immediate readiness to control the vehicle whenever necessary. Accordingly, current owner's manuals provided by the manufacturers stress the driver's responsibility to supervise system and environment when using assistance systems, as a change in automation levels (from Level 2 to Level 1 or Level 2 to Level 0) can occur suddenly. To this end, the driver's hand posture is often monitored during PAD, with both hands required on the steering wheel at all times following the instruction provided in the owner's manual (Ford Motor Company, 2019; Volvo Car Corporation, 2019). A warning is issued when the driver removes the hands from the steering wheel for a defined time. The warning is followed by the disengagement of lateral assistance in case it is disregarded.

Following the definition provided for an exemplary Level 2 system by Gasser et al. (2012), system limits can be either announced by the system on short or long notice or have to be detected by the driver without notification. Encountering system limits is followed by disengagement of the system and subsequent manual driving. Especially the possibility of short-termed notifications on takeover and of unannounced limits imposes high importance on continuous monitoring to enable sufficient levels of SA.

The successful transition of control between system and driver presents a key challenge in interaction with automation (Flemisch et al., 2012), with responsible monitoring behaviour as a prerequisite for takeover readiness. The main requirement for the driver during partial automation is to stay in the loop, that is, to remain involved in the driving task although not actively contributing to its execution. Failure to attribute attention to the environment prevents an accurate situation assessement and reduces the likelihood for a detection of system failures, "thereby retarding the necessary transition between automated and manual operations" (Ward, 2000, p. 400). Next to the ability to "detect anomalies" (Seppelt & Victor, 2016, p. 142), this out-of-the-loop performance problem (OOTL; see, e.g., Endsley, 1995; Endsley & Kiris, 1995) has been connected to the ability to "form correct expectations" and "control the system manually" (Seppelt & Victor, 2016, p. 142). Accordingly, OOTL has been discussed as one of several contributing effects explaining changes in the interaction with automation.

A schematic overview on the driving task being shared between a driver and a partial automated system can be found in Figure 2-1. The driving task is defined by driver, automation (system) and driving context. The term *context* is used in this work instead of the frequently used term *environment* to refer to all external, neither user nor system related factors defining the interaction, including factors not inherent to the current driving situation such as the duration of automation use. System limits, leading to a control transition between driver and system, are either detected and announced by the system (TOR) or detected by the driver. The perception and comprehension of either system limit or TOR is the necessary precondition for the initiation of a takeover by the driver.

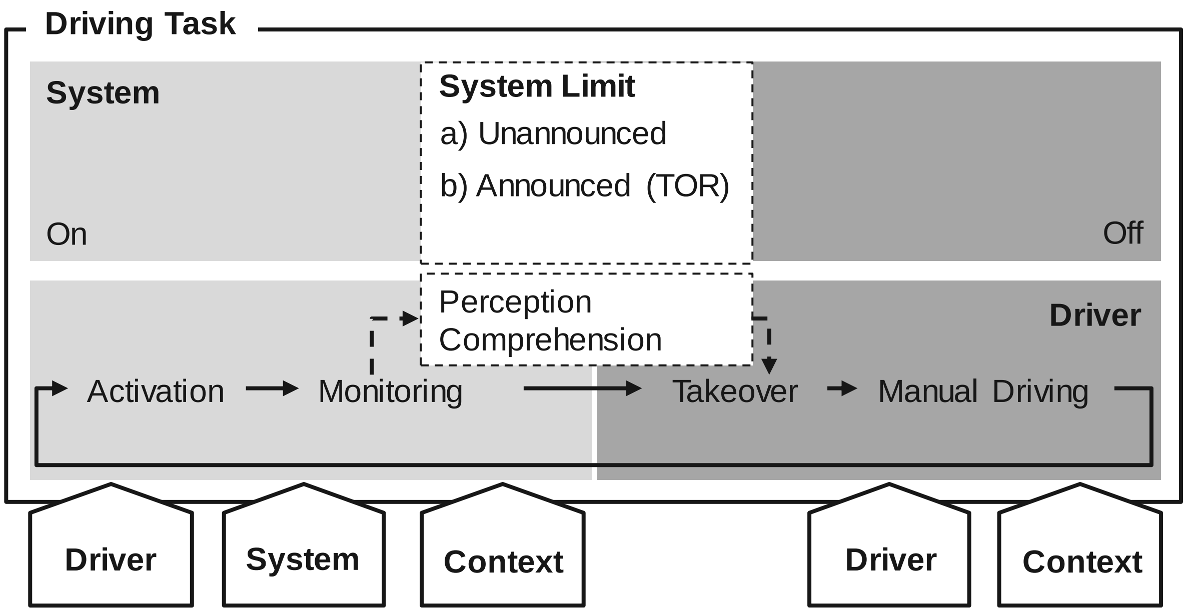

Figure 2-1: Overview on the interaction with partial automated driving (PAD) systems with focus on the driver's tasks. The driving task is defined by characteristics of driver, PAD system and driving context.

In a general overview on replacement of human tasks by automation, Manzey (2008) cited the loss of SA as one of three major problems of interaction with automation next to ill-adjusted trust (see also Hoff & Bashir, 2015; Lee & See, 2004; Parasuraman & Riley, 1997) and the possible loss of skill due to the use of automation. In a literature-based framework on performance inhibitory effects in automation, Ward (2000) distinguished three processes related to monitoring and SA, namely simplification, OOTL and complacency, with negative effects on control transitions. The simplification of the driver's task is expected to decrease his arousal and subsequently SA by means of reduced vigilance. Similarly, OOTL connects lowered SA to the reduction of feedback during passive monitoring. Finally, complacency, following Ward (2000), links SA reductions to a reduced arousal with subsequent lower investment of effort into supervision as the demands of the monitoring task are underestimated, but also to adaptations in behaviour in response to a perceived higher safety when using automation.

Performance and safety in transitions from automated to manual control can be assumed to be strongly related to SA (de Winter, Happee, Martens, & Stanton, 2014). The current level of SA, as detailed in the reviews cited above, can be connected to the driver's workload or the driver's voluntary effort to stay in the loop. As Endsley (1995) pointed out, SA may be low despite low workload, that is, despite sufficient resources, “because of inattentiveness, vigilance problems, or low motivation” (p. 53), indicating thus a “certain degree of independence between SA and workload” (Endsley & Kiris, 1995, p. 384). Additionally, high workload must not immediately result in degraded performance, but might so after a sudden increase in task demand (Parasuraman et al., 2008).

Performance after system-initiated control transitions with notification of the driver (TOR; see Figure 2-1) may be degraded by low SA due to the supposedly hampered re-orientation process (Ward, 2000), with the need for a complete assessment of the situation before taking over control and guidance of the vehicle. Endsley and Kiris (1995) demonstrated decision time to correlate with SA. The self-detection of system failures depends mainly on monitoring performance, with lesser monitoring efforts leading to an incomplete state of knowledge about the current situation (i.e., to lowered SA). Endsley (1995) subsumed that “system operators working with automation have been found to have a diminished ability to detect system errors and subsequently perform tasks manually in the face of automation failures as compared with manual performance on the same task (Billings, 1991; Moray, 1986; Wickens, 1992a; Wiener and Curry, 1980)” (p. 53). In addition, methodological aspects of the study design seem to play a role for monitoring quality. De Winter et al. (2014) have assumed that instructions motivating a thorough visual hazard perception search could be the reason for enhanced SA found in some studies after highly automated driving.

It was stated for unannounced system limits in partial automation that “drivers cannot be completely relied upon to fill in for automation because there may not be enough time for them to act once they notice the automation does not act” (Victor et al., 2018, p. 1113). In addition to SA, the mental model of the driver is thus considered relevant for transitions, especially for the prediction of system limits. When interacting with driver assistance, the term mental model relates to the driver’s understanding of and expectations towards the technical system in general or in the current situation of use (cf. Endsley, 1995; Flemisch et al., 2012; see also Othersen, Petermann-Stock, & Vollrath, 2014, on mode awareness and SA). Following the definition by Rouse and Morris, mental models are “mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future states” (Endsley, 1995, p. 43). SA, describing the driver’s understanding of and expectations related to the current driving situation (Baumann & Krems, 2007), “applies to more rapidly evolving situations” (Parasuraman et al., 2008, p. 145) and is to be distinguished from mental models (Parasuraman et al., 2008; see also Wickens, 2008a). When using automation, drivers do thus not only need to be aware of the current driving situation, enabling predictions of potentially critical driving manoeuvres (SA), but also of the system's functioning (mental model), enabling predictions based on capabilities of the system to handle the current and expected future driving situation (see, e.g., Vollrath & Krems, 2011, p. 205; Othersen et al., 2014). According to Endsley (1995), mental models facilitate the establishment of SA by directing attention to relevant cues, in making correct predictions about future states of the environment and by linking "typical actions" (p. 44) to recognised types of situations. Expertise with the system at hand benefits the mental model and thereby SA (Endsley, 1995). The role of mental models, with a focus on the consequences of the correctness of mental models for the interaction with driver assistance has also been shown in recent studies (Beggiato & Krems, 2013) and is detailed in the following chapters.

As has been reviewed, different frameworks and reviews discuss similar concepts to explain effects of automation on human performance. Frequently discussed is a change in workload associated with automation use, with workload either decreasing (de Winter et al., 2014; Stanton & Young, 1998, 2005), due to lowered active involvement, or increasing, due to heightened monitoring demands (Endsley & Kiris, 1995). Other concepts discussed with regard to automation use include changes in feedback and information processing during monitoring, decreased vigilance, increased secondary task activity and an unwarranted reliance on the system used (see, e.g., de Winter et al., 2014; Endsley & Kiris, 1995; Parasuraman et al., 2008; Vollrath & Krems, 2011; Ward, 2000). In addition, the correctness of the driver's mental model is considered relevant for human-automation-interaction (Beggiato & Krems, 2013; Endsley, 1995; Stanton & Young, 1998). Many of these processes and influences can be hypothesised to culminate in a less accurate assessment of the situation (e.g., Ward, 2000). As a consequence, SA is used frequently as an explanatory concept for automation effects, although the concept has not been undisputed in scientific discourse (see, e.g., Parasuraman et al., 2008, in counter to criticism on the value of SA, workload and trust as explanatory concepts).

It might be important to note at this point that automation does not necessarily go along with adverse effects such as reduced SA, but has also been shown to reduce human error or to increase performance for certain tasks (as pointed out by Endsley & Kiris, 1995). Considering crash risk as the most important safety relevant indicator, Victor et al. (2018), investigating countermeasures against adverse automation effects before and during use of partial automation, reported furthermore that a majority of participants (72 %) was able to successfully handle system limits even without notification. In order to predict under which circumstances problems arise out of automation use and to derive recommendations for future system design, human behaviour has to be examined under controlled variation of potential influencing factors.

In the following, factors that were shown to influence the interaction with automated systems will be reviewed, with a focus on vehicle automation of SAE Levels 1 to 3 (SAE, 2018), but also including the use of technology in other environments. In particular, the focus is on vehicle automation primarily introduced to increase driver comfort by taking over either longitudinal or lateral vehicle guidance (Level 1; ADAS, e.g. ACC or LKA) or by combining longitudinal and lateral assistance for use in more or less defined driving contexts (Level 2; SAE, 2018; PAD, e.g. TJA). Primarily safety-oriented assistance systems, such as forward collision warning, are not in the focus of this overview. The unspecified term *automation* is used when referring to principles of interaction that are not unique to one specific level or type of (vehicle) automation, including other environments for human-automation interaction. The following overview is concerned with four major aspects related to the interaction with automated systems, as defined in Figure 2-1:

1. *Takeover performance* as an indicator for effects of automation use (2.3)
2. *System characteristics* that influence the decision for or against system use and / or the quality of interaction (2.4)
3. *User characteristics* that influence the decision for or against system use and / or the quality of interaction (2.5)
4. *Contexts of use* that influence the interaction with automation (2.6)

2.3 Takeover Performance

As depicted in Figure 2-1 and evident from literature (e.g., Flemisch et al., 2012), transitions in control from automated to manual driving can be regarded as a, if not the most, critical phase of human-automation interaction. Studies "typically look at driving stability or reaction time metrics post-automation to gauge the success of a transfer of control" (Wright et al., 2016, p. 270). The time it takes the driver to execute the first input after a TOR is a frequently referred to metric quantifying behaviour in terms of performance decrements after automated driving (see, e.g., Eriksson & Stanton, 2017; Gold, Damböck, Lorenz, & Bengler, 2013; Zeeb, Buchner, & Schrauf, 2015; for ACC, e.g., Stanton & Young, 1998). A general effect of automation is the prolongation of intervention times, either in comparison to interventions after unassisted driving, with manual driving being considered the baseline level of performance (e.g. Damböck, Farid, Tönert, & Bengler, 2012; Gold, Damböck, Bengler, et al., 2013), or under variations of the automated driving episode (e.g., Eriksson & Stanton, 2017; Merat et al., 2012; Strand, Nilsson, Karlsson, & Nilsson, 2014).

The driver’s response to a TOR can be expected to follow the general process of human response execution, consisting of sensory processing, perception, response selection and response execution (as detailed in, e.g., Wickens, 2002; see also

Schlick, Bruder, & Luczak, 2010, on human information processing; Olson, 1989, on perception response times). The takeover process has been differentiated into single actions, providing the basis for a more detailed assessment of when and why delays arise during transitions. These actions are considered as either consecutive (Damböck, 2013; Eckstein, 2015; see also Petermann-Stock, Hackenberg, Muhr, Josten, & Eckstein, 2015) or as partly parallel (Gold, 2016).

A model which provides the theoretical basis for a classification of the takeover process after automated driving has been introduced by Eckstein (2015; see also Petermann-Stock et al., 2015). It incorporates models relating to the general execution of the driving task and drivers' reactions to critical events in manual driving (see, e.g., Donges, 2009; Hoffmann & Gayko, 2009; Luczak, as cited in Schlick et al., 2010). The model distinguishes four steps towards a complete resumption of control, labelled as t_g (glance), t_c (contact), t_a (action) and t_e (end of correction). The re-direction of visual attention away from potential secondary tasks towards the traffic scene (t_g) to gain relevant cues for action selection is followed by the contact (t_c) with relevant input devices (i.e., pedals, steering wheel). If PAD is used in compliance with current law and rules, the re-orientation towards the road (t_g) should not be necessary, as the execution of the driving task by the system has to be monitored at all times (SAE, 2018). The contact with control devices enables the execution of necessary actions (t_a, i.e., longitudinal or lateral changes undertaken by the driver). After handling the transition of control, the baseline level of driving performance is re-established by the driver (t_e). This last step has been defined as the congruence of the current state with a target state (Petermann-Stock et al., 2015) and provides a link between performance-based and time-based metrics. Morgan, Alford, Williams, Parkhurst and Pipe (2018), who distinguished between the terms takeover (the re-engagement with vehicle controls) and handover (regaining the baseline level of driving) have sought an equivalent inclusion of performance and timing aspects in their data analysis. A similar approach to quantify performance has been described by Neukum and Krüger (2003) for the assessment of driver controllability of steering angle faults instead of control transitions in automated driving. The authors compared the driver's maximum steering wheel angular velocities during undisturbed driving to those after the occurrence of faults, defining the duration of failures by discrepancies between the current and a driver-individual target state.

Gold, Damböck, Lorenz, et al. (2013) measured different driver actions undertaken during the takeover process after HAD with different TOR-limits, that is, time limits between TOR announcement and system limit. Participants were encouraged to engage in a secondary task before the takeover occurred. The authors distinguished between gaze reaction, road fixation, hands-on time, intervention time, side mirror glance and the application of the turn signal. The actions identified largely correspond to those introduced by Eckstein (2015) albeit a slightly higher degree of differentiation, for example between gaze reaction and road fixation, or between steering wheel contact and turn signal application.

Next to the timing of actions, “adverse after-effects” (de Winter et al., 2014, p. 211) or “carryover effects” (Skottke, Debus, Wang, & Huestegge, 2014, p. 1279) of automation on driving performance quality were identified by comparison of driving input or vehicle movements for variable intervals after control transitions (Merat, Jamson, Lai, Daly, & Carsten, 2014; Morgan et al., 2018; Skottke et al., 2014). Zeeb, Buchner and Schrauf (2016), by investigating the effect of different kinds of secondary tasks in conditionally automated driving, showed that analysing driving performance allows for the detection of differences between experimental conditions that are not apparent when analysing the timing of actions only. Although the timing of driver actions for the takeover process did not differ between distraction conditions in their study, differences existed for quality, here defined by the average deviation from the lane centre, in the 10 s after takeover. The authors concluded “that for a comprehensive understanding of driver take-over, both response times and take-over quality must be considered” (Zeeb et al., 2016, p. 239).

Metrics of driving performance are often compared within drivers instead of between groups, using a driver-dependent baseline for comparison (Morgan et al., 2018; Skottke et al., 2014; see also Neukum & Krüger, 2003). For estimating effects of automation, manual driving performance without prior use of automation can be used as the benchmark for driver performance under the assumption that the uninterrupted manual drive represents uninfluenced, optimal driver performance within the current context. In their review, de Winter et al. (2014) stated that "there are some indicators that after seeing or experiencing automated driving, drivers show poorer lane keeping performance, shorter headways, or delayed reaction times as compared to drivers who have not seen/experienced automated driving" (p. 211). However, while late interventions are clearly undesirable as they raise the risk for collisions, the interpretation of differences in driving performance after automation, following de Winter et al. (2014), seems to be less clear-cut and additionally “to be characterised by small effect sizes” (p. 211). Albeit a higher difficulty to define their relevance for driving safety, addressing the quality of driving performance after automation seems important to identify influencing factors when considering the results by Zeeb et al. (2016).

One way to assess the quality of takeover performance is through overall differences in the execution of the driving task in one defined timeframe after takeover, as done by Gold, Körber, Lechner and Bengler (2016), Körber, Gold, Lechner and Bengler (2016), Zeeb et al. (2016) or Eriksson and Stanton (2017). Other analyses did not only approximate existing quality differences during takeover, but also the duration after which performance did no longer differ significantly between experimental conditions, starting with the driver resuming control (Merat et al., 2014; Morgan et al., 2018). In a similar manner, Eckstein (2015), in his model on the takeover process, considered the point where the execution of the driving task is comparable to what it would have been without use and consecutive disengagement of an automated system. The relevant measurement point (i.e., t_e) represents the end of the takeover process, where driving performance has again reached a baseline performance level.

Next to the quality of takeover actions and time elapsed after TOR, a slightly different focus on post-automation driving performance has been the detection of carryover effects in terms of "behavioural adaptations to automation systems" (Skottke et al., 2014, p. 1273). An example is the time headway (THW) to a lead vehicle after automated convoy driving, during which the system kept a small THW. The authors considered "low-level sensory adaptation (i.e., driver might get used to small spacing [during automation use] and subsequently exhibit a tendency to underestimate spatial distance), or […] higher-level learning processes (i.e., driver might experience that small spacing during automation did not yield hazardous events, and subsequently adopt a more risky driving strategy" (Skottke et al., 2014, p. 1274) as reasons for reduced THW found after automated driving. Re-gaining control was explicitly disregarded for this analysis purpose by excluding the first 2 km after automation disengagement from comparisons of THW pre- and post-automation, in difference to studies focussing on the quality or duration of the control transition.

The timeframe and the number of intervals considered for the performance assessment do thus differ between studies and with the purpose of analysis. They range from a single, comparatively short interval for the assessment of takeover quality (e.g., 8 s in Kerschbaum, Lorenz, & Bengler, 2014, or 10 s in Zeeb et al., 2016, both for HAD) to the consideration of manual driving performance for several minutes (e.g. Neubauer, Matthews, Langheim, & Saxby, 2012) or several kilometres (e.g. Skottke et al., 2014, for behavioural adaptations) after automation use. For HAD, steering input has been used to identify the time it takes the driver to resume control, here around 10 s, as well as the process of re-gaining control, "seen by an exaggeration in steering corrections in the next 10-15 s, which then steadied after around 35-40 s" (Merat et al., 2014, p. 280). Merat et al. (2014) found the duration of degraded lateral control after control resumption and the variability in visual attention to be further related to the logic of automation disengagement, being either specified by the duration of use or by driver behaviour.

Measures used in prior research on post-automation driving include the standard deviation of steering angular rate (Eriksson & Stanton, 2017), the standard deviation of steering wheel position (Brookhuis, de Vries, & de Waard, 1991; see Mok et al., 2015, for an application to automation effects), the standard deviation of lane position (SDLP; e.g., Kerschbaum et al., 2014; Merat et al., 2014; Morgan et al., 2018), the average deviation from the lane centre (Zeeb et al., 2016), average speed (Merat et al., 2014; Morgan et al., 2018) and maximum accelerations (longitudinal or lateral; e.g. Körber et al., 2016; Morgan et al., 2018). Overall, metrics considered for the analysis of driver performance are either characteristic for a certain driver, thereby reflecting driver strategy and thus baseline driving performance, or can be connected directly to driving safety. An example for the first type of measure is the mean lateral position (MLP; Knappe, Keinath, & Meinecke, 2006). SDLP, with higher values reflecting “poorer vehicle control” (Neubauer et al., 2014, p. 2055), or the maximum steering wheel angle, “directly influenc[ing] current acceleration” (Kerschbaum et al., 2014, p. 1688), could be considered as examples for the latter category.

Although mostly shown for aspects of driving performance, after-effects of automation can also be shown in the response to events occurring after system use and control transitions, as was done by Neubauer, Matthews and Saxby (2014), who concluded accordingly that “problems may persist following termination of automation” (p. 2057) based on response times after both cruise control (CC) use and HAD. Likewise, Mok et al. (2015) performed a study focussing on “post-transition driving performance” (p. 1167) by variating the time (2 s, 5 s or 8 s) between a TOR, here indicating an immediate system failure, and the negotiation of an imminent road hazard, here road works in a curve. The condition with 2 s was found to result in significantly worse performance in terms of lane offsets and collisions. Additionally, more time spend in the manual driving mode before the critical event resulted in higher comfort of the drivers.

Further, the time available for preparing the transition itself was found to influence takeover performance and the magnitude of differences to manual driving. For partial automation, this effect mainly relates to the visibility and predictability of the need for a control transition, as a sufficient takeover time need not be provided by the system (Gasser et al., 2012). Damböck et al. (2012) found performance after HAD to equal that of manual driving after 6 s to 8 s of preparation time, but not after 4 s, with detected performance differences depending on the quality criteria applied. Equivalently to the results reported by Mok et al. (2015) for post-transition performance, differences were largest in comparison to the shortest takeover limit in driving data and comfort ratings.

For Level 3 or Level 4 automation, allowing the withdrawal from the driving task (SAE, 2018), a main focus of research is therefore how much time needs to be provided for the driver to successfully take back control after notification (e.g., Damböck et al., 2012; Wright et al., 2016). With Level 2 automation, in need of continuous monitoring and the readiness for immediate interventions, the focus of research is less on how much time needs to be provided to the driver, but on the circumstances under which a driver can use the system safely. The aim is to define factors that endanger a successful takeover even under exclusion of misuse of the system[2]. In the following, influencing factors of system, user and context on driving quality and time-based metrics of takeover performance are detailed.

2.4 System Factors Influencing Interaction

The automation level, defining the responsibilities of system and driver as introduced above (SAE, 2018; Gasser et al., 2012), is probably one of the most prominent factors of human-automation interaction. However, within a defined automation level, system reliability, influencing trust in the system, system settings (as discussed in a review by de Winter et al., 2014), as well as the feedback provided to the driver, e.g. by the human machine interface (HMI; Beggiato et al., 2015; Petermann-Stock et al., 2015; Stanton & Young, 2005; Wulf, Zeeb, Rimini-Döring, Arnon, & Gauterin, 2013), influence the safety of interaction between system and driver.

When using partial automation, the driver has to stay in the loop, a requirement that relates mostly, but not only, to the constant visual monitoring of system and environment. As de Winter et al. (2014) pointed out, reviewing partial automated systems, drivers are required to “intermittently touch the steering wheel” (p. 197). Likewise, Othersen (2016) stated that current partial automated systems require at least one hand on the steering wheel during use. This requirement is based on current specifications for the design of such systems (UNECE, 2017b). UN regulations for Advanced Driver Assistance Steering Systems state that hands-on detection must be part of any such system and that an optical hands-on reminder has to be issued latest after 15 s of hands-off driving (UNECE, 2017b). After 30 s of hands-off driving, an additional acoustical warning is issued, followed by system deactivation after additional 30 s of hands-off driving. The maximum duration of driving hands-off following current regulations of the UN is thus one minute. For steering assistance, system status and especially changes in system status need to be signaled to the driver optically and additionally by acoustic or haptic warning (UNECE, 2017b).

---

[2] For a detailed definition of the term misuse, please refer to Parasuraman and Riley (1997). The term is used here to include all types of driver behaviour not warranted by the definition of Level 2 automation.

Accordingly, Gasser et al. (2012) stated that taking the hands off the steering wheel is not necessarily in accordance with the law, even when active input is not required. Rather, as stated by the authors, the lack of an explicit prohibition to drive hands-off does not allow the conclusion of its general legitimacy (Gasser et al., 2012, p. 68). The authors conclude that data is needed on which recommendations to adapt current legal frameworks can be based. Especially critical, in this regard, are short-termed takeover situations that require an active steering input compared to situations that might be resolved solely by longitudinal driver input (Gasser et al., 2012). De Winter et al. (2014) further discussed the influence of defined steering wheel thresholds for system disengagement on the feasibility of hands-on driving, arguing that low thresholds condition the hands-off use of systems to prevent unwanted disengagements by the driver. System implementation for research should keep this requirement in mind for questions of system design.

Apart from legal considerations and the driver still being part of the control loop, removing the hands from the steering wheel can also be relevant when considering the drivers' assessment of the current system status. Feedback has already been discussed as an important factor for SA during automation (e.g., Manzey, 2008) and can be considered as the critical process motivating a restriction of hand posture from the psychological point of view. Endsley and Kiris (1995) reviewed studies on the influence of feedback for different kinds of automated systems, fostering the hypothesis that visual feedback alone does not provide for an equally good detection of system failures than visual feedback enhanced by other senses such as proprioception or smell. However, since being free to remove the hands from the steering wheel is also considered a comfort feature, as suggested by Othersen (2016), and might enhance the users' intention to use automated system, hand posture as an influencing factor on takeover performance has received attention in different studies on vehicle automation. The importance of drivers' attitudes towards PAD should not be underestimated. As, for instance, Brookhuis et al. (2009) stated, “mental workload and acceptance […] are crucial for the implementation success of new systems” (p.1020).

Schaller, Schiehlen and Gradenegger (2008) investigated ACC systems with and without lateral guidance and the option to remove the hands from the steering wheel when driving below 40 kph. Hand posture during monitoring was not investigated as a controlled variable, but instead interpreted as an indicator for the acceptance of lateral guidance. The authors found that hands-off time increased to around 80 % of the time of use when lateral guidance was being provided by the system. No statistical comparison of transitions with different hand postures was reported, but the authors stated that intervention times were not distinctly higher after hands-off driving before the TOR.

Damböck, Weißgerber, Kienle and Bengler (2013) investigated partially automated hands-off driving as a supposedly higher level of automation in comparison to hands-on and assisted driving (ACC) under secondary task conditions (visual-verbal task) in a driving simulator. The driving task was introduced as the primary task in the way that additional engagement with the secondary task was a voluntary decision of the driver. Three different automation failures, two of them requiring a longitudinal, reaction by braking, were implemented with a maximum time contingent of 6 s after the need for a control transition became apparent. The hands-off condition was the only automation condition that resulted in significant intervention time differences in comparison to the manual drives regardless of the takeover scenario. Furthermore, subjective workload was significantly reduced in the hands-off condition, supposedly as the driver “is not only mentally but physically relieved from the driving task” (Damböck et al., 2013, p. 1661). Gaze data could not confirm different levels of fatigue between conditions. However, the manual driving data showed a more focused view on the traffic scene as compared to the automated conditions.

In a driving simulator study, Gold, Damböck, Bengler, et al. (2013) investigated hand posture variations between subjects (16 subjects per groups) in relation to monitoring episodes during system uncertainty in high automation. Drivers, driving hands-off, received either a hands-on or a hand posture unrelated monitoring request 6 s prior to a point of system uncertainty. In two cases, the situation evolved critical 4 s before reaching the system boundary. No significant difference between hand postures was found, although drivers monitoring hands-off reacted 300 ms later than drivers monitoring hands-on. Hand posture did thus not influence intervention times significantly after short episodes of monitoring. Gaze orientation towards the road scene differed however significantly between the hand posture groups with faster reactions in the hands-on group. Furthermore, a slightly higher number of non-responders was found after hands-off monitoring. The effects found were attributed to a supposedly higher urgency of the monitoring request in combination with the request to place the hands on the steering wheel, while the magnitude of differences in intervention times, following the authors, corresponded to expected movement times towards the steering wheel.

In a simulator study by Naujoks, Purucker, Neukum, Wolter and Steiger (2015), two different maximum hands-off durations (10 s versus 120 s) for PAD were compared in critical takeover situations requiring a braking manoeuvre to a vehicle at standstill in the current lane. Hands-off driving was voluntary and thus not experimentally controlled. It was found that nearly all drivers were driving hands-on before the critical situation became apparent and all had contact to the steering wheel one second after that. No difference in intervention times between the two hand posture conditions was found.

In her doctoral dissertation, Othersen (2016) investigated the regulation of hand posture (12 participants in each hand posture group) next to other strategies to keep drivers in the loop while using a PAD system on a German highway. The other strategy was the selection of driving manoeuvres either by the driver or the system, thereby actively engaging the driver on a cognitive level. Hands-on driving was realised with at least one hand on the steering wheel. Two types of uncritical takeovers were investigated, a planned takeover at an instructed highway exit and an unplanned takeover. The unplanned takeover simulated a system limit in the own lane, leaving 3 s for the driver to regain control before an emergency stop was initiated by the system. Apart from the emergency stop, no critical situation emerged from late driver interventions. As many drivers deactivated the system before encountering the planned takeover, behavioural data on hand posture was analysed only for a sub-sample in the unplanned situation. Drivers in the hands-on group intervened on average about one second faster. This difference was attributed to longer movement times and OOTL performance problems. The author further found lesser gazes to driving relevant areas when driving hands-off, although most comparisons, apart from the number of gazes to the right side-mirror, were significant by tendency only. Taking the hands off the steering wheel was the option preferred by participants in a post-hoc comparison questionnaire. This finding is in line with the significantly higher rating of comfort and lower workload for the hands-off system, although hands-on driving increased trust in the system. Thus, albeit the small sample, being further reduced by early deactivations, hand posture was proven to be a relevant factor for the subjective and objective evaluation of automated systems.

Some studies on PAD have suggested that the process of reaching for the steering wheel is independent of the level of engagement with the driving task, whereas takeover time is not (Lorenz & Hergeth, 2015). In the same line of argument, context variations with no influence on executional aspects of the takeover process (i.e., establishing motor readiness) were found to influence takeover time after HAD (Gold et al., 2016). Finding similar results for takeover after acoustic TOR in HAD under variation of driver distraction, Zeeb et al. (2015) concluded that “there seems to be no influence of visual driver distraction on the time at which the driver establishes motor readiness to take over the vehicle" (p. 220). With regard to the studies cited above which found effects of hands-off in comparison to hands-on monitoring, the question arises whether effects of manual involvement are attributable only to the process of establishing motor readiness, but independent of other factors influencing monitoring performance.

Whereas the haptic feedback from the steering wheel might be important for maintaining SA, the visual feedback from steering wheel movements is likely not often used by drivers, at least not in HAD (Kerschbaum et al., 2014). In general, however, visual feedback in automation, in contrast to haptic feedback, is deemed important for different contexts due to its relation to SA (Manzey, 2008). As Parasuraman & Riley (1997) stated, “the more removed the operator is from the process, the more th[e] feedback must compensate for this lack of involvement” (p. 248). The HMI of a system provides a means to display relevant information regarding system availability, current system status and the need for driver intervention by TOR. Information relating to the supervision of the system has been found to be essential both to HMI experts and system users regarding PAD and HAD (Beggiato et al., 2015). Several studies have thus engaged in the task to design HMI concepts for automated driving, supporting the driver’s mode awareness or to encourage the driver in maintaining a continuously high level of monitoring activity (e.g. Beller, Heesen, & Vollrath, 2013; Othersen, 2016; Petermann-Stock et al., 2015; Victor et al., 2018).

Louw et al. (2015) argued for a differentiation between physical and cognitive components of disengagement from the driving task, that is, of being OOTL, during use of Level 2 automation. The authors proposed that being OOTL should be connected to impaired performance after automation use. To investigate the role of cognitive disengagement, they investigated different levels of deprivation from visual information during automation use. Screen manipulations in their driving simulator study (light and heavy fog) prevented drivers from monitoring their environment during HAD. Although tendencies for higher disengagement with less visual information were found, no significant effects of the manipulation on takeover quality for a lead vehicle braking manoeuvre could be shown. However, general adverse effects of automation use for responding to critical events in terms of driving performance were found. Instead of a TOR, automation uncertainty with the request to monitor, similar to Gold, Damböck, Bengler, et al. (2013), was investigated, with difficulties resulting for the assessment of response timing.

Furthermore, the system settings chosen when interacting with automation have been in the focus of research. It was found that drivers sometimes choose more critical settings regarding speed and headway when using ACC or automation than when driving manually (Hoedemaeker & Brookhuis, 1998). Other studies on ACC did however find a reduction in driving speed and no increase in secondary task engagement, arguing against an adoption of riskier behaviour while being relieved of driving tasks (Vollrath, Schleicher, & Gelau, 2011). Differences in ACC settings between groups of users have been assumed to arise out of uncertainties in the mental model of drivers (Dickie & Boyle, 2009). Setting preferences were found to vary with user characteristics and available system settings such as headway or speed have thus been found to influence the subjective evaluation of the system (as reviewed by de Winter et al., 2014). Personality factors, specifically sensation seeking, have been related to the degradation of lateral driving performance when using longitudinal driver assistance (Rudin-Brown, Parker, & Malisia, 2003). Changes to the design and interaction logic of a system should thus be analysed not only regarding the safety of interaction, but also regarding the subjective assessment and differences for groups of users.

2.5 Individual Differences in Interaction with Automation

The analysis of user characteristics provides a means to explain variance in the interaction with automation. User traits and characteristics have been found to influence the opinion on future automated vehicles (Kyriakidis, Happee, & de Winter, 2015), influencing the likelihood of their actual use as well as the actual interaction with available ADAS (Rudin-Brown et al., 2003; see also Venkatesh & Morris, 2000, for findings on general technology acceptance). In their review on both ACC and HAD systems, de Winter et al. (2014) discussed experience with a system, age, gender and current driver state, such as fatigue, as variables moderating the interaction quality. Similarly, Körber and Bengler (2014) reviewed dispositions, traits, states, attitudes and demographics with potential relevance for human-automation interaction. The authors concluded that, although some factors have been studied extensively in the driving context, such as age, there is further need for research to define the interrelationship between driver characteristics and their influence on automation use. In the IT context, research has established beliefs and attitudes "as the key determinants of both initial […] usage (acceptance) and long-term usage (continuance) intention and behaviour (Bhattacherjee 2001; Davis et al. 1989)" (Bhattacherjee & Premkumar, 2004, p. 230).

Trust has frequently been considered as an important attitude when comparing different user groups for the automation of work environments, driver assistance or vehicle automation (e.g. Gold, Körber, Hohenberger, Lechner, & Bengler, 2015; Hoff & Bashir, 2015; Lee & See, 2004; Muir & Moray, 1996; Parasuraman et al., 2008). The reason for this is its relation to the actual use of the system under investigation, as was found by Muir & Moray (1996) who stated that "operators used automation they trusted and rejected automation they distrusted, preferring to do the control task manually" (p. 429). Lee and See (2004) considered the appropriate use of a system to be the result of a successful matching between the operator's level of trust and the system's capabilities (i.e., its trustworthiness). This so-called *calibrated trust* is to be differentiated from *overtrust*, a notion describing the overreliance on a system that is not warranted by its technical capabilities, and *distrust*, that is, an underestimation of the system's capabilities, leading to disuse (Lee & See, 2004). "Trust stands between beliefs about the characteristics of the automation and the intention to rely on the automation" (Lee & See, 2004, p. 54). It "guides reliance when complexity and unanticipated situations make a complete understanding of the automation impractical" (p. 50).

Changes in the level of trust have been observed between different phases of interaction, that is, between ratings based on information and ratings based on experiences made with a system, although not all authors found trust to change considerably with experience (e.g. Gold et al., 2015; Feldhütter, Gold, Hüger, & Bengler, 2016; Muir & Moray, 1996; see also Hoff & Bashir, 2015). Rudin-Brown et al. (2003) found trust to increase after using ACC in a test track study, regardless of ACC failures. Hoff and Bashir (2015) differentiated between trust before use of automation, influencing the initial reliance strategy, and dynamic learned trust during interaction, influencing reliance on the system. In their model, trust before use of automation was further segmented into dispositional, situational and initial learned trust and hypothetically influenced, amongst others, by pre-existing knowledge, age, gender and personality variables (Hoff & Bashir, 2015).

Beggiato and Krems (2013) manipulated the initial mental model of ACC users by either presenting correct information on system limits, omitting potential problems, thus idealising the system's capabilities, or presenting correct as well as non-occurring problems. To restrict the influence of other driver characteristics on results, the instruction groups were matched regarding, amongst others, age, driving experience, ADAS experience and gender. Omitted system limits in the initial mental model were found to decrease trust and acceptance during continued system use, whereas initially established system limits did not negatively influence drivers' attitudes. Experience with a system was found to update the initial, theoretically based mental model towards a correct system model. Beggiato and Krems (2013) concluded that "studies of system trust and acceptance should include, and attempt to control, users' initial mental model of system functionality" (p. 56).

As Beggiato and Krems (2013) showed, trust and acceptance of ACC are strongly related to the correctness of the (initial) mental model (see also Feldhütter et al., 2016, for HAD), that is, to the match of mental model and experienced system capabilities. This indicates a need to investigate changes in user expectations when introducing changes in system capabilities. Beggiato et al. (2015) further found trust to be related to the amount of feedback requested from the system as well as to monitoring behaviour, with less informational demand and higher secondary task activity related to higher levels of trust in a system. Muir and Moray (1996) found a similar effect of increased monitoring of automated systems with low levels of trust.

Bhattacherjee and Premkumar (2004) distinguished between a "pre-usage stage" and a "usage stage" (p. 234) when investigating user beliefs and attitudes in the IT context and their change over time, similar to the model of trust in automation proposed by Hoff and Bashir (2015). Experience with a system serves to calibrate the user's mental model and trust into the system (e.g. Beggiato & Krems, 2013; Gold et al., 2015; Hoff & Bashir, 2015). Initially, the mental model is based on information, that is, reading a manual (Flemisch et al., 2012; Beggiato & Krems, 2013) or media consumption (Feldhütter et al., 2016), or on expectations based on experience with (similar) systems (e.g. ACC experience; Zeeb et al., 2016). A correct mental model is needed to predict system limits or system behaviour and most studies on user-automation interaction thus include a detailed system instruction to ensure at least a basic understanding of the system under evaluation (e.g. Beggiato & Krems, 2013; Rudin-Brown et al., 2003; Vollrath et al., 2011). However, a survey amongst ACC users by Dickie and Boyle (2009) revealed that, although prolonged use is beneficial for awareness of system limitations, even users frequently interacting with a system might be unaware regarding limitations of the system with adverse effects on the use of such systems. The awareness of limitations is related to the appropriateness of trust in the system, as information and beliefs provide the basis for trust (Hoff & Bashir, 2015; Lee & See, 2004).

Overall, experience is considered an important factor for the quality of system interaction, especially for performance in takeover situations. In a field operational test on ACC use, Weinberger, Winner and Bubb (2001) analysed the learning phase of users with a new system. They reported significant changes in behaviour, specifically in brake application, during the first two weeks of use as well as a learning process of two to three weeks until users reported to have had a sufficient mental model of takeover situations. Kopf and Simon (as cited in Saad, 2007) distinguished three learning stages of interaction, namely operating the system, internalising system limits and learning to use the system appropriately in the situational context. Vollrath et al. (2011) discussed the necessity to include experienced users in analyses to account for changed interaction behaviour with increased experience.

As longitudinal studies on system use are rare due to the high effort related to such studies, effects of increasing experience with a system on attitudes and interaction quality are mostly analysed in form of learning effects after repeated takeover encounters during experiments, as in Happee, Gold, Radlmayer, Hergeth and Bengler (2017) or in Körber et al. (2016). Körber et al. (2016) found significant improvements in takeover performance after HAD with repeated exposure, which were independent of driver age. Payre, Cestac and Delhomme (2016) found no effect of a larger number of practised situations on takeover time, but later reactions for higher levels of trust. Zeeb et al. (2016) found learning effects in response times for conditional automation, with faster response times after repeated takeover, only for those drivers who had never used ACC before, but not for regular ACC users. Thus, not only experience with the system currently under investigation, but also prior experience with similar ADAS influences takeover performance. “Regardless of its specific effect, past experience almost always plays a role in guiding human-automation interaction” (Hoff & Bashir, 2015, p. 421).

Apart from prior ADAS experience, driving experience in general has been shown to benefit vehicle control (Duncan, Williams, & Brown, 1991) and hazard perception (see, e.g., a review by Horswill & McKenna, 2004). It could thus be considered as a relevant influencing factor on takeover performance (as also hypothesized by Strand et al., 2014). For hazard perception, driving expertise has been attributed to a higher accuracy of the experts’ mental model for identifying hazards, although high effort is still required for good performance (Horswill & McKenna, 2004). Latent hazard perception, as an indicator of SA, has been used to quantify the effect of driving experience on control transfers in high automation (Wright et al., 2016). Whereas no differences between middle-aged, experienced drivers and young, less experienced drivers were found in manual driving, the scanning behaviour for hazards differed significantly in conditions with 6 s TOR time, with more hazards detected by experienced drivers. Experience in this study was quantified in years of driving experience.

Generally, caution has to be exerted to avoid confounding effects of (higher) age with those of cumulated driving experience, as discussed in Duncan et al. (1991). In general, high age is often considered as an influencing variable on the takeover process and on attitudes towards automation. The former can be attributed to the fact that driving performance in time-critical or complex situations was shown to decline with high age (see, e.g., Vollrath & Krems, 2011, for an overview on age effects on driving performance; see Olson, 1989, for a review on age and perception-response times), questioning the feasibility of short-termed takeover for according groups. Warshawsky-Livne and Shinar (2002) found perception-reaction time in brake reactions to increase with age when comparing young, middle-aged and senior drivers, whereas brake-movement time did not vary. Körber et al. (2016) compared the takeover performance after HAD between two age groups with drivers below 28 years and drivers above 60 years. It was found that older drivers are not slower to react to TOR when left with 7 s time for a response, even under increased traffic complexity. However, drivers of higher age braked more often, indicating a change in strategy when handling a takeover situation. Petermann-Stock et al. (2015) found that after HAD with secondary task involvement, older drivers between 50 and 70 years reacted similarly fast as younger drivers between 25 and 35 years. Albeit the overall similar intervention times, differences were found regarding the time between the different steps of the takeover process, with older drivers taking more time to establish contact with input devices than younger drivers, who took more time to intervene after early establishing the readiness to respond. However, the effect of age on driving performance is mediated by overall driving experience and compensatory driving strategies applied (Vollrath & Krems, 2011; see also Körber et al., 2016). Additionally, in the context of automation, higher age was found to positively influence the intention to use and the trust in automation (Gold et al., 2015) and might thus also influence the frequency of interaction with automation.

Another frequently considered demographic factor is gender. Following a review by de Winter et al. (2014), gender influences the settings preferred when interacting with ADAS and the frequency and types of secondary task engagement. Payre, Cestac and Delhomme (2014) conducted a survey on fully automated driving and found a significant influence of gender on the intention to buy and on attitudes towards automated driving. In a driving simulator study on fully automated driving, the same authors found higher levels of trust for men than for women, but no difference in takeover timing (Payre et al., 2016). The same effect was found by Feldhütter et al. (2016) for HAD. Higher male ratings in trust and intention to use in this study were however only found before experiencing automation, but not after gathering experience with the system. Gender might thus be considered as a mediating variable regarding the affinity towards new technology or the acceptance of specific settings especially in a priori contexts. Other studies (Warshawsky-Livne & Shinar, 2002) have shown gender to influence brake-movement time (i.e., movement from accelerator to brake), but not perception-reaction time to the brake lights of a lead vehicle. This was attributed to the potentially lower driving experience of females in distance driven albeit similarly long driving experience in years (Warshawsky-Livne & Shinar, 2002).

2.6 Automation Use in Different Contexts

As has been detailed before, user, system and environment are closely interrelated, as, for example, expressed by the notion of calibrated trust that defines appropriate use and depends on the match between expected and actual capabilities of the system within the current situation of use (Lee & See, 2004). This chapter provides an overview on relevant situational variables, that is, the context of use, and their influence on human-automation interaction, specifically on monitoring behaviour and takeover performance. Variables defining the complexity of the situation and thus the demand for monitoring the correct execution of the driving task and for maintaining (Bolstad, 2001) and regaining SA as well as variables influencing the driver's state and monitoring strategy need to be considered.

The complexity of the driving situation has been varied by means of different traffic densities. Gold et al. (2016) compared three conditions, namely no surrounding traffic, ten and 20 vehicles per kilometre, with the expectation that high traffic densities have an adverse effect on takeover performance after HAD as “the effort to regaining situation awareness is dependent on the number of relevant vehicles to be considered” (p. 643). Study results confirmed longer reaction times and higher accelerations as well as a higher crash probability with higher traffic densities, although differences were only found in comparison to the situation with no surrounding traffic. As motoric aspects, in this case, the time it took to place the hands on the steering wheel, were not influenced by traffic density, the authors concluded that these processes “can be executed without extensive situational assessment or understanding” (Gold et al., 2016, p. 648). The authors additionally discussed strategy changes in takeover behaviour, as lane changes, for example, are generally facilitated by lower numbers of surrounding vehicles. Körber et al. (2016) found, consistently for both of the age groups considered, a significant increase in takeover time for medium and high traffic densities compared to no surrounding traffic. Congruently, Radlmayer, Gold, Lorenz, Farid and Bengler (2014) reported an adverse effect of dense traffic (30 vehicles per kilometre) on takeover quality after HAD, especially in combination with secondary tasks.

Regarding secondary task uptake and eye-movements under different traffic conditions during HAD, Jamson, Merat, Carsten and Lai (2013) showed that more visual attention was attributed to the roadway in dense traffic. No effect of traffic density on gaze data was present in manual drives. Overall, more visual attention was attributed to the roadway in manual driving than during automation use. Congruently with results on traffic density, Beggiato et al. (2015) found a higher number of control glances and less secondary task engagement in HAD in complex driving scenarios such as construction zones. These results suggest that drivers adapt their monitoring effort to the traffic situation even in higher automation levels where continuous monitoring is not a driver responsibility.

The duration of the automated drive is another important variable with its effect mostly attributed to a decrease in vigilance or a decreasing motivation to monitor a correctly working system over prolonged time of use (e.g. de Winter et al., 2014; Endsley, 1995; Thiffault & Bergeron, 2003). “Vigilance refers to the ability of organisms to maintain their focus of attention and to remain alert to stimuli over prolonged periods of time (Davies & Parasuraman, 1982; Parasuraman, 1986; Warm, 1984a, 1993)” (Warm, Parasuraman, & Matthews, 2008, p. 433). It has also been referred to as “the ability to maintain sustained attention within the road environment” (Thiffault & Bergeron, 2003, p. 381) and is generally understood to decrease under monotonous conditions, with monotony describing either a low variability or, more specifically, a high predictability of visual input (as reviewed by Thiffault & Bergeron, 2003). A typical reaction to monotony includes the “loss of interest of performing the task at hand” (Thiffault & Bergeron, 2003, p. 383). “The quintessential finding in vigilance research is that detection performance declines over time, a result known as the *vigilance decrement*" (Warm et al., 2008, p. 434) and of high relevance for the detection of automation limits during PAD. Further, a decrease in vigilance has been connected to slower responses in critical driving situations (de Winter et al., 2014), similar to effects found in monotonous manual driving contexts also associated with a decrease in vigilance (Schmidt et al., 2007). Performing a vigilance task, such as monitoring for automation failures, has additionally been shown to increase workload, thereby countering one considered benefit of automation (de Winter et al., 2014; Warm et al., 2008).

Even short periods of driving (around 10 minutes) with a highly automated TJA have been found to decrease vigilance and induce OOTL performance problems (Dogan, Deborne, Delhomme, Kemeny, & Jonville, 2014). In general, depending on task demand, decrements in vigilance tasks have been observed after 5 to 15 minutes (as reviewed by Warm et al., 2008). Investigating the effect of monitoring duration in PAD, Othersen et al. (2014) measured the detection rate of takeover situations after 5, 15 and 25 minutes of supervision. The authors found an increase in secondary task engagement and thus attention averted from the road as well as a higher number of non-responders to a braking lead vehicle after prolonged supervision, with the largest effect after 15 minutes of automated driving. Focusing on takeover performance instead of self-detected takeover need, Feldhütter, Gold, Schneider and Bengler (2017) found no differences between transitions after 5 and 20 minutes of HAD. However, driving duration influenced gaze behaviour as glances towards the road scene became shorter, but more frequent. A group being offered a highly visually distracting task during HAD to reduce underload, as drivers do not need to monitor automation during HAD in opposition to PAD, took longer to avert the eyes from the secondary task in case of a TOR after longer durations of automated driving.

A related and frequently voiced concern is an overall increase in secondary task activity, which can be considered adverse to safety albeit the passive driver role because reduced monitoring lowers SA (e.g., Wickens, 2008a; Wulf et al., 2013). Secondary task performance was found to deteriorate the quality of reactions to critical events after HAD in comparison to manual driving, specifically the reduction of speed (Merat et al., 2012). Indeed, drivers have been found to perform more secondary tasks when using automation (Jamson et al., 2013; Rudin-Brown et al., 2003; Neubauer, Matthews, & Saxby, 2012) and the number of tasks completed was found to increase with higher levels of automation, that is for manual driving compared to ACC and HAD (as reviewed by de Winter et al., 2014). Rudin-Brown et al. (2003) found drivers to complete more secondary, visual search tasks when using ACC, although the safety of the primary driving task was stressed as a main goal and no reward was connected to secondary task performance. However, whereas secondary task engagement was reported to increase during longer episodes of use, it was found to decrease with the complexity of the driving situation, similarly to increased supervision observed in dense traffic. Othersen et al. (2014) found differences in secondary task activity, quantified in terms of gaze behaviour, to manual driving only for monotonous driving situations. Eriksson and Stanton (2017), finding longer takeover times with secondary task engagement, suggested to provide more time for takeover in HAD depending on driver state. Zeeb et al. (2016), however, investigating conditional automation, did find only little to no influence of secondary task engagement on the time it took drivers to return their hands on the steering wheel and to intervene. In their study, distraction was found to impair takeover quality instead.

Next to driver strategies related to situational complexity, the type of secondary task used has an influence on the size of effects found. De Winter et al. (2014) discussed the influence of modality for secondary task effects, suggesting that a non-visual task can be timeshared with the driving task, while a mainly visual task cannot. Following driver distraction research, cognitive and visual distraction can indeed be assumed to have different effects on gaze behaviour, with “an increased concentration of gaze towards the forward roadway” found for cognitive load (as reviewed in Engström, Markkula, Victor, & Merat, 2017, p. 746). Cognitive tasks can even enhance reaction times by countering effects of monotony during automation (Neubauer, Matthews, & Saxby, 2012). The effect of task modality for interference with the monitoring task was proven important for PAD by Lorenz and Hergeth (2015), whereas Radlmayer et al. (2014) did find no differential effects between visual and cognitive tasks for HAD. For PAD, in difference to higher levels of automation, involvement with secondary tasks other than those compliant with manual driving has to be considered as adverse, as the driver is responsible for the continuous supervision of system and situation. Results suggest, however, that drivers might be more inclined to engage in secondary tasks when using automation and that this effect increases with higher levels of automation, longer duration of use and higher monotony.

The anticipation of an upcoming transition in control (i.e., a system limit), is considered helpful for takeover quality (Dogan et al., 2014). Anticipation is however related to the driver’s monitoring behaviour in the current driving situation, that is, the opportunity to perceive relevant elements, and to the driver’s mental model, to (correctly) interpret the elements perceived and predict the system’s behaviour. In general, longer time intervals, either by earlier announcement of TOR or by anticipation of system limits before TOR, are seen as beneficial as they give the driver the opportunity for a prepared, more controlled takeover, reducing the occurrence of critical situations (e.g. Damböck et al., 2012). Dogan et al. (2014) investigated the benefit of anticipatable takeover situations. They found a (non-significant) trend for longer takeover times after surprising TOR compared to situations where cues should have prepared the driver for the upcoming control transition, such as leaving the system's functional range at the end of a traffic jam. However, HAD driving with secondary task engagement was investigated in this study, constituting different conditions for an anticipation of system limits as in PAD.

Not only should the context of testing, meaning the characteristics of the driving scenario, be taken into account to assess the effects of automation, but also its operationalisation and the overall methodology used. Saad (2007) suggested that in order to explain variance in adaptation processes found across different studies, "the context in which the studies have been carried out (driving simulator, closed tracks or real driving scenarios) should be specified as well as the various scenarios and driving tasks in which the behavioural changes have been identified" (p.152). Frequently discussed in this context, not only when investigating automation effects, is the validity of driving simulation results for real-world driving applications. Stanton and Young (1998) stated that "physical fidelity may help to convince the experimental participant that the task should be taken seriously, which would be less convincing in a more abstract environment" (p. 1018). In general, the validity of results from different methodologies depends on the metrics used and conclusions drawn, as in the difference between absolute and relative comparisons between conditions (Kaptein, Theeuwes, & van der Horst, 1996; for an overview, see also Abendroth, Schreiber, Bruder, Maul, & Maul, 2012).

2.7 Summary on Partial Automation from a User Perspective

A guiding principle when introducing automation into a human performed task is often an assumed benefit in safety and reliability of task performance (Manzey, 2008). In the area of driving, being a complex, cognitively and visually demanding task (Abendroth & Bruder, 2009; Vollrath & Krems, 2011; Ward, 2000), an additional goal is the increase in comfort by relieving the driver of the active conduction of the driving task (Stanton & Marsden, 1996). The responsibilities assigned to driver and system vary with the level of automation, with higher levels of automation placing lesser restrictions on the user's behaviour during automation use. However, automation is also associated with risks, with the three central problem areas of interaction, according to Manzey (2008), being ill-adjusted trust into automation (see also Lee & See, 2004), loss of SA and loss of skill due to the use of automation. Whereas the latter area seems initially less relevant for PAD based on the likely limitation of the use cases for this automation level (Gasser et al., 2012), the other two areas need be considered for a safe interaction with partial automation.

Current Level 2 systems (SAE, 2018) require the driver to act as a supervisor of the system. Although the driver does not need to actively perform the driving task, he needs to retain a sufficiently high level of SA nonetheless, similar to manual driving. To achieve good SA, information from the environment and the system needs to be perceived and aggregated to establish an understanding of the current driving situation, on which a correct estimation of the future status of elements can be based (Endsley, 1995). This definition of SA mirrors the importance of visual and cognitive resources for driving. The use of partial automation changes little about the demands associated with the continuous supervision of system and surroundings and therefore the importance of SA. Deficits in establishing continuously good SA during PAD become visible primarily in case of takeover situations (Manzey, 2008).

Takeover situations, meaning control transfers from automated to manual driving, are often used to quantify the effects of automation use. A general disadvantage of automation in comparison to manual driving has been shown in the form of later responses to (sudden) critical events (e.g., Damböck et al., 2012; Gold, Damböck, Bengler, et al., 2013). Variability in takeover performance, defined by quality and timing, has been associated with different characteristics of driver, system and context of use (see, e.g., de Winter et al., 2014, for a review). Additionally, quality has been considered in terms of manoeuvre success or operationalised by non-binary surrogate reference metrics such as the variation in lane position (e.g., Happee et al., 2017). The questions of whether and for how long after takeover different factors and degrees of monitoring influence driving performance, as indicated as the end of correction in the model by Eckstein (2015), have been considered to a lesser degree for PAD use. In general, persisting effects of automation use have however been shown in lateral as well as longitudinal control (Merat et al., 2014; Morgan et al., 2018).

Although the manual performance of the driving task is not required during PAD, drivers currently need to keep their hands on the steering wheel (Gasser et al., 2012) more or less permanently (Ford Motor Company, 2019; UNECE, 2017b; Volvo Car Corporation, 2019). As hands-off monitoring is associated with a reduced subjective workload and increased comfort (Damböck et al., 2013; Othersen, 2016), the necessity of this precaution should be evaluated. In 2012, Gasser et al. called for further evaluation of this matter and several studies have since delivered input on this subject (Damböck et al., 2013; Naujoks et al., 2015; Othersen, 2016). Movement time seems to be a factor in comparison to hands-on supervision (Gold, Damböck, Bengler, et al., 2013). Still missing to pinpoint the role of haptic feedback on monitoring and takeover performance is however more data under variable conditions of interaction. Especially critical cases in this regard are short-termed control transfers requiring steering input (Gasser et al., 2012).

The context of automation use has mainly been considered in terms of complexity and monitoring demand, that is, for variables which were shown to influence the driver’s monitoring strategies, their SA and takeover times, such as traffic density (Gold et al., 2016; Lu, Coster, & de Winter, 2017). Monitoring quality has further been associated with vigilance effects, leading to the question of whether and when performance degrades after prolonged use of automation (Feldhütter et al., 2017; Othersen et al., 2014). Another issue frequently targeted is an assumed increase in secondary task involvement during monitoring in automated driving (Jamson et al., 2013; Rudin-Brown et al., 2003), indicating a reduced effort to stay involved in the driving task.

Further evident from literature is that any evaluation of the interaction with different automated systems needs to take into account the user, both as an influencing factor on interaction quality and for designing accepted and useful technology. De Winter et al. (2014), reviewing studies on ACC and HAD, concluded that “individual differences in age, gender, and driving experience may explain a large share of the variance of the observed workload and situation awareness scores” (p. 210). Similarly, Strand et al. (2014) argued for an inclusion of automation experience as well as general driving experience and age in studies. Lee and See’s conceptual model of trust (2004) further states the importance of information and experience on beliefs as they form the basis of attitudes, in this case with trust being an attitude translating into automation reliance as the consecutive behaviour.

Characteristics of the user, the system and the context of use thus need to be taken into account to ensure that drivers can interact safely with specific partially automated systems within relevant contexts of use. Especially factors targeting the level of SA established and maintained by the driver seem important in this regard. Saad (2007) described two major concerns for introducing new technology into the market, namely “drivers’ opinions and their acceptance” (p.150) and behavioural adaptations. Both subjective and behavioural changes as well as physiological metrics, particularly gaze behaviour, should be addressed for assessing the interaction with PAD. Research questions derived from the current state of the art and a framework for addressing these questions are detailed in the following chapter.

3 Research Framework

Interaction quality and especially the success of control transitions between driver and automation are influenced by characteristics of user, system and context (see Figure 2-1). As Zeeb et al. (2015), after reviewing the literature, concluded equivalently, “take-over time is, within limits, specific for a particular set of situation variables (e.g., traffic density, action alternatives, HMI concepts, secondary task type, level of driver distraction) and driver variables (e.g., age and skill of the driver)” (p. 213). Performance problems observed in interaction with PAD systems or with automation in general have often been attributed to degraded SA (de Winter et al., 2014; Endsley & Kiris, 1995; Manzey, 2008; Wickens, 2008a). For instance, Endsley and Kiris (1995) defined degraded SA as the major contributor to OOTL performance decrements in decision time. In particular, SA is thought to deteriorate due to four major automation related problems: *insufficient monitoring* of the system, limited or changed *feedback* channels, lack of *system transparency* or insufficient *system understanding* (Manzey, 2008, who based his overview on Endsley, Bolté, & Jones; see also Endsley & Kiris, 1995, for a similar classification). Due to good SA being an important precondition for good performance (Wickens, 2008a), factors contributing to these problems in interaction with PAD systems need to be addressed by research. Factors connected to SA should be considered for their influence on the interaction, the acceptance and intended use of PAD. For example, expertise guides attention, thereby influencing perception, and shapes the mental model of users, which in turn influences prediction (as reviewed by Wickens, 2008a) as well as the inclination to rely on technology (e.g., Beggiato & Krems, 2013).

Most experimental factors targeted in this work were therefore selected based on their assumed relation to achieving or maintaining sufficient SA by influencing either the monitoring strategy of users or the feedback received during automation use. The change in performance quality and safety when interacting with PAD systems under presumably adverse conditions is in the focus of this work, with the aim to identify potential conditions under which PAD cannot be safely used. The reasoning behind the factor selection for the investigation of interaction behaviour is therefore similar to that behind implicit SA measures, where “the loss of SA can be inferred from changes in performance on tasks for which good SA is essential” (Wickens, 2008a, p. 399). Additionally, changes in the attitudes towards automation are considered. An overview on the influencing factors investigated can be found in Figure 3-1.

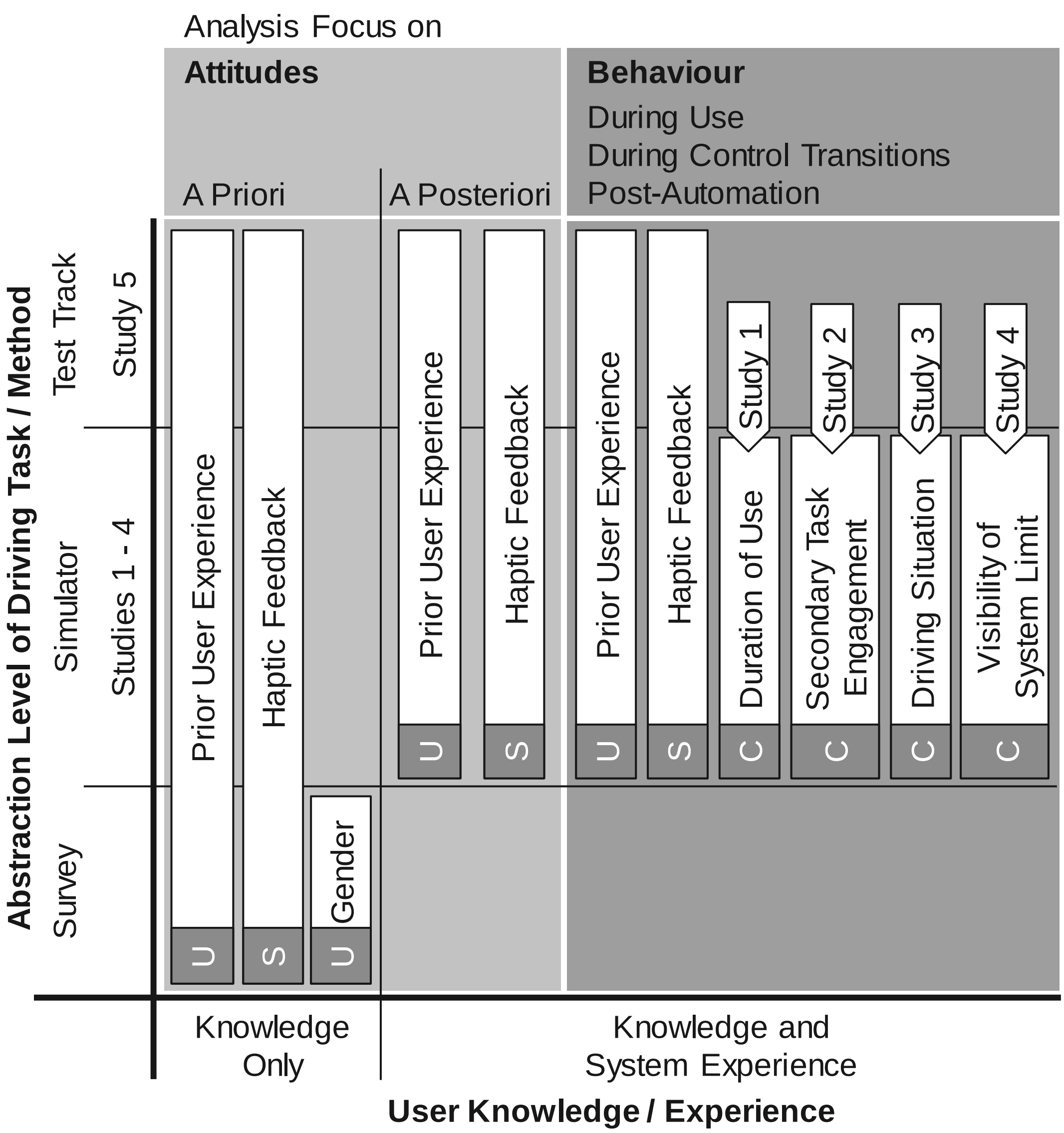

Figure 3-1: Research framework with experimental factors of user (U), PAD system (S) and context of use (C). Analysis is guided by two superordinate axes of comparison, *user knowledge / experience* (x-axis) and the *method of investigation* (y-axis). Two analysis levels, *attitudes* and *behaviour*, provide insights into human-automation interaction.

Of the four automation-related problems, the *limitation of feedback*, specifically of haptic feedback, is in the focus of the current work. The physical contact with the steering wheel provides an additional level of feedback from the system (haptic feedback; Damböck et al., 2013), with feedback being considered beneficial to SA (Manzey, 2008; see also Endsley & Kiris, 1995; Parasuraman & Riley, 1997). Inconclusive results regarding the significance of haptic feedback on takeover performance over prior studies (Damböck et al., 2013; Gold, Damböck, Bengler, et al., 2013; Naujoks et al., 2015; Othersen, 2016), and changes in driver comfort with less continuous haptic involvement (Othersen, 2016) argue for a closer investigation of the current limitation of hands-off monitoring during PAD (Gasser et al., 2012). Further research seems necessary, as the beneficial notion of hands on the steering wheel from the regulatory perspective (Gasser et al., 2012; UNECE, 2017b) is in opposition to the possibly lowered driver comfort and acceptance from a human factors perspective. This work seeks to establish a data basis to estimate the benefit of continuous haptic feedback not only on driver interactions with PAD systems during use and control transitions, but also on drivers' attitudes towards the use of PAD systems.

Therefore, the degree of haptic feedback received by the steering wheel is compared in two configurations, in form of continuous hands-on or hands-off monitoring, in all of the studies in this work. Haptic feedback is considered as a characteristic of the PAD system. Hands-off monitoring is either mandatory or prohibited during automation use in this work for reasons of experimental control, instead of allowing for a more natural, but uncontrolled variation of hand posture by choice of the user. Some prior studies were conducted with small sample sizes (e.g., Othersen, 2016) and others exerted little experimental control over hand posture variations, making it difficult to draw conclusions on the impact of haptic feedback (e.g., Schaller et al., 2008), thus calling for a controlled experimental variation of this factor. Other studies have considered takeover situations requiring braking manoeuvres (e.g., Naujoks et al., 2015) and thereby possibly underestimated the effect of hand posture. Movement time to the steering wheel has been considered a factor contributing to the prolonged intervention times observed in some studies after hands-off monitoring (Gold, Damböck, Bengler, et al., 2013; Othersen, 2016).

SA-adverse factors did not influence the time taken to place the hands back on the steering wheel after PAD or HAD, whereas the complete takeover process was prolonged under SA-adverse conditions (e.g., Zeeb et al., 2015). The time to establish motor readiness did neither change in case of driver distraction nor under higher demands to achieve sufficient SA, such as for higher traffic density (Gold et al., 2016; Lorenz & Hergeth, 2015; Zeeb et al., 2015). The hands-on movement was instead described as "mostly reflexive following the acoustic warning signal" (Zeeb et al., 2015, p. 220) and assumed to be performed "without extensive situational assessment or understanding" (Gold et al., 2016, p. 648).

However, the irrelevance of the context of use on the "execution level" (Gold et al., 2016, p. 648) of the *takeover process*, as investigated in prior research (Louw et al., 2015; Zeeb et al., 2015), has to be distinguished from a possible relevance of the context of use when users are provided with different levels of haptic feedback *while supervising* automation. If the disadvantageous effect of hands-off monitoring on takeover performance lies only in additional movement time and the lack of haptic feedback has no influence on the involvement with the driving task, its effect should be rather small. Moreover, the disadvantage of hands-off automation use should be stable over conditions of different demand for maintaining or establishing SA. Takeover performance is a means to assess the relevance of any changes in system perceptions or interaction with the system, here considered under variations of haptic feedback and context of use.

In addition to takeover performance, the driver's monitoring strategy during automation use is evaluated. Lesser monitoring in combination with lesser physical involvement with the driving task might condition earlier or more pronounced reductions of SA, with possible consequences for control transitions. Other authors have already suggested that lesser haptic feedback might invite a lesser degree of supervision (Damböck et al., 2013; Othersen, 2016). However, when effects of hand posture on gaze behaviour have been reported, they were rather small (Gold, Damböck, Bengler, et al., 2013; Othersen, 2016). Next to the assessment of a general effect of reduced haptic feedback on drivers' monitoring behaviour, its stability over SA-adverse contexts is analysed. Variations of the context of use in this work are designed to resemble conditions inviting a reduced supervision of automation, based on findings of prior studies on PAD or other automation levels. The problem of *insufficient monitoring* is thus considered in this work by addressing changes in gaze behaviour under the concurrent limitation of haptic feedback.

In addition to the analysis of interaction and monitoring behaviour, the subjective relevance of hand posture on comfort and intention to use as well as its impact on the perceived capabilities of the system should be addressed. For instance, the correctness of the initial mental model was found to influence trust in a system (Beggiato & Krems, 2013) and trust, in turn, has been found to influence monitoring quality (Beggiato et al., 2015). Lee and See (2004), pointing out the importance of actual interaction with automated systems (see also Hoff & Bashir, 2015) in their consideration of trust as an attitude towards automation, state that "if the system is not trusted, it is unlikely to be used; if it is not used, the operator may have limited information regarding its capabilities, and it is unlikely that trust will grow" (p. 68). Resulting from this is a general importance of initial attitudes towards a system, as they influence the decision for or against the use of and gained experience with a system (Lee & See, 2004; Brookhuis et al., 2009). Furthermore, attitudes can provide explanations for behavioural effects observed, such as reduced monitoring efforts with increased trust (Beggiato et al., 2015; see also Manzey, 2008). Thus, users' initial attitudes need to be considered as they form the basis of behaviour (Ajzen & Fishbein, 1977; Lee & See, 2004).

A comprehensive understanding of preconditions for a safe interaction of users with partial automation thus needs to consider not only characteristics of system, here changes in haptic feedback, or in the context of use, but also the user. A user characteristic worth considering is prior driver experience, either in general due to its link to SA (i.e., hazard perception; Horswill & McKenna, 2004) or ADAS specific due to its influence on mental models apart from initial information (see also Wickens, 2008a, on the link between experience and SA). The user's prior experience with similar systems was further found to moderate the interaction with automation (Zeeb et al., 2016) and should thus be taken into account for the analysis of attitudes towards automation as well as interaction behaviour. As other research has identified gender as a relevant characteristic for the acceptance of automation in a priori settings (Feldhütter et al., 2016; Payre et al., 2014), an assessment of different PAD systems should include gender-specific effects for intended system use after receiving standardised initial information on PAD.

Although not being in the focus of this work, the relevance of *automation transparency* should be considered by implementing a similar HMI in all of the studies and by explaining system and HMI before use. Automation transparency as well as visual feedback from the system do thus not vary over the studies conducted. Similarly, variations in system complexity over studies can be prevented by using systems of the same technical scope and behaviour (see Chapter 5.1). Standardisation is also needed to induce the behaviour of interest, specifically different hand postures, in an experimentally controlled way while preventing the misinterpretation of automation levels regarding the driver's responsibility during automation use. Generally, standardisation over studies is considered beneficial to prevent changes in user behaviour solely due to a different instruction of the monitoring task (as discussed by de Winter et al., 2014). It seems further indicated to achieve at least a basic level of *system understanding* for all users by providing standardised instructions and initial information before first use.

The use of a system defines the point in time when knowledge (i.e., the initial mental model based on information) is augmented by experience with a specific system. Interaction strengthens or corrects parts of the initial mental model (Beggiato & Krems, 2013). Similarly, Hoff and Bashir (2015) differentiated between *initial learned trust* in automation prior to automation use, which is influenced by pre-existing knowledge, and *dynamic learned trust* during interaction. In accordance with Payre et al. (2014), the user's attitudes before interaction with the system are referred to as *a priori attitudes* in this work (as compared to *a posteriori attitudes* after use, see Figure 3-1; see also Bhattacherjee & Premkumar, 2004).

Results from other contexts, such as use of IT software or electric vehicles (Bhattacherjee & Premkumar, 2004; Jensen, 2014), and from vehicle automation (Gold et al., 2015) suggest an influence of repeated interaction on the attitudes of users. Effects of practice have also been found on the behavioural level (Zeeb et al., 2016). Changes in interaction strategies with increased exposure to partial automation should thus be investigated not only for attitudes and perceptions, but also for interactions in form of practice effects, similarly to analyses conducted by Körber et al. (2016) or Zeeb et al. (2016; see also Strand et al., 2014, for a discussion of the *first-failure* effect).

A methodological consideration concerns therefore the combination of different metrics, that is, subjective and objective metrics, and the stability of effects found in different phases of use to define the impact of an influencing factor. Whereas monitoring quality, manoeuvre success, takeover timing and quality are frequently used metrics, fewer studies considered the duration of differences in driving performance after termination of automation. With reference to a model proposed by Eckstein (2015), the descriptive assessment of driving performance after automated driving episodes aims at providing insights into post-automation effects in the contexts investigated in this work, as well as into the duration of such effects on driving performance after control transitions. In this work, manual driving performance is considered as the baseline level to anchor performance after or during interaction with an assistance system, similar to comparisons for response times or driving performance in other studies on automation (Damböck et al., 2012; Gold, Damböck, Bengler, et al., 2013; Merat et al., 2012; Skottke et al., 2014). “For the transport domain, the evaluation of an automated system is founded on a comparison to baseline values for these measures during unassisted driving. A fundamental issue for these methodologies is the identification of an appropriate baseline condition” (Ward, 2000, p. 404). By using a driving simulator and a defined test track (see Figure 3-1), a satisfying level of comparison can be established in driving conditions, road design and the timing of events. Unassisted, manual driving as a baseline of performance effects seems thus feasible for the current purpose. Metrics can be selected based on those used for the assessment of general driving performance in prior studies (see Chapter 5.2.4).

The study design should follow a worst-case oriented approach to give a conservative estimation of the relevance of haptic feedback on interaction safety. Challenging takeover scenarios in need of short-termed, immediate interventions are therefore used in the studies, as recommended in Gasser et al. (2012). To further ensure the subjective understanding of the necessity for re-engagement and to limit variance for a better analysis of takeover quality (Happee et al., 2017), the focus is on driving situations with clear need for steering reactions (see Chapter 5.1.2). Self-detected takeover need, where limits are encountered without a system-issued warning, should be especially sensitive to the detection of reduced SA (Louw et al., 2015). However, this system behaviour can be expected only in case of malfunctions or an insufficient situational assessment by the system. To assess a common use case, acoustic and visual feedback when reaching a system limit should be provided at the moment of disengagement, with an immediate need for action at TOR. Providing TOR as feedback instead as a timely warning for an upcoming system limit has been considered the worst case for higher levels of automation (Mok et al., 2015), as only the latter case allows the user to prepare for the control transfer. However, TOR as feedback can be assumed a common use case following the definition of PAD (Gasser et al., 2012). For a more comprehensive assessment, the effects of reduced haptic feedback on the self-initiation of control transitions when approaching visible system limits will however be addressed in addition to takeover performance after TOR in one of the studies conducted.

The main research question (RQ) concerns the role of haptic feedback in partial automated driving. Specific influences of user characteristics, the context of use as well as methodological aspects on the interaction with PAD in general and for haptic feedback variations in specific are addressed in separate studies. Aggregated data sets are used for analyses in need of larger sample sizes. The investigation of effects of PAD on post-automation driving performance, with a focus on the return to individual performance, constitutes the third research focus of this work.

RQ 1: Does the lack of haptic feedback during hands-off monitoring influence the assessment of or the interaction with partial automation?

RQ 2: Is the influence of haptic feedback stable over different user groups, driving contexts, methodologies and phases of interaction?

RQ 3: Does PAD influence the quality of post-automation driving performance?

The analyses conducted are aligned with the two axes of comparison in the research model (see Figure 3-1), differentiating when effects of driver-system-interaction are measured (*user experience*), considering influences on attitudes and behaviour alike, and the methodology applied to investigate effects (*method*). The second axis of comparison is established to address methodological influences on the size and stability of feedback effects. Results found for different abstractions of the driving task, here a survey (for attitudes), driving simulator studies and a test track study (for attitudes and behaviour), will be compared, similar to the work of Stanton and Young (1998). Characteristics of user and context are considered in separate data sets, whereas the main experimental variation of haptic feedback (as a system characteristic) is considered in all sections.

Attitudes towards partial automation:

The first analysis section concerns users' *a priori attitudes* towards automation. An online survey (Chapter 4) explicitly addresses the effects of haptic feedback as part of initial user information on mental models and attitudes. Relevant user characteristics, as identified by the literature review, are considered. The question behind this a priori assessment is whether the assumed comfort feature (Othersen, 2016) of waiving the need for a continuous haptic involvement with the driving task changes users' expectations towards the system or their intended usage behaviour. A *change in attitudes* through system use is examined by comparison before and after use in the last analysis section of this thesis, using questionnaire data collected in simulator and test track studies (Chapter 7).

RQ 2.1: Does the continuous need for hands-on supervision have an influence on the a priori system assessment, the intention to monitor or to use PAD in different user groups? (Survey, Chapter 4)

RQ 2.2: Is the continuous need for hands-on supervision a relevant factor for the acceptance and the perceived comfort of partial automation after use? (Aggregated data set; Chapter 7)

RQ 2.3: Does practical experience with a system change the perception of the system in case of correct initial mental models? (Comparison for a priori and a posteriori attitudes for an aggregated data set; Chapter 7)

Interaction with partial automation in different contexts of use:

The second part of the analysis section concerns the results of four different driving simulator studies[3]. Focus of this section is the user's interaction with a system in different *contexts of use* (Chapter 5). Based on the literature reviewed, the duration of monitoring, the availability of secondary tasks and the complexity of the driving situation seem relevant for monitoring automation.

---

[3] Results on takeover timing, subjective ratings and parts of the quality assessment for three of the driving simulator studies (Study 1, 2 and 3) as well as the test track study (Study 5) have prior been published in Josten, Zlocki and Eckstein (2016). To provide a comprehensive overview, results from these studies are reported here in similar, but partly adapted form. Data published prior can be found in Chapter 5.3, 5.4, 5.5, 6.1, 6.2, 6.3 and 7. Data reported in Josten et al. (2016) for Studies 1, 2, 3 and 5 of this work further provides the basis for additional analyses conducted on takeover quality and monitoring behaviour to complement the results published prior. Data on short-termed takeover situations was collected in a project funded by the Forschungsvereinigung Automobiltechnik e. V. (FAT) and the Bundesanstalt für Straßenwesen (BASt) (see Josten et al., 2016).

A longer duration of use should reduce monitoring quality due to vigilance effects, as was found for monitoring tasks in general as well as during PAD for the detection of system failures (e.g., Endsley, 1995; Othersen et al., 2014). Its effect in combination with hands-off monitoring is thus investigated in a driving simulator study (Study 1; Chapter 5.3). In addition, the complexity of a takeover situation is addressed in this first study by comparing two scenarios placing different demands on the situational assessment before takeover, that is, a lane change scenario (where performance relates directly to gaze behaviour; e.g. Metz & Landau, 2015) and a lane keeping scenario on a curved road.

An increased engagement in visual secondary tasks, especially during monotonous driving situations (Jamson et al., 2013; Othersen et al., 2014), can be assumed to impact SA negatively. Visual-manual tasks have accordingly been shown to influence takeover performance (Eriksson & Stanton, 2017; Merat et al., 2012; Zeeb et al., 2016). The effect of visual-manual driver distraction under variation of haptic feedback is addressed in a second driving simulator study (Study 2; Chapter 5.4).

In Study 3 (Chapter 5.5), the complexity of the situation of use is targeted by manipulating traffic density. Based on prior research, higher traffic densities, leading to a larger number of elements in the immediate surroundings of the vehicle, should increase the motivation to monitor the system (Jamson et al., 2013) and thus increase the level of SA maintained. Then again, a higher number of elements might also lead to a longer re-orientation phase at TOR with effects on takeover time, but not on motor readiness (Gold et al., 2016). Here, the effect of traffic density is investigated for variations of haptic feedback during automation use.

A fourth driving simulator study (Chapter 5.6) addresses the question whether haptic feedback and an assumed change in involvement with the driving task influence the self-detection of takeover situations. The anticipation of takeover situations allows for a quantification of changes in monitoring behaviour aside from glance-based metrics, by assessing the frequency of self-initiated control transitions before a TOR is issued by the system when reaching a system limit.

RQ 2.4: Does haptic feedback change the user behaviour in SA- or monitoring-adverse contexts of use? (Study 1 – Duration of use, Chapter 5.3; Study 2 – Secondary tasks, Chapter 5.4; Study 3 – Traffic density, Chapter 5.5)

RQ 2.5: Does haptic feedback influence the occurrence of self-initiated, planned control transitions in view of anticipated system limits? (Study 4, Chapter 5.6)

Context-independent stability of the interaction with partial automation:

Alongside the analysis of user, context and system variables and the change in attitudes by experience, the *stability of behavioural effects* is a focus of the current work. For behavioural data, different phases of use and corresponding performance indicators will be considered. The influence of the level of abstraction of the driving task (i.e., the method of investigation) is analysed by comparing results of driving simulator studies to a test track study with a comparable experimental design (Chapter 6.1). The stability of influences on user groups is examined for objective data (interaction quality; Chapter 6.2) in addition to the analysis of system preferences (see Chapter 4 for a priori preferences; see Chapter 7 for a posteriori preferences). Further considered is the influence of *practice* on takeover performance (Chapter 6.3).

RQ 2.6: Is the effect of haptic feedback stable over different phases of interaction, specifically during use, during control transitions and after use? (Study 1-5 and aggregated data sets)

RQ 2.7: Is the effect of haptic feedback independent of the methodological approach, specifically for different abstractions of the driving task? (Study 5; Chapter 6.1)

RQ 2.8: Do user characteristics or practice change the interaction with partial automation or the influence of haptic feedback on interaction quality? (Aggregated data sets; Chapters 6.2 and 6.3)

Post-automation driving performance:

For all studies conducted, post-automation driving performance (see also RQ 2.6) is analysed with a special focus on the end of corrections after control transitions, following a model proposed by Eckstein (2015). Manual driving performance is used as a baseline to define when individual driving performance is re-established after takeover. Next to a general assessment of post-automation effects, analyses are conducted to establish whether system, context or user characteristics influence the time needed to regain the individual level of driving performance after use of partial automation.

RQ 3.1: Does the duration or magnitude of effects on post-automation driving depend on user, context or PAD system factors? (Descriptive data analyses; Chapters 5 and 6)

4 A Priori Attitudes

An online survey was conducted to measure the influence of instructed haptic involvement (hands-on or hands-off monitoring) on a priori system assessment, assumed system capabilities and the intention to use. An additional aim of the survey was to investigate the general influence of selected user characteristics on a priori attitudes towards automation with regard to the sample selection for following studies. Ratings were based on initial information concerning driver responsibilities during system use and system capabilities (see Figure 4-1).

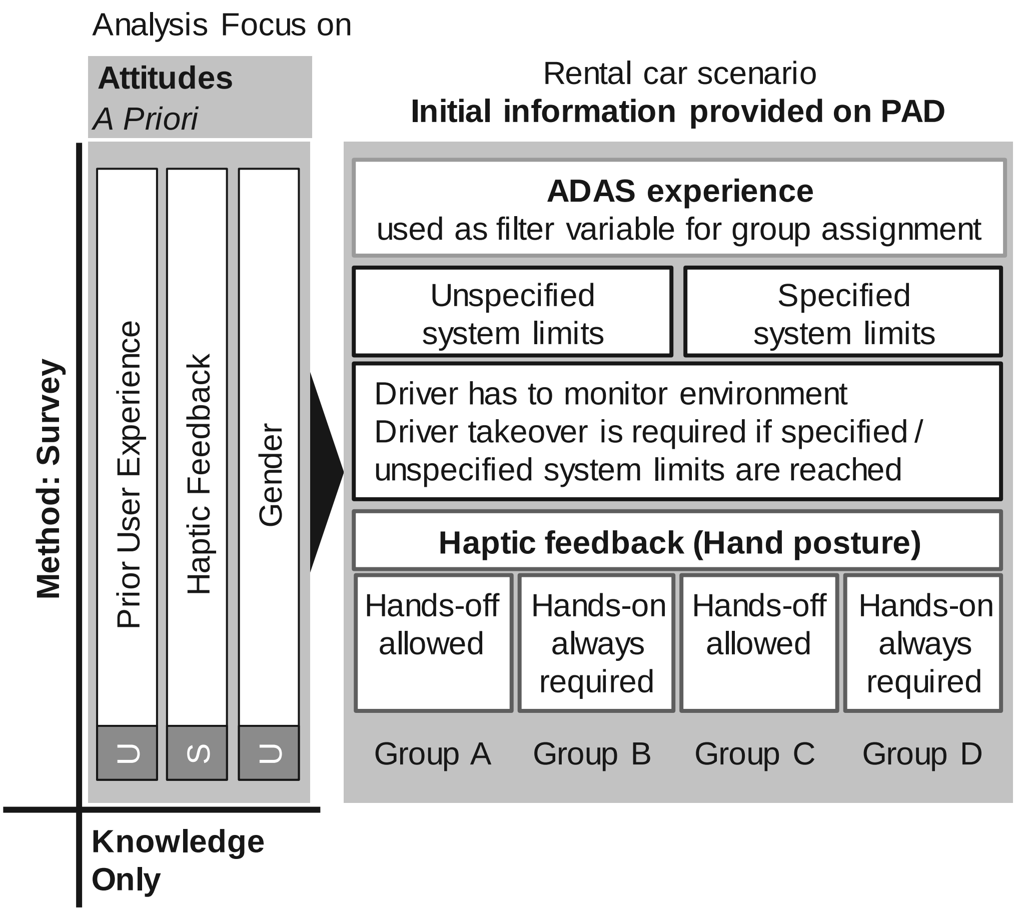

Figure 4-1: Overview on user (U) and system (S) characteristics included in the survey (left hand side) and the experimental groups defined by the initial information provided (right hand side).

4.1 Method

The measurement of a priori attitudes towards partial automation was part of a larger survey on information needs and privacy concerns for different levels of automated driving systems. Two automation levels, SAE Level 2 und Level 3 systems (SAE, 2018), were presented to participants in balanced order. Only aspects concerning partial automation (Level 2) are considered in the following analysis. A detailed description of all other aspects included in the survey has been provided in Josten, Schmidt, Philipsen, Eckstein and Ziefle (2017; 2018)[4].

Participants were recruited using emailing lists for students and employees of RWTH Aachen University as well as social networks. Based on demographic information, participants were assigned to one of four experimental groups (see Figure 4-1). A selection mechanism ensured an equal number of participants with prior ACC experience in each of the groups, thereby balancing prior ADAS experience as one of the central user variables at least partly over groups. Completing the survey took about 30 minutes.

Bhattacherjee and Premkumar (2004) addressed the possible influence of external information available to users via, for example, media channels when discussing their results on a priori user attitudes (see also Feldhütter et al., 2016). To address this problem for the current survey, the four experimental groups received standardised system descriptions in the beginning of the survey to introduce a different mental model within each group. This has also been done in other studies to introduce new driver assistance technology (Beggiato & Krems, 2013; Kyriakidis et al., 2015). Participants were informed that they had received a rental car with a new driver assistance system for automated highway driving and that the information displayed was taken from the manual of the vehicle. Always included was information on the tasks the system would be performing when activated, that is, longitudinal and lateral guidance as well as regulation of the distance to a lead vehicle. The driver's task, the supervision of the system within the current driving situation, was stressed as well as the necessity to always be prepared to take back control should the system reach its limits. All further information was varied between groups and concerned the hand posture while supervising the system (Groups A and C versus Groups B and D, see Figure 4-1) as well as the amount of detail provided on system limits (Groups A and B versus Groups C and D, see Figure 4-1). Participants were either given specific information about the system's limits, which were introduced as non-detectable lane markings or the end of the own lane, or were simply told that the system had limits, but not which. The manipulation of information on system limits was included to raise uncertainty regarding the system (uninformed mental model) compared to a mental model of higher detail (informed mental model) on the basis of limited information about system capabilities being an assumed obstacle for trust (Lee & See, 2004). Each information group was split into two groups receiving either the information that

[4] Data reported in this chapter was collected in a project funded by the Profile Area ICT of RWTH Aachen University (RWTH-PH-ICT).

at least one hand had to be on the steering wheel at all times while the system was active or that both hands could be taken from the steering wheel if the system was active. As the manipulation of initial information was introduced to test effects on information needs of naïve users later on in the survey (see Josten et al., 2017; 2018), results will be reported for the aggregated groups A and C (Group Off) and groups B and D (Group On) only, excluding any effects of initial information. The full instructional text (for unspecified system limits) can be found in Appendix 12.1 for hands-off driving (Group Off) and in Appendix 12.2 for hands-on driving (Group On).

For the survey, attitudes towards the system were measured on 6-point Likert Scales ranging from “I do not agree at all” / “Not at all” to “I completely agree” / “Completely”. Participants were asked to rate the usefulness of the system, the comfort of use, the attractiveness, the assumed change in effort when driving with the system and their trust in a successful handling of highway driving situations by the system. Finally, participants were asked to indicate their intention to use the introduced system on highways (binary measure). To further get a first impression on how lesser restrictions of hand posture influence the a priori reliance respectively the necessity to supervise the system, three static highway-situations, as viewed through the windscreen of a vehicle, were displayed within the survey. Each situation featured cues related to system limits, either the end of one’s lane, roadworks with narrowed lanes and temporary lane markings or a straight road section with missing lane markings. Ratings on the subjective importance to be visually involved in these situations when using automation were collected on the above-described Likert scale.

Driver characteristics collected at the beginning of the survey include gender, general driving experience (in km driven per year), ADAS experience and technical affinity (measured on the scale by Karrer, Glaser, Clemens, & Bruder, 2009). All independent variables apart from gender are considered as continuous measures. ADAS experience is reported as the summed experience with four assistance systems, namely ACC, CC, lane keeping assist (LKA) and park assist. Experience was measured on a four-point scale for each system including the options “I don’t know this system” (0), “I know this system but have never actually used it” (1), “I used this system once” (2), “I have used this system multiple times” (3) and “I use this system regularly” (4).

4.2 Survey Results

The survey was completed by 130 participants, of which 123 were considered for analysis as the remaining seven participants answered less than half of all relevant questions concerning PAD ratings. Sample size varied slightly over analysis due to missing values.

Of the 123 participants, over 70 % held a university degree, indicating a high level of education within the collected sample. Three participants did not hold a drivers licence. They were included in analyses as the focus of the survey was on information-based a priori attitudes towards automation and driving experience was one of the user characteristics of interest. Due to the number of dropouts after group assignment by ADAS experience, including all participants not completing the survey, a different number of valid participants was considered for each of the experimental groups. The hands-off system was rated by 66 participants (n = 22 of 66 in Group A with unspecific system limits, i.e., 33 %) and 57 participants rated the hands-on system (n = 26 of 57 in Group B with unspecific system limits, i.e., 46 %). Gender was balanced within the overall sample with 58 female participants (47 %), as well as within both experimental groups, as can be seen from Table 4-1. Some of the user characteristics within the sample, such as age, did cover only a narrow range of values, as can be seen by the percentiles provided in Table 4-1.

General intention to use the systems was high with 73.2 % of affirmative answers (n = 90 of 123). The same percentage of users was affirmative in both feedback groups (Group Off: 72.7 %; Group On: 73.7 %). Attitudes towards automation were generally positive, that is, above or at the scale's median, with median ratings of 4 or 5 for the hands-on system and median ratings of 3 or 4 for the hands-off system. The lowest median rating was trust in the handling capabilities of the hands-off system with a median of 3. For all items, the full scale was used for ratings (i.e., range from 1 to 6). All items were correlated (Spearman correlations between usefulness, comfort, attractiveness, effort and trust; all $p < .001$; all $\rho > .51$). An overall score, labelled *attitude*, was therefore calculated for comparison of the two systems, using the mean[5] of all single items per participant instead of separate analyses of medians for each item. The hands-on system was perceived as more positive compared to the hands-off system by tendency ($z = 1.84$, $p = .066$, $n = 120$; $M_{on} = 4.13$, $SD = 1.21$; $M_{off} = 3.74$, $SD = 1.19$). Regarding the subjective intent to be visually involved with the driving task, a higher mean subjective need was found for the hands-off system over all three situations ($z = -2.54$, $p = .011$, $n = 115$; $M_{on} = 4.81$, $SD = 1.02$; $M_{off} = 5.26$, $SD = 0.81$).

The relationship between user characteristics, attitudes towards and intended use of partial automation was investigated by correlational analysis (using Spearman's ρ; see Table 4-2 for the summarised correlation matrix). Of all characteristics

[5] The mean of multiple ordinal items provides a continuous measure (see Van der Laan, Heino, & de Waard, 1997). However, as the assumption of normality was not met by data, non-parametric tests are applied.

investigated, prior ADAS experience and general technical affinity were most strongly associated with attitudes towards automation. Age, being included with a restricted range from 22 to 51 years, was not associated with any of the attitudes collected. Apart from age, all user characteristics investigated were moderately associated with the intention to use the system.

Table 4-1

Overall and group specific sample characteristics for the online survey

User Characteristic	Overall; Group Off / Group On
Age (years)	Overall: *M* = 30 (9); IQR: 24 - 31; range: 22 – 51; Group Off: 30 (10) / Group On: 29 (9)
Gender (-)	Overall *N*: 65 male / 58 female; Male: n_{Off} = 36 / n_{On} = 29 Female: n_{Off} = 30 / n_{On} = 28
Driving Experience (kilometres driven annually)	Overall: *M* = 10771 (13350); IQR: 1000 - 15000; range: 90 - 35000; Group Off: 11372 (15774) / Group On: 10034 (9684)
ADAS Experience (-)[a]	Overall: *M* = 8 (4); IQR: 5 - 10; range: 3 - 13; Group Off: 7 (3) / Group On: 8 (3)
Technical Affinity (-)[b]	Overall: *M* = 4.11 (0.61); IQR: 3.75 - 4.50; range: 3.17 - 5.17; Group Off: 4.15 (0.66) / Group On: 4.07 (0.55)

Note. Displayed are means (*M*) and standard deviations (*SD*; in brackets). IQR = interquartile range; range: 5^{th} to 95^{th} percentile. The hands-off and hands-on system were rated by two groups of participants (Group Off: *n* = 66 / Group On: *n* = 57).
[a]ADAS experience reported as experience points (0-16) for prior use of four different systems (ACC, CC, LKA, park assist).
[b]Technical affinity measured on the scale by Karrer et al. (2009; 6-point-scale).

Additionally, the user characteristics investigated were correlated to at least small degrees. ADAS experience was positively correlated with the amount of kilometres driven per year ($\rho < .37$, $p < .001$) and the technical affinity reported ($\rho < .41$, $p < .001$) as were technical affinity and driving experience ($\rho < .22$, $p = .018$). Female participants were associated with driving less kilometres per year, lower technical affinity and lesser prior use of ADAS (all $\rho < -.41$; for all correlations: $p < .001$).

Table 4-2

Summarised correlation matrix of user characteristics, the subjective importance of system monitoring and five a priori attitudes investigated (Spearman's ρ).

	Gender	Age	Driving Experience	ADAS Experience	Technical Affinity
Visual Monitoring	**.28****	.10	-.12	-.08	-.01
Usefulness	-.14	.00	.15	**.40****	**.30****
Comfort	-.03	.01	.09	**.25****	**.23****
Attractiveness	**-.20***	.01	.15	**.38****	**.34****
Reduction of Effort	.20	-.06	.03	**.20***	.17
Trust	-.15	.07	.06	**.37****	**.21***
Intention to Use	**-.24***	.11	**.22***	**.33****	**.26****

Note. The number of valid responses ranges between $n = 111$ and $n = 123$ for correlations. Significant correlations are printed in bold. Asterisks indicate the level of significance as follows: * $p < .05$ and ** $p < .001$.

4.3 Discussion of Attitudes Before Use of Automation

An online survey was conducted to identify relevant user characteristics on a priori attitudes towards partial automation in general and under variation of feedback. Beliefs and attitudes, determining the intention to use as well as the actual use of technology should be taken into account when evaluating new technology (see, e.g., Bhattacherjee, 2001; de Winter et al., 2014; Körber & Bengler, 2014; Lee & See, 2004; Payre et al., 2014; Rudin-Brown et al., 2003). One aim of the survey was to identify user characteristics to consider for sample selection when investigating the interaction with partial automation. Additionally, the survey was intended to provide first indication of the subjective relevance of haptic feedback for the intention to use PAD and for the (a priori) importance of system monitoring.

Due to dropouts (i.e., participants who aborted the survey at an early stage, but were originally assigned to one of the four instructional groups), the number of the 123 participants considered for analysis was not balanced over the different instructional groups. This issue relates mainly to the instruction of system limits, which was not part of the analyses reported here, and lead to a larger number of participants with specific knowledge of system limits in both experimental groups. Thus, effects of imbalanced instructional effects might affect the results, but at least so in the same direction for each PAD system group (On and Off).

No clear differentiation for attitudes was found between systems, but a significantly higher subjective importance of system supervision for the hands-off system was found. The finding of a higher subjective importance for monitoring might further be related to the descriptively lower trust reported by the hands-off instruction group, as trust was shown to influence monitoring behaviour (Beggiato et al., 2015). Behavioural measures in future studies should validate whether driving hands-off results in an actual improvement of visual monitoring as a strategy to counter reduced haptic feedback.

The survey sample consisted mainly of young participants (IQR: 24 - 31 years), possibly affecting correlation coefficients concerning age by range restriction as was also argued by Kyriakidis et al. (2015) who found non-significant effects for age with a similarly young, but much larger sample (N = 4886). Gender proved to be the only user factor related to the subjective importance to monitor automated systems and was further associated with the general intention to use. This is in line with prior findings on gender differences. Kyriakidis et al. (2015), focussing on higher levels of automation, found that men were willing to pay more for automation and were “somewhat less worried about fully automated driving vehicles than women” (p. 134). A post-hoc analysis of the importance to be visually involved revealed significant differences between women and men for the hands-off system, but only by tendency for the hands-on system (Group Off: z = 2.83, p = .005; Group On: z = 1.78, p = .076) with higher ratings for women (overall means; M_{female} = 5.32, SD = 0.72; M_{male} = 4.79, SD = 1.00). This finding aligns with the above cited result that women were found to be more worried of higher levels of automation, in this case of exerting lesser control over the system after letting go of the steering wheel (cf. Damböck, 2013, for the idea of hands-off driving as a higher level of automation). Considering the correlations found for attitudes and ADAS experience, the combination of lesser driving and ADAS experience for women might also be a reason for their more critical view on automation. Overall, the results on women being generally more sceptical towards new technology are in line with those of other studies (Feldhütter et al., 2016; Kyriakidis et al., 2015; Jensen, 2014; Payre et al., 2016).

Of the user characteristics included, only prior ADAS experience was congruently associated with positive attitudes towards automation. Users who have regularly used other ADAS before might be more inclined to use new assistance systems as well, as indicated by results of Kyriakidis et al. (2015), finding the strongest correlations “between income, mileage, driving frequency, and current ACC use, on the one hand, and willingness to pay [for automation], on the other” (p. 136). The hypothesis of Kyriakidis et al. (2015) that the best predictor for future behaviour (i.e., buying or using new ADAS) might be current behaviour (i.e., use of current ADAS) is not rejected by the results of this survey, but remains to be proven by actual data of system use in daily life.

As nearly half of the sample reported no or only very limited experience with ACC or other ADAS systems, some participants included in this survey might have lacked a general understanding of what partial automated driving might feel like. Furthermore, the possibility to take the hands from the steering wheel also includes the option to keep both hands permanently on the steering wheel if the driver wishes to do so. However, no data on the planned extent of hands-off driving was collected. To generate a standardised first impression of driving with different hand postures, instruction for future studies should enforce different hand postures at least for initial contact to enable a controlled measurement of changes in attitudes and supervision behaviour before and after use for both hand postures.

Future studies should investigate in more detail how experience with a system changes the attitudes towards said system and whether the degree of change is associated with prior experience with similar systems. As changes in attitude, at least in the IT context, “tend to be more prevalent during the initial phases of […] usage” (Bhattacherjee & Premkumar, 2004, p. 249; see also Feldhütter et al., 2016, for the difference between changes induced by information and use), prior experience with similar systems could provide a more robust basis for a priori beliefs and attitudes that should be reflected in fewer changes after use, at least if prior experience does not contradict current experience (cf. Beggiato & Krems, 2013). It seems advisable to consider experience with ADAS such as ACC or CC for sample selection to better explain possible behavioural differences between users, also with regard to expected changes in the detection of control transitions and mental model stability (e.g. Seppelt & Victor, 2016).

5 Interaction with Partial Automation in Different Contexts of Use

The main focus of the following analyses is on behavioural measures (see Figure 5-1). Additionally, subjective ratings concerning the context of use realised within different studies will be considered in this section. Concerning unique aspects of each study, these ratings provide a background against which to interpret behavioural effects found for specific context variations. Beforehand, an overview over methods (Chapter 5.1) and the choice of metrics (Chapter 5.2) will be provided.

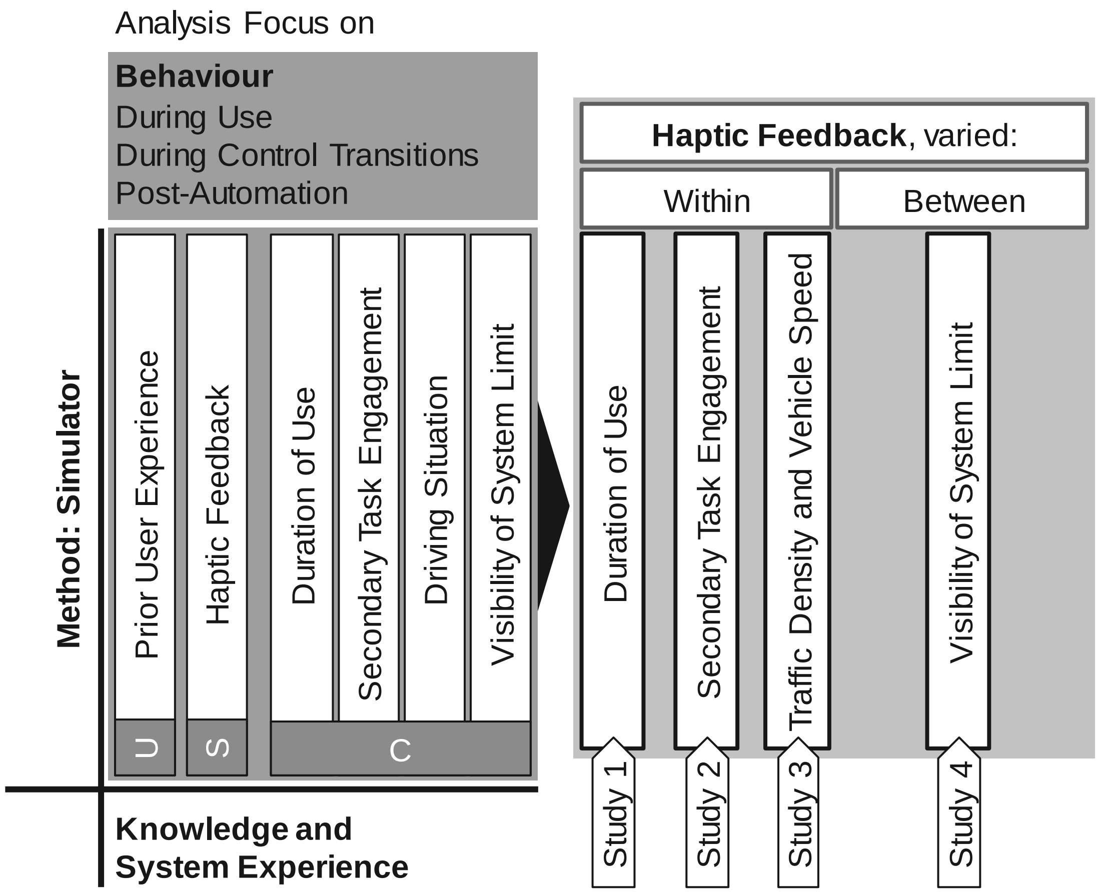

Figure 5-1. Characteristics of user (U), PAD system (S) and context of use (C) included in the driving simulator studies (left hand side) and their assignment to the four studies (right hand side).

5.1 Simulator Study Setup

For the following studies, performance is measured predominantly in unexpected, short-termed takeover situations, requiring immediate steering input from the driver for a successful situation management. This procedure was chosen to prevent preparatory actions, which lead to problems regarding the clear assignment of hands-on and hands-off conditions for analysis in prior studies on hand posture (e.g., Naujoks et al., 2015; Othersen, 2016). Furthermore, close compliance with one specific hand posture (hands-on or hands-off) was encouraged by instruction for the same reason. Preferences for hand posture in automated driving are thus not deductible from the performance data, but only from subjective ratings. This procedure allows instead for a clear differentiation of haptic feedback within the contexts investigated.

In all scenarios investigated, a TOR was issued by the system in the moment of system deactivation. The TOR served as acoustic feedback on the change in system status, not as an advance warning. An immediate takeover was required from the driver. Situations with erroneous system behaviour or system malfunctions, relying on the driver as supervisor to detect the need for takeover, were not investigated. Situations with immediate takeover need without ambiguity were implemented to rule out unwanted variance arising out of the subjective decision of if and when to intervene. However, driver-initiated control transitions in preparation for upcoming TOR-events were additionally investigated in one of the studies (Study 4; see Chapter 5.6).

All simulator studies were carried out using the same vehicle, the front half of a BMW 5 series with fully functional control elements. Data acquired for short-termed takeover situations (Study 1) was partly collected in a moving-base simulator (Hexapod with 6 degrees of freedom). For analysis purposes, no differentiation was made between static and dynamic driving simulator data. Due to the within-subject design and depending on the measures chosen for performance evaluation, differences in the test method should not influence comparisons. The two driving simulators used for testing provided a field of view of about 180° x 40° (Study 1-3), respectively 220° x 40° (Study 4). Rear view was enabled via the side and rear view mirrors by additional screens positioned behind the vehicle. Driving simulator data were recorded with 100 Hz. Surrounding traffic and track design were realised with SILAB (Würzburger Institut für Verkehrswissenschaften GmbH; Studies 1-3) respectively VTD (VIRES Simulationstechnologie GmbH; Study 4). Further details on specific experimental set-up and procedure are provided in the following chapters.

5.1.1 General Study and Automation System Design

One challenge in interpreting the results of the survey was the between-subjects design for variations of hand posture. As user characteristics have been shown to influence system preferences and thus likely the interaction with the system, for example the subjective need to monitor a system, experimental variation in the simulator studies was realised within subjects wherever possible (the exceptions being Study 1 regarding traffic scenarios and Study 4 regarding hand posture). Experimental conditions were counterbalanced to avoid interference with learning effects, time-on-task or general drowsiness (cf. Happee et al., 2017; Körber et al., 2016; studies on prolonged driving, e.g., Neubauer et al., 2014). Next to conditions including the use of partial automation (hands-on and hands-off; i.e., *on* and *off*), a manual baseline drive (*man*) within the same setting was included in each study to gather data on unassisted driving performance.

Two highway scenarios were chosen as exemplary takeover situations requiring a steering intervention. A control transition was indicated either shortly before entering a sharp curve or shortly before the own lane ended due to a beginning construction zone (see Figure 5-2). The curve scenario was included as a reference for comparisons over all within-subject studies. Both situations required the driver to immediately conduct corrective actions and enabled the measurement of takeover timing as well as manoeuvre success and driving performance. Following considerations of Kerschbaum et al. (2014) on TOR during curve driving, who assumed a higher difficulty to grasp the steering wheel at a high steering wheel angle, TOR were always issued with steering wheel angles around 0°. This prevents effects of study- or vehicle-specific reset forces for disengagements at higher steering wheel angles. To this end, a straight road element (200 m) was implemented adjacent to the takeover scenario in each track design.

Degrees of freedom in takeover scenarios were found to influence the variation of performance (Happee et al., 2017). Available space was thus restricted by surrounding traffic when entering a sharp curve, encouraging drivers to start steering immediately after the takeover request and preventing evasive manoeuvres onto other lanes. If drivers did not regain control immediately, the vehicle collided with surrounding traffic on the left lane (Figure 5-2).

For the construction zone scenario, any neighbouring vehicles on the left lane left a gap large enough for the driver to change lanes without causing a collision. If drivers did not act in time, the vehicle collided with the traffic cones positioned on the temporary lane markings. Lead vehicles also changed lanes before the construction zone (for the exception, see 5.1.3), starting approximately 800 ms before the TOR was issued (Study 1) or starting at the time of TOR with a lane change duration of around one second (Study 4). After each TOR, participants drove manually until the next automation offer occurred to enable analyses on manual driving performance. After the construction zone, participants had to change back onto the right lane. The track design was kept the same for different drives within one study to enable performance comparisons between conditions under similar affordances of the driving task, apart from adaptations necessitated by balancing different context conditions. Different environment configurations were included to prevent expectations based on track design in studies applying within-subject designs.

Both types of situation required the driver to conduct a steering manoeuvre, including the need to establish motor readiness (Zeeb et al., 2015) in difference to longitudinal manoeuvres. However, whereas curve negotiation places demands solely on lane keeping performance (i.e., stabilisation; Donges, 2009), the construction zone places demands on the awareness of surrounding traffic for lane change manoeuvre planning and on consecutive lane keeping performance (i.e., guidance and stabilisation; Donges, 2009; see also Vollrath & Krems, 2011).

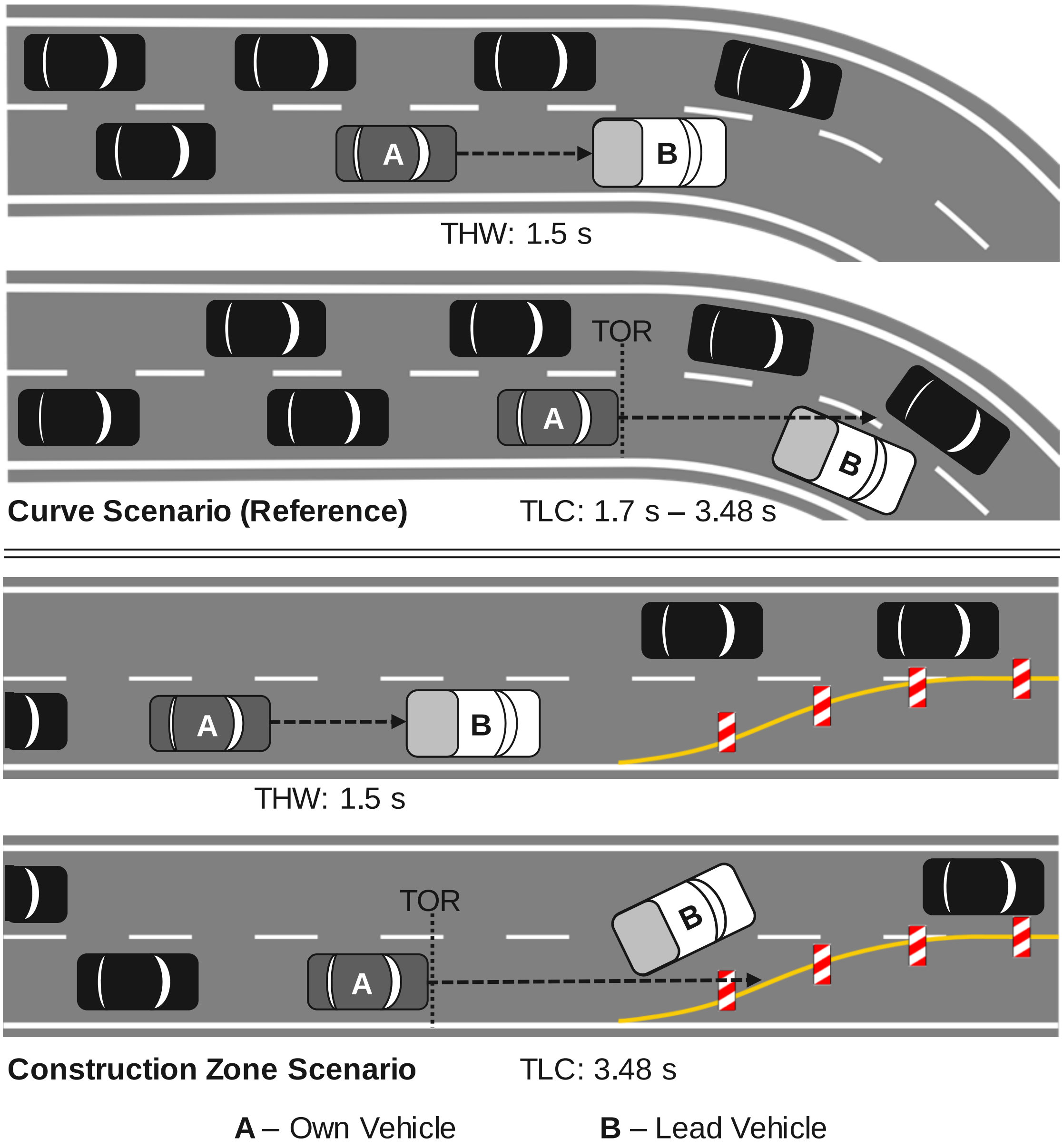

Figure 5-2: Automated drives and takeover scenarios (upper part: curve; lower part: construction zone) for Studies 1, 2 and 3. Time to line crossing (TLC) for curve scenarios varied between context variations and studies according to speed and radius of the curve. Following distances (THW) were the same for all studies on short-termed takeover (Figure adapted from Josten et al., 2016, p. 28).

Driving performance after automation was compared to continuous manual driving in uncritical situations (i.e., curve driving) and after specific events (i.e., late visibility of lane change necessity in construction zone). The timing of actions in response to unexpected events after automation use can only be compared to manual driving for the construction zone. However, performance quality can be compared between unassisted and post-automated driving in both scenarios.

The PAD systems implemented in this work were prototypical systems, designed to incorporate all relevant aspects of PAD while at the same time keeping instructions for system use and the effort for system implementation at a functional and manageable level. Relevant for the systems' design were the driver's responsibilities and thus a realistic interaction with PAD systems based on the theoretical definition of Level 2 automation, albeit the partially simplified system description and behavioural logic. Control transitions investigated in the studies were changes between Level 0 and Level 2 only, meaning that both longitudinal and lateral control were transferred to either driver or system at the same time, excluding partial control transitions of either longitudinal or lateral control (change from Level 2 to Level 1; SAE, 2018).

All drives took place on two-lane highways. To be able to better control the surrounding traffic in takeover situations, apart from using the right lane where possible, participants in all studies were asked not to overtake lead vehicles except when driving safety required such a manoeuvre or the experimenter explicitly asked for it. The implemented systems allowed activation only on the right hand lane and participants were told that the PAD systems could not perform lane changes on their own. Implemented speed restrictions for engagement, with engagement only below 50 kph for traffic jams, and the necessity of visible lane markings further prohibited untimely re-engagement after a TOR to measure post-automation driving performance. Following distance to the lead vehicle (THW) could not be freely chosen by participants as it influences the criticality of the takeover by design. THW was instead set to 1.5 s for short-termed takeover situations and to 2.0 s for Study 4. Details of system instruction differed only slightly between the studies and can be found in the Appendix (German instruction; see Appendix 12.3 and 12.4). The hands-on and hands-off system were implemented as the same system, only the instruction for drivers differed. To participants, hands-on and hands-off system were thus presented as two different systems which had to be tested and compared.

As no technical solution for hands-on detection was available in the simulator, the experimenter ensured the correct hand posture via camera surveillance and reminded participants, if necessary, of the correct hand posture shortly after automation engagement. Hand posture was additionally recorded for post-hoc analysis with 120 frames/s (Study 1-3) respectively 25 frames/s (Study 4).

The system state was displayed in the instrument cluster at all times during the automated drives. Four different system states (not available, thus no engagement possible; available, thus engagement possible; active, that is, automation engaged; TOR; see Appendix 12.4 for an overview that has also been published in Josten et al., 2016) were introduced as part of the general instruction. The HMI featured all elements relevant for system supervision, with (predefined) distance information to the lead vehicle, lane markings detected as well as a steering wheel icon to indicate the transfer in driving responsibility (for similar HMI solutions or suggestions, see, Alessandretti, Amditis, Metzner, Johannson, & Fahrenkrog, 2014; Louw et al., 2015, UNECE, 2017b). System offer and system disengagement were accompanied by different auditory signals to ensure immediate driver perception.

5.1.2 Short-Termed Takeover Situations

Three driving simulator studies on short-termed takeover situations were conducted (Study 1-3), focussing on interaction with automation in different contexts of use. Each study started with general information on the procedure, a demographic questionnaire and a short familiarisation drive. This first drive included familiarisation with the vehicle and with specifics of the driving task (for 3 km of highway driving), such as keeping a defined distance to a lead vehicle during manual driving. Each participant engaged automation once during a short traffic jam in this drive (around 2 km of highway driving) and manually disengaged automation by steering at the end of the drive. No system-initiated takeover was experienced within the familiarisation drive, but the according HMI signal and sound were presented during a general introduction of the system, ensuring the correct understanding of the TOR signal and the according shift in responsibilities. Three experimental drives (man, on, off) followed for each participant, each of them featuring two measurement situations. The order of conditions was balanced over the sample. A questionnaire on a priori attitudes and a follow-up questionnaire on the automated system and takeover situations after the drive were filled in for automated drives. For manual drives, a short follow-up questionnaire on driving demand was filled in. After all three experimental drives, participants filled in a questionnaire comparing the three drives and were compensated for participation.

Participants were told to use automation during the drives as often as possible, but were also shown how to disengage automation if needed. This could be done using a steering wheel button (same as for activation), by steering (2 rad difference in steering angle), braking (15 % of maximum pedal travel) or accelerating (40 % of maximum pedal travel). Thresholds for manual disengagement were set deliberately high to prevent unintentional disengagement. No situation occurred during the drives in which manual disengagement would have been necessary, as the system disengaged itself simultaneously with the TOR. In case of driver-initiated disengagement, participants had to wait for the next offer to re-activate automation (3 s wait time). The system, once engaged, followed the current lead vehicle with a fixed THW while keeping to the centre of the own lane. The lead vehicle for the complete drive was either a van or a truck, high and broad enough to block the view on the upcoming traffic situation and to reduce early recognitions of the takeover situations. Generally, speed varied slightly during automated driving, but stop and go was prevented during traffic jams to increase monotony. Each drive began and ended on a highway rest area.

To compare performance after different automation conditions without enabling planned changes in hand posture, likely mediating differences between conditions, takeover situations with immediate need for intervention were designed. Nonetheless, the time between automation disengagement (TOR) and the lane departure in case no driver input occurred needed to be long enough to enable a successful takeover for immediate driver interventions, but urgent enough to present the need for immediate action. Warning times for collision avoidance systems present an adequate benchmark, as their purpose is similar to a TOR in partial automation, which is to raise awareness to the current driving situation with need for immediate driver action. An overview by Reinisch (2012) on warning strategies for collision avoidance systems placed visual or acoustic warnings at time to collisions (TTC) between 3.4 s and 1.7 s with initial warnings mostly issued between 2 s and 3 s. The point of system disengagement by TOR was defined by the time to lane departure at current speed without course correction, labelled as the time to line crossing (TLC) in the following. For takeover in a curve, TLC was defined by current speed and the curve’s radius, with no transition curves included by design. For the construction zone, the current speed and the distance to lane end defined TLC. The lane end was defined as the point where the s-shaped construction zone marking intercepted the centre of the rightmost lane.

Each of the three studies featured a *reference scenario*, allowing for data aggregation under similar conditions for analysis of user characteristics and sequence effects as well as quality analyses with a larger sample (Figure 5-2). The reference scenario was defined as an unpredictable, short-termed takeover with low speed (40 kph at TOR) in high traffic density (i.e., traffic jam) after three to five minutes of automation use with no secondary tasks available at the time of takeover. This scenario was labelled the reference scenario as the traffic jam context is closest to today’s application case of partial automation (e.g., ERTRAC, 2017; Gasser et al., 2012).

All simulator studies on short-termed takeover further featured one scenario under variation of context (see Table 5-1), either the duration of use (Study 1, Chapter 5.3), secondary task involvement (Study 2, Chapter 5.4) or speed and traffic density in the takeover scenario (Study 3, Chapter 5.5). Each participant experienced three drives, two of them including the use of automation, and two takeover situations in each drive. In total, four TOR were experienced by each participant. Two of these TOR occurred in the reference scenario, one with each hand posture. In the manual drive, the same driving setting as in the automated drives was implemented. Situational parameters likely underlie higher variance, for example regarding THW, in the manual drives and will be assessed for comparability during data analysis.

Table 5-1

Overview on the experimental variation of context in the driving simulator studies

	Duration of use	Traffic situation	Secondary task offer at TOR	Takeover scenario
Study 1	Short / Long (~ 15 min)	Traffic jam	None	Curve (400 m) / Construction zone
Study 2	Short	Traffic jam	None / With	Curve (336 m)
Study 3	Short	Traffic jam / Lead vehicle only (120 kph)	None	Curve (320 m) / Curve (720 m)

Note. Information on variations of context deviating from the reference scenario (i.e., TOR before entering a curve at 40 kph after a short traffic jam of about three minutes and without secondary task offers) is given in brackets. For each takeover scenario, the radius of the curve, defining the time available after TOR, is displayed in brackets.

For Study 1, two variations of context were realised, the first between subjects regarding the takeover scenario (curve and construction zone) and the other within subjects regarding the duration of automation use (short and long duration of use). The timeframe for successfully solving the situation (i.e., the TLC at which the TOR occurred) was kept constant between the two takeover situations (3.48 s), thereby enabling a comparison of takeover performance between the two scenarios. The TLC was chosen slightly above the upper limit of the suitable range for warning signals (Reinisch, 2012) to enable a lane change of the lead vehicle in the construction zone at the time of TOR. The lead vehicle began to change lanes 800 ms before a TOR was issued, resulting in a realistic lead vehicle behaviour within the simulation. This might have provided attentive drivers in this group with a slight advantage in preparation for the TOR compared to the curve scenario and was considered in data analyses.

To investigate a hypothesized earlier occurrence of OOTL performance problems after hands-off monitoring, a longer period of automation use in monotonous driving conditions (15 minutes) was compared to a shorter period (3 minutes) under similar traffic conditions. The long duration of use was based on that investigated in prior studies (e.g., Dogan et al., 2014; Othersen et al., 2014). To maximise differences between the conditions, the duration in the reference drive was set slightly lower (about three minutes) in Study 1 than in following studies (about five minutes).

For Study 2, the monitoring behaviour of participants was targeted by offering a secondary, non-driving related task during the drives. The curve radius was lowered, resulting in smaller TLC values (3.1 s). The general system instruction was complemented by an instruction on secondary task use, which was introduced as a side research project on display designs. However, instructions stressed that supervision of the automated system, respectively safe driving performance, was the main task during the drives. To help drivers focus on both goals, the secondary task was programmed as a self-paced visual-manual task, in which addresses had to be transferred from an address book into a navigation entry mask. The task was displayed on a 10.1-inch-tablet attached to the middle console of the vehicle for easy access and usability during manual driving as well as to enable control views onto the street. The task was offered two to three times at specified points during each traffic jam. In one of the traffic jams, the offer was activated 30 s before TOR. In all other cases (reference scenario), the last offer ended 2 min before TOR. Offer or task stayed active for 50 s. No timer or achievement count was provided to participants to discourage inappropriate focus on the secondary task. Tasks were offered in all traffic jams of this study to avoid TOR expectations based on secondary task offers in the within-subject design.

Focus of Study 3 was a variation of traffic conditions by speed and traffic density. Automation effects were compared after use in traffic jams against use at higher speeds with lower traffic density under the assumption of different user expectations and monitoring behaviour in the two traffic situations (Beggiato et al., 2015; Jamson et al., 2013). The TLC of the reference scenario was again slightly adapted. Adopting prior TLC for higher speeds in the curve scenario would have resulted in a road section with nearly non-perceptible curvature, limiting the subjective need for takeover. Following a different approach, the minimum radius for highway curves according to German road design regulations was chosen for TOR at 120 kph (i.e., 720 m; Arbeitsgruppe Straßenentwurf, 2008). The steering wheel angle necessary for driving through this curve when following the right lane's centre was then used to design the curve for comparison in the traffic jam scenario (resulting radius: 320 m). This procedure ensured a similar need for action, but the TLC values are not comparable between the two situations, with less time for successful takeover at higher speeds (1.7 s versus 3.3 s). However, both TLC values fell into the acceptable time range for warning signals (Reinisch, 2012). The speed limit on the track was set to 130 kph to keep drivers from overtaking the lead vehicle which was driving with 120 kph. The instruction was adapted to stress visible lane markings as a precondition for system activation, needed to justify a period of unassisted driving after takeover without the possibility for re-engagement of automation. The instructed functional range of the system in this study was between standstill and 130 kph with no lead vehicle necessary for activation of the system.

To take part in the studies, participants in all studies had to have at least used ACC or CC once. This was done to ensure a minimum of experience with entrusting a technical system with vehicle guidance as well as with supervising ADAS. Specifying minimal experience with longitudinal driver assistance was considered a trade-off between validity of driver interaction with a similar assistance system during a relatively short drive and enough variance within the sample to enable analysis of ADAS experience as an influencing variable. As age differences were not in the focus of the current studies, only participants between 25 and 65 years were invited, excluding those age groups with the highest accident risk (Vollrath & Krems, 2011). Other requirements were the possession of a valid driver's licence as well as naivety towards the study purpose, excluding participants from repeated participation in the three studies as well as participants from prior studies on partial automation.

5.1.3 Planned Takeover Situations

Focus of the first three studies was the performance in sudden, unexpected takeover situations, requiring immediate action for a successful handling. However, short-termed takeover scenarios represent a worst-case scenario, in which even attentive drivers cannot react in anticipation of a control transition. The a priori survey indicated, however, that drivers might adapt their monitoring behaviour to a level seen fit for the current level of system feedback, motivating a comparison of different systems regarding behaviour in evident, early-visible takeover situations. In a fourth simulator study, performance after short-termed takeover was therefore compared to that after anticipatable takeover, where anticipation related to the visibility of the system limit before a system-initiated disengagement by TOR. Closely related to the expectation or anticipation of control transitions is a valid mental model of the system used. Consequently, the possibility for driver initiated takeover before TOR and the mental model needed to be accounted for in system design and instruction (see Appendix 12.4).

ADAS experience was not explicitly required for participation in the study on expected TOR. Instead, similar to the approach of the survey, detailed initial information was used to establish a mental model of partial automated systems before first use. This procedure approximated the rental car scenario of the survey, investigating how informed users prepare initially for takeover situations and whether preparation behaviour differs between feedback conditions.

As one of the main objectives of this study was the analysis of differences between planned (early visibility of limit) and sudden (late visibility) system limits, the number of TOR-repetitions per participant was kept at a minimum to reduce overall expectation of TOR or the recognition of scenarios. Sudden TOR situations were masked by the lead vehicle as in prior studies and meant an immediate takeover necessity at TOR. Hand position respectively automation type (hands-on, hands-off and manual driving) was varied between three groups of subjects. Anticipation of TOR was realised as a within-subject factor, resulting in two different takeover situations per participants. The construction zone was selected as a suitable, easy recognisable, takeover scenario. Easy to instruct in the initial information, a construction zone should immediately be identified as a relevant situation for a control transition when approached. Similar to Study 1, a construction zone that required a lane change and included temporary lane markings was placed on the right lane, thus incorporating two clearly recognisable system limits for users.

As a correct mental model is the predisposition for anticipation of system limits, participants received a manual-like instruction of system limits in the beginning of this study (see Appendix 12.4), similar to the approach taken by Beggiato (2015), establishing comparable initial mental models by pre-drive information. The function described in the manual was based on current state-of-the-art systems. Instructed system limits were simplified and adapted to fit the takeover scenario described above as well as the function used for testing. The manual contained a detailed description of the general functioning of PAD, including the HMI with TOR-announcement, speed limits of the function, orientation in traffic, activation and deactivation, system limits and hand posture during PAD. It also stressed the driver's responsibility to monitor the system throughout the whole drive. Instructed system limits comprised of the inabilities of the system to detect vehicles on the neighbouring lane, thus forcing the driver to monitor and prevent overtaking on the right, and to conduct automated lane changes, making it necessary to deactivate the system whenever a lane change was necessary. Finally, the instruction stressed system uncertainty in case of temporary, missing or dirt-covered lane markings, resulting in a raised possibility of a TOR being issued. In cases where the system detected an obstacle that necessitated an evasive manoeuvre, control would immediately be transferred to the driver with notification by TOR.

Reading the manual was followed-up by a questionnaire on a priori attitudes and on the capabilities of the system, to ensure a correct mental model (as suggested by Seppelt & Victor, 2016; see Chapter 5.2.1). Another questionnaire, administered in the beginning of the study, focused on demographics. After reading the manual, a familiarisation drive together with the instructor stressed correct activation as well as correct hand posture and included a driver-initiated deactivation of the system, but no TOR. Furthermore, handling of the vehicle was practiced before activating the system in about five minutes of manual driving. As before, participants were told to never overtake slower vehicles except when told to and to always use the rightmost lane available. For the manual group, the familiarisation drive was used to stress the importance of keeping the constant distance of one reflexion post between the own and any lead vehicle. Participants in this group were told that they would be testing an automated system in a second drive and given a cover story necessitating a constant following-distance in the first experimental drive. This was done to ensure similar THWs in the relevant measurement situations for automated and manual drives. After the familiarisation drive, the main experimental drive, described in detail below, followed. Afterwards, participants again indicated attitudes and trust in automation, as well as anticipation of TOR, actions taken in case of TOR and their intention to use a similar system in daily traffic.

The main experimental drive took place on a two-lane highway section (Figure 5-3). Other vehicles apart from the lead vehicle were following in large distance to not influence lane change behaviour in the TOR situations. Signs limiting velocity to 120 kph were placed at the beginning of the drive and again after each TOR situation. No signs announced the upcoming construction zone and lane end. The system adapted its velocity after encountering the lead vehicle (large van) travelling with 100 kph and kept a THW of 2 s. THW was increased for this study to time the visibility of the TOR situation and the lead vehicle's lane change. In difference to the other simulator studies, the reference for THW was the centre of gravity (COG) of the own vehicle instead of its front. If participants attempted to overtake the lead vehicle, the experimenter, as before supervising hand posture and simulation on a separate screen, instructed the return to the rightmost lane to ensure reproducible visibility of TOR-situations during the drives. Two different track designs were used, one featuring the early-visible situation first and one featuring the short-termed, late-visibility situation first. Each design was applied to half of the participants of each hand posture group. The only difference between the two TOR-situations was the behaviour of the lead vehicle. To enable anticipation of the upcoming construction zone, the lead vehicle changed lanes 7 s before the participant's vehicle reached the lane end when travelling at 100 kph (early-visibility takeover). Based on the results by Mok et al. (2015) and Damböck et al. (2012), more than 4 s of prediction time (i.e., the critical event being visible to the driver) should enable a successful and comfortable takeover without secondary task application. The lead vehicle completed its lane change 6 s before the own vehicle reached the lane end. If the driver did not intervene, the own vehicle continued travelling at 100 kph until 3.1 s before reaching the lane end, where a TOR was issued and the system was deactivated.

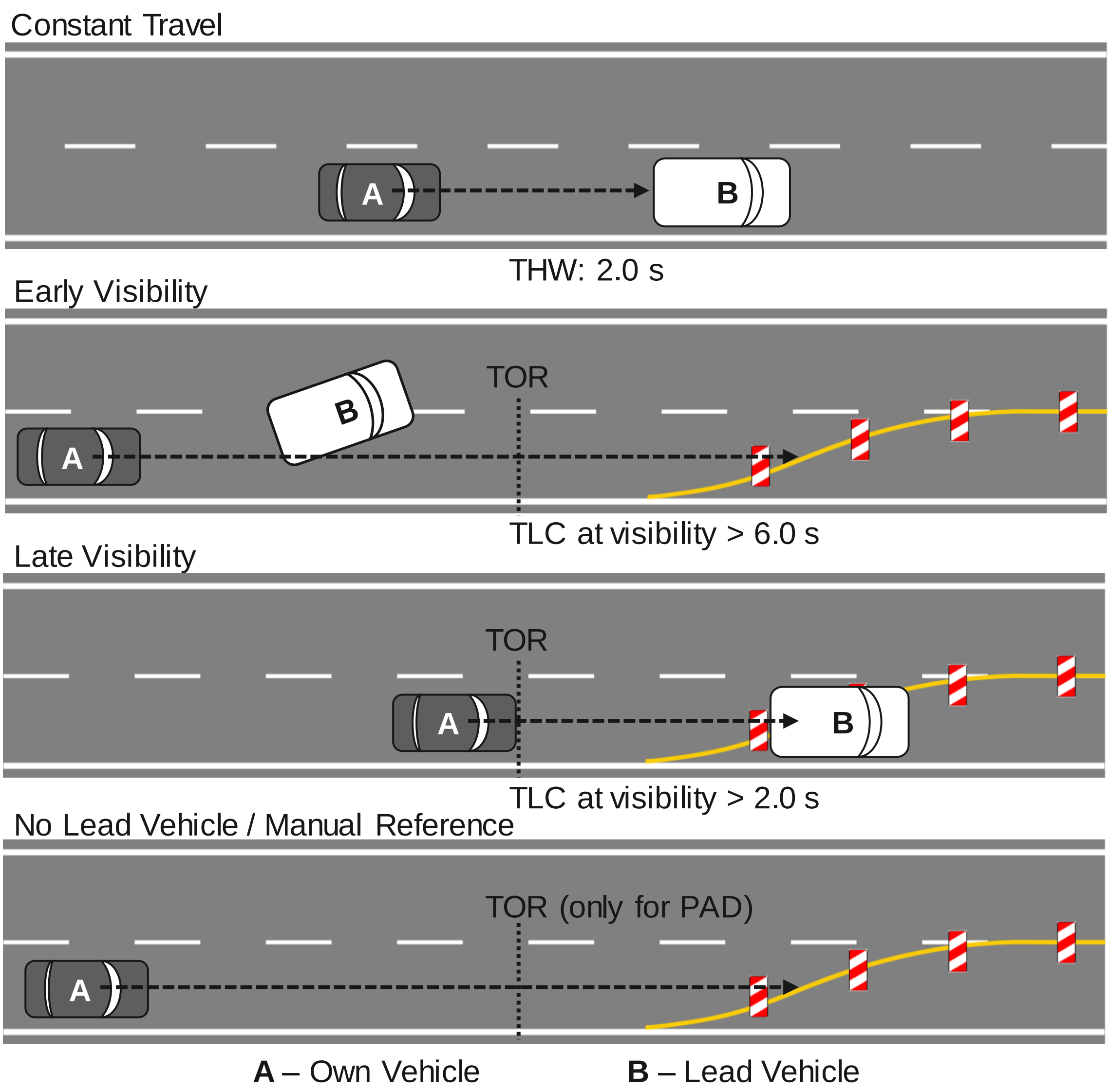

Figure 5-3: Lead vehicle behaviour during automated drives and in takeover scenarios with lead vehicle (early visibility and late visibility). The scenario without lead vehicle was used as the manual reference and driven with automation only by the manual group.

After passing the construction zone, the lead vehicle stopped on the kerb and was replaced with a different lead vehicle shortly afterwards for the second automated driving section. A yellow s-shaped lane marking stretched over 100 m and was additionally lined with traffic cones and poles to enforce a timely lane change.

For the late-visibility, immediate takeover scenario, the lead vehicle did not change lanes but ran over the traffic cones and came to a full stop well behind the lane end. Full braking of the lead vehicle was initiated 2 s before the lane end[6]. At the time the TOR was issued, the construction zone markings as well as the first traffic cone were visible from the driver's perspective. The lead vehicle passed the first traffic cone 0.6 s after the TOR was issued. Thus, at 2 s before the lane end, no lead vehicle blocked the view onto the construction zone. However, due to the s-shape with traffic cones, the construction zone was already visible at the time the TOR was issued (3.1 s until lane end). This timeframe is similar to those used before in studies on short-termed takeover and hypothetically enabled the initiation of an immediate, but manageable lane change in the non-automated condition, although no warning was issued in the manual drive. For this TOR situation, the lead vehicle remained stopped behind the traffic cones and was replaced with a different lead vehicle after the construction zone.

In addition to the main experimental drive, all participants conducted a manual lane change without a lead vehicle in a final short drive (Figure 5-4). Participants drove on the right hand lane towards a construction zone with 100 kph, given the instruction to 'change lanes at the last possible moment before the lane change manoeuvre would become critical'. No speed limit was included in the construction zones of this study to enable the manipulation of system limit anticipation and to focus on lateral guidance for the analysis of driving performance. In this drive, manual lane change data from each participant was collected in a non-critical manoeuvre to serve as a baseline for quality comparisons and for comparison on lane change initiation for drivers who did not initiate a control transition before TOR. Participants of the designated manual driving group without system use in the main experimental drive tested a hands-off automated system in a third, additional short drive. Visibility of the necessary takeover situation and lane change was maximised by omitting the lead vehicle to observe strategies and timing of actions in a completely uncritical takeover situation. Participants of the manual group received all information on automated

[6] The behaviour of the lead vehicle in this situation was not realism-oriented, but combined the affordances to create a situation of higher urgency that enabled the detection of takeover need at TOR, while keeping the remaining drive comparable to the early-visibility takeover in terms of THW.

driving prior to the according drive, but did not conduct a familiarisation drive with the system beforehand.

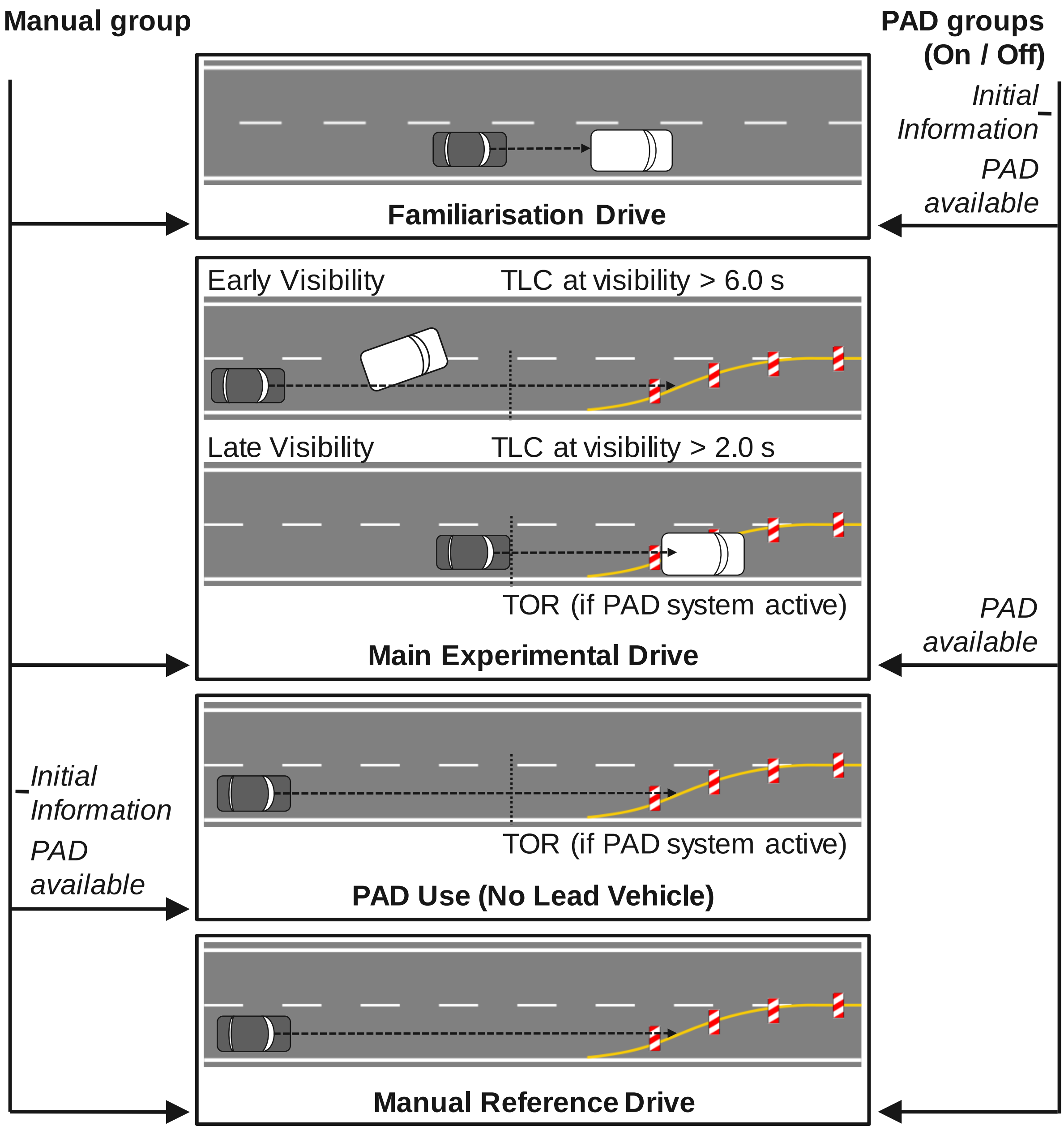

Figure 5-4: Experimental drives for the three participant groups in Study 4 (one manual and two automation groups). The manual group used automation once without lead vehicle. The sequence of scenarios in the main experimental drive was balanced over participants.

The same vehicle type, system set-up, HMI symbols and sounds as in prior studies were used. Deactivation thresholds were adapted for this study as driver initiated deactivations were part of the research focus. The system could be deactivated by pressing the instructed button (button for pre-selection of radio-station), or by either depressing the brake (by 40 % maximum pedal travel) or the accelerator pedal (by 20 % maximum pedal travel), as well as by steering input. Due to an error in the software version implemented for testing, the steering input required to deactivate the system was set to a difference of 6° between requested and actual steering wheel angle, instead of an originally intended lower value for easier de-activation.

5.2 Measures and Performance Indicators

Following the research model (Chapter 3), indicators were defined according to the separate phases of use, here defined as before use (a priori), during use and after use of automation (a posteriori; see Figure 5-5). The phase during use consists of the automated drive itself, where the user acts as a supervisor of the system, the control transition, where the system hands back control to the driver, and the manual driving phase following the control transition.

As the role of the user changes during phases of use, specific indicators are needed to adequately describe the interaction behaviour, including gaze metrics, metrics for takeover initiation and quality as well as metrics for the quality of manual driving performance. The track design of each study was planned to maximise comparability between experimental conditions, for example, regarding driving task complexity, curvature, speed of lead vehicles or traffic density. The comparability between automation conditions and variations of context is especially important for the analysis of driving performance (5.2.4) and regarding gaze data (5.2.2). Relevant for comparability of data within the first category are, amongst others and depending on the indicators used, variations of speed and road curvature (Knappe et al., 2006). Data of the second category requires comparable driving context variations such as lead vehicle presence (Morando, Victor, & Dozza, 2019). In the following, indicators used within this work are described, starting with subjective metrics for a comparison of attitudes and preferences before and after use.

Phases of User-System-Interaction		Subjective Measures	Objective Measures	
After Use		- Acceptance - Attractiveness - Comfort of Use - Intention to Use - Effort - Drowsiness - Trust		
During Use	Post AD		Post AD Driving Performance: - Mean Lateral Position (MLP) [m] - SD Lateral Position (SDLP) [m] - SD Steering Wheel Angle (SDST) [deg]	Performance During Manual Driving
	Control Transition	- Criticality of Control Transition	- Intervention Time [ms] - Type of First Driver Input (Longitudinal / Lateral) [-] Quality of Manoeuvre: - Lane Departure [-] - Maximum Steering Wheel Angle [deg]	
	During AD		Monitoring Behaviour: - Percent Road Centre (PRC) [%]	
Before Use		- Acceptance - Attractiveness - Trust - Mental Model		

Figure 5-5: Measures used for evaluation (AD: automated driving; SD: Standard deviation). Performance during manual driving served as the baseline for comparisons on driving performance, handling quality and monitoring.

5.2.1 Attitudes Before and After Use of Automation

For the measurement of change in preferences and attitudes through use of the system (see Chapter 7), questionnaires were applied before and after using each of the two systems (hands-on and hands-off). Usefulness and satisfaction with the system were measured using a five-point semantic differential (scored from -2 to 2 with higher values indicating more positive attitudes; Van der Laan et al., 1997). On the same scale, the attractiveness of the system was measured. Both acceptance (i.e., usefulness and satisfaction) and attractiveness are analysed regarding changes through use as well as general system and user preferences. Intention to use, respectively the intended frequency of use if the system was available, and comfort of use were collected after use of each system, allowing for a comparison between systems and different user groups. To further include the manual drives of the within-subject studies for the analysis of the effect of automation, the experienced effort and the drowsiness induced by the drive were collected after automated and manual drives. After each automated drive, participants were asked to rate the criticality of each takeover situation, indicating the influence of context variations on takeover criticality. All ratings were collected on a 16-point scale divided into six verbal categories (Figure 5-6), similar to the scale applied by Buld et al. (2002) or the 15-point scale applied by Skottke et al. (2014) for similar purposes.

User characteristics were collected in a pre-questionnaire in the beginning of each study. Next to age, gender and general driving experience (defined by kilometres driven per year), ADAS experience was measured on the same four point scale as in the survey (see Chapter 4.1 and Appendix 12.5) for four systems considered to be relevant prior experience (i.e., CC, ACC, LKA and TJA). Short system descriptions of these four ADAS, stressing tasks executed by the system, were provided to participants in order to ensure a correct assessment of the own prior experience. An ADAS experience sum was calculated for each participant of the within-subject studies (Study 1-3). Note that to be invited for participation, contrary to the online survey, users were required to have at least once used CC or ACC in traffic before (except for Study 4) and to be in possession of a valid driving license (see also 5.1).

gar nicht	sehr wenig			wenig			mittel			viel			sehr viel		
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Figure 5-6: German version of the 16-point-scale used in the questionnaires with six verbal sub-categories. Verbal categories translate to *not at all* (0), *very little* (1-3), *little* (4-6), *average* (7-9), *much* (10-12) and *very much* (13-15).

For the between-subject design (Study 4), user groups were additionally compared regarding trust and their mental model of the presented system, as both likely influence the interaction with the system in situations where the driver needs to decide whether to disengage or not. For measurement of trust, the scale by Jian, Bisantz and Drury (2000; German translation by Beggiato, 2015) was used. The mental model was measured using statements about the system, similar to the questionnaire designed by Beggiato (2015), with statements being adapted to the system at hand and rated either as ‘true’ or ‘not true’. The statements included lane change abilities and lane marking detection as the two system limits encountered in the study. Other items included the reported awareness of driver responsibilities (7-

point scale with higher values indicating higher awareness), and, as in the within subject studies, acceptance (Van der Laan et al., 1997) and attractiveness of the introduced function (5-point scale with higher values indicating higher attractiveness). Trust and mental model serve primarily as validation criteria for appropriateness of between-sample comparability.

5.2.2 Gaze Metrics During Use of Automation

No driver control input is available during automated driving. Instructions as well as data validation criteria for hand posture (see 5.2.5) prohibit variance in activation times of the system, which might be interpretable in regard to trust and system preference. Thus, the monitoring behaviour of the driver is most indicative for effects of context and system during use. Although interaction with vehicle controls is not necessary, monitoring of the surrounding traffic situation is still required by definition of Level 2 automation (SAE, 2018) and failure to do so may prohibit the timely detection of erroneous system behaviour or critical driving situations. Gaze metrics have accordingly been used to classify drivers' monitoring behaviour, SA and vigilance (Körber, Cingel, Zimmermann, & Bengler, 2015; Feldhütter et al., 2017; Louw et al., 2015; Othersen et al., 2014).

Morando et al. (2019) compared on-path and off-path glance durations as well as percent road centre (PRC) to detect changes in driver attention when using longitudinal and lateral assistance. Significant changes between automated and manual driving were found for PRC and for the proportion of on-path glances. PRC was also found to be sensitive to driving context, specifically to time-of-day and car following (Morando et al., 2019), as well as to changes in driving task complexity (Victor, Harbluk, & Engström, 2005) and cognitive demand during secondary task involvement (Wang, Reimer, Dobres, & Mehler, 2014). Generally, PRC has been considered a sensitive indicator of driver distraction (Ahlstrom, Kircher, & Kircher, 2009). Additionally, PRC has been associated with less effort in calculation, due to the necessity for data filtering and segmentation (Morando et al., 2019; Wang et al., 2014), than other measures such as off-path glances. Whereas it has originally been defined as "the percentage of gaze data points labelled as fixations during a task that fall within the road centre area" (Victor et al., 2005, p. 181), it can also be calculated from raw data, as was suggested by Victor et al. (2005). Proof for this suggestion was provided in Ahlstrom et al. (2009) and Wang et al. (2014). No distinction has to be made between fixations and saccades, as negligible differences in PRC were found for approaches using unsegmented data (Ahlstrom et al., 2009). Thus, the analysis of monitoring behaviour was conducted using the percentage of valid gaze data points towards the road centre during a specified driving interval. As a normal distribution cannot be assumed for PRC (Morando et al., 2019) and ceiling effects can be expected (Wang et al., 2014), non-parametric tests were used for analysis.

For automated drives, the analysis interval used in most analyses began with the last activation of the automated system before the TOR and ended 500 ms before the TOR was issued. The endpoint was defined to exclude influences of situation recognition from analysis, although variation in track design and surrounding traffic should likely prevent recognition effects. For manual drives, where no system activation defined the interval of analysis, the sample's mean system-activation time in the corresponding condition was selected for comparison. Exceptions from this definition, as for secondary task engagement, are described in the according analysis sections. Additionally, vehicle speed was controlled to ensure inclusion of the correct traffic condition, such as speeds below 55 kph for traffic jam conditions. Due to the balancing of conditions over the sample, influence of track design on PRC, especially by curvature or lead vehicle presence, is largely prevented.

For the defined interval, all recorded data points directed at the road ahead were filtered using a quality indicator provided by the measurement software (Smart Eye Pro by Smart Eye AB). From the remaining data points, the road centre was calculated for each participant in each drive and initially defined as a circle of 8° radius (Ahlstrom et al., 2009) around the mode of valid data points (cf. Victor et al., 2005). The percentage of valid data points within the road centre based on the overall number of valid data points was then calculated. As the analysed interval ranged from less than a minute to up to 15 minutes, depending on the experimental condition, large variations in the number of overall valid data points were expected. Although PRC results have been considered robust for different calculation durations (Wang et al., 2014), data sets were excluded from analysis if the z-value – calculated for each study separately over all conditions and for Study 1 separately for traffic jam durations – was smaller than $z = -2$. Under the assumption of comparable lengths of intervals within conditions, very small numbers of valid data points indicate a general problem with tracking quality for the according drive.

5.2.3 Performance During Control Transitions

In the following, indicators for assessment of takeover performance during control transitions are detailed. The focus for this phase of interaction is twofold, including aspects of timing and the general success of the takeover. Measuring the time it takes the driver to execute the first vehicle input is one of the most frequently referred to indicators of takeover performance. It has been used for analysis of performance decrements after automated and assisted driving for a variety of conditions and systems (see, e.g., TORrt in Eriksson & Stanton, 2017; for other examples, see also

Happee et al., 2017; Körber et al., 2016; Naujoks et al., 2015; for ACC, see Stanton & Young, 1998).

For timing of driver actions after TOR, the first driver control input (FDI) after automated driving, also referred to as the “intervention time” (Happee et al., 2017, p. 211), was analysed. Different thresholds have been suggested by different authors, depending on scenario and context (for examples, see, e.g., Feldhütter et al., 2017; Happee et al., 2017; Zeeb et al., 2015). Steering wheel angles after TOR were analysed for Study 1 to find an adequate threshold that indicates FDI in short-termed takeover situations. Only data sets with comparable steering wheel angles at TOR, indicative of a comparable system behaviour and thus lane position, were included. The selected threshold needed to be low enough to capture initial corrective steering input. For pedal input, separate thresholds were defined. The first threshold detected, either steering or pedal input, defined the intervention time. FDI could only be compared between automated drives, as no reference behaviour for measurement of changes in driver input was available for manual (curve) driving.

For Study 4, due to the differences in visibility of the takeover situation, FDI was defined by deviations from input of the automated system collected 100 m before the point at which the lead vehicle changed lanes (scenarios without lead vehicle with expected earlier interventions: 200 m). For Study 4, timing was referenced onto the according trigger point, defined as the point where the lead vehicle began to change lanes in the early-visibility condition or the point where the TOR was issued in the late-visibility condition. For the short-termed takeover studies, FDI was referenced onto the point of system disengagement (TOR).

Planned lane changes on highways, in difference to short-termed takeover situations at low speed, need not necessarily go along with large changes in steering wheel angles. In some cases in a study by Happee et al. (2017), an “obstacle was successfully avoided with peak steering wheel angles as low as 3°” (p. 213). Thus, an adapted threshold for FDI likely has to be used for expected control transitions in the construction zone scenario in Study 4. Additionally, the moment of lane change initiation (LCI), being also available for manual driving data, could be used in this study for comparison of timing and success of driver behaviour. Takeover initiation after automated driving (i.e., FDI with adapted thresholds) and LCI were compared in a first step for data of Study 4. Aim of this comparison was a decision for or against an analysis using only LCI as the relevant indicator for timing of actions in Study 4. The decision was based on differences found in the timing of LCI and FDI after automated drives.

In contrast to FDI, marking the first driver decision to influence vehicle guidance, LCI marks the point in time where the driver begins a lane change to the left lane or otherwise reacts to the lane end by significant braking. The threshold defined needs to differentiate between corrective steering behaviour on straight road segments in manual driving and the lane change initiation. Going back from the lane change, defined as the point when the COG passes the lane marking, the last point in time with a steering wheel angle above the defined threshold and a positive alteration rate (steering to the left) was identified. The threshold for manual driving was defined as the individual mean steering wheel angle within a 100 m interval plus two standard deviations when driving straight. The threshold after automated driving was the same as for detection of first steering input under additional consideration of the alteration rate. To further capture different driver strategies such as braking before LCI, brake pedal depression by more than 10 % (Happee et al., 2017) was considered when analysing the timing of actions.

The timing of interventions is only one aspect influencing the success of the takeover, with late interventions raising the risk for critical driving situations. Many studies have additionally analysed the performance in the seconds following first input, thereby including the quality of the takeover manoeuvre to estimate the impact of a control transition on driving safety (e.g., Happee et al., 2017; see also de Winter et al., 2014). The simplest (binary) quality classification is the overall success of the manoeuvre, defined as the occurrence of collisions with surrounding objects or other traffic participants as well as lane departures due to late or inappropriate interventions.

In TOR situations where few or no collisions or lane departures occur, quality metrics (labelled “surrogate safety metrics” by Happee et al., 2017, p. 212), need to be defined for a meaningful comparison of conditions. The timeframe for quality analyses begins with automation disengagement, but goes beyond the initial takeover manoeuvre. For all studies conducted, the immediate success of the takeover manoeuvre, defined as a control transition without lane departures, is reported next to the timing and type of first driver input after TOR, complemented by a description of driver input during control transitions. Manual drives served as the comparison baseline for success of the manoeuvre, following up on the aim to construct time-critical, but manageable control transitions. Additionally, manual baseline drives were considered in more detail regarding the quality of driving performance during and after takeover, especially steering performance, as detailed in the following.

5.2.4 Performance After Control Transitions

Literature has shown that automation influences driving performance for longer than the takeover initiation with its potential for lane departures and collisions resulting mainly from late driver interventions. Prior studies have therefore also investigated the duration and magnitude of effects on post-automation driving performance (e.g. Merat et al., 2014; Morgan et al., 2018). Within this work, driving performance after automation was compared to manual drives within the same context of use to define the end of corrections after a control transition (cf. Eckstein, 2015; Petermann-Stock et al., 2015; for a similar approach, see Louw et al., 2015).

For a valid comparison of driving performance between conditions, any additional influences on driving performance need to be controlled for. As described for the study set-up, the investigated scenarios were designed to induce steering reactions for successful takeover, targeting the difference in hand posture between automation conditions. Measures of lateral acceleration were not considered within this work due to using a static driving simulator in most studies. Instead, measures underlying visual control by the driver, such as lane offset, and indicators of direct driver control input were preferred. Steering performance and consecutive lane keeping performance were of special interest for the analysis of differences. Relevant influences on lane keeping are road curvature and speed (for an overview, see Knappe et al., 2006). Within the studies conducted, road curvature was defined by the track design. The behaviour of the simulated surrounding traffic restricted variation in speed. Nonetheless, mean speed was analysed descriptively for the selected analysis intervals with regard to differences in speed between conditions. Curvature was completely comparable between at least two of the three drives per driver (i.e., same track design) and did never incorporate narrow curves, thereby establishing similar road conditions between all drives of each participant. Together with proof of similar driving speeds in each condition, road curvature can be assumed to be comparable in each condition for the different analysis intervals.

Additionally, measures of lane keeping performance have been found to be driver dependent. SDLP was found to “differ greatly between individual drivers” (Verster & Roth, 2011, p. 363), but to be “a stable measure within subjects across time” (p. 363). Similarly, the mean lateral position (MLP) reflects a general driver strategy and is most insightful when compared to a baseline drive of the same driver (Knappe et al., 2006). Similar to the need for the analysis of driving impairment under visual occlusion or substance abuse (as reviewed by Brookhuis, de Waard, & Fairclough, 2003), performance after automated driving and in manual drives was compared within participants instead of between samples, using each driver’s manual driving performance within the same scenario as the normative baseline.

The estimation of automation effects on driving performance after takeover within this work was assessed for different timeframes. Performance in the first seconds after TOR captured missing input from the driver by inclusion of the time until driver intervention as well as the initial takeover, whereas the manual drive provided data on uninfluenced manual driving within the corresponding road section. Maximum steering wheel angels ($steer_{max}$) were used to describe the quality of the initial takeover (cf. Kerschbaum et al., 2014). Analyses over longer timeframes were conducted to describe the hypothesized harmonization of performance indicators after potential initial differences have subsided (i.e., the end of correction; Eckstein, 2015).

For short-termed takeover situations, the analysis interval began either with the deactivation of the automated system (t_0) or at the corresponding track coordinate for manual drives without automation. For the lane change scenario in Study 4, driving performance was analysed starting with the transgression of the COG into the new lane for better comparison between manual and automated drives after self-initiated lane changes. Additionally, $steer_{max}$ was analysed in the first three seconds after deactivation of the system for automated drives similarly to short-termed takeover situations. The maximum interval for analysis was set to 40 s based on track design. A shorter timeframe was chosen if one of the following events occurred earlier than after 40 s, namely the re-activation of automation, the exiting of the highway at the end of the drive or passing the end of the construction zone that was followed by a lane change to the right.

Initial differences in takeover handling were compared using $steer_{max}$ during the first three seconds of the analysed interval, covering the timeframe for a successful takeover in short-termed scenarios and reflecting the execution of initial input respectively the lane change[7]. SDLP and the standard deviation of steering wheel angle (SDST) are reported for increments of 3 s for a timeframe of 12 s overall, using four analysis intervals, thereby including the initial takeover and subsequent driver performance. Metrics were calculated for intervals beginning with the first measurement point within the defined timeframe up to the last measurement point before the beginning of the next interval. Thus, the first interval began with the first measurement point after deactivation and lasted until the last measurement point earlier than 3 s after deactivation (resulting in an interval duration of about 3 s).

SDST reflects direct driver input, therefore being especially sensitive to differences in steering input, whereas SDLP, covering the vehicle's reaction to driver input, is directly interpretable in terms of driving safety (cf. Neubauer et al., 2014). Focus of this analysis was the identification of differences persisting after the initial driver takeover. For SDLP, higher values reflect "poorer vehicle control" (Neubauer et al.,

---

[7] Based on TLC values, successful steering interventions need to occur within 3 s after TOR. The type and timing of driver interventions were validated before analysis.

2014, p. 2055) and SDST has been interpreted in the same manner, with larger variation reflecting a less controlled driving behaviour (Brookhuis et al., 1991; Mok et al., 2015). However, the chosen interval length of 3 s might not allow for a meaningful investigation of SDLP. As was pointed out by Verster and Roth (2011), “if the length of the segment approaches zero, SDLP does too” (p. 365). Nonetheless, slightly shorter intervals than those used in prior research (e.g., 5 s as in Merat et al., 2014), were chosen within this work to capture initial differences as well as possible fast convergence between conditions. Data of Study 1 was used to decide whether SDLP should be discarded as a metric because of the expected lack in variation, in which case the analysis should focus solely on SDST and MLP instead.

MLP, reflecting driver strategy, is reported in 3 s increments for a 40 s interval as the focus was on the (descriptive) harmonization of driving performance and the stability of the harmonization point in time over different contexts. Initial differences in MLP were expected not only due to delayed driver input after automation, but also due to differences in curve cutting, as the automated system positions the vehicle exactly in the centre of the lane before TOR. In a first step, SDLP, SDST and MLP were analysed descriptively for the studies on context variation. Significance tests, after a first descriptive assessment, were applied to the reference situation for which a larger sample, aggregated from all studies on short-termed takeover, was available (see Chapter 6).

For curve driving, cases of lane departures were excluded from further analysis of quality and were classified as non-successful takeovers instead. All other takeover cases were analysed further using the defined metrics. For the construction zone, lane departures, other than the one necessary for keeping the target course, were excluded and classified as non-successful takeovers. Lane departures after the necessary lane change, reflecting a timely driver intervention with insufficient follow-up quality, are reported separately, but included in the analysis of driver performance.

5.2.5 Data Analysis and Aggregation

To capture the effect of context variations on interaction, data of each study was analysed separately considering all phases during use (see Figure 5-5). The same indicators were additionally used for the analysis of individual user differences and sequence effects (see Chapter 6). For some analyses, additional data validation using specific indicators was undertaken, such as for the analysis of secondary task engagement, and is described in the according results sections.

Means were calculated participant-wise for the according automation condition (hands-on or hands-off) before calculating sample means, if not stated otherwise. Validity criteria for inclusion of datasets into behavioural analysis were the correct (i.e., the instructed) hand posture during system monitoring, the correct system status in the measurement situation, as well as comparable steering wheel angle, THW and speed at TOR. Hand posture had to be correct either the moment the situation became visible, that is, when the lead vehicle began to change lanes, or the moment the TOR was issued.

The level of significance for all analyses was set to $\alpha = .05$. Post-hoc t-tests following up on significant main effects of ANOVAs are reported under adjustment of alpha-level (Bonferroni-Correction). Greenhouse-Geisser-corrected values are reported in case of violation of sphericity. Other adjustments are reported where applicable. Questionnaire results and gaze data were analysed using non-parametric procedures where indicated by scale or data distribution.

5.3 Study 1: Duration of Use

The first driving simulator study[8] focused on behaviour after long and short system use during traffic jams (factor *duration*: *short* versus *long*) with hand posture and automation use varied within participants (factor *automation*: hands-*on* versus hands-*off* versus *man*ual). The takeover scenario was varied between subjects (factor *scenario*: *con*struction versus *curv*e). Problems with the measurement equipment lead to a number of data sets being excluded from analyses. For ten participants – three of them assigned to the curve scenario – hand posture or interaction with the system was found to be against instructions in at least one of the situations of interest. These participants were excluded from data analysis. Two of these cases were excluded due to deviating steering wheel angles in the hands-on condition, indicating that the driver was working against the system. Violations in the hands-off condition were identified using the video recordings of hand posture. One participant was included in the analysis with driving data for PAD only.

Steering wheel angles, lane position and vehicle speed at the TOR trigger point were examined. For PAD, no deviations from the defined system behaviour (0 m lane offset, 40 kph, 0° steering wheel angle) were found. For manual driving, minor variations were found with mean lane offsets ranging between 0.2 m (SD = 0.3) and -0.55 m (SD = 1.5) after short and long traffic jams before curve and construction zone. Mean steering wheel angles were slightly higher for the construction scenario (short: M = 6.4°, SD = 6.4; long: M = 4.7°, SD = 4.6°) than for the curve scenario (short: M = 3.9°, SD = 1.5, long: M = 3.7°, SD = 1.3). Mean distances to the lead

[8] Data of this study can be found in adapted form in Josten et al. (2016).

vehicle were around 25 m before entering the curve and around 17 m when entering the construction zone.

Of the 38 participants (5 female) included in analysis, $n = 22$ were assigned to the curve scenario. Mean mileage per year was 20750 km (*SD* = 12900) with 61 % of the overall sample using highways at least once a week or more often. Mean age was 41 years (*SD* = 13). Participants were licensed since on average 23 years (*SD* = 13; minimum = 7), 97 % had driving experience with CC systems and 37 % had used ACC systems at least once.

Continuous activation of the system before the takeover request was verified to ensure a clear difference between long and short durations of use. Mean activation time in long traffic jams was 14:40 min (*SD* = 00:07 min; minimum = 14:19 min) and in short traffic jams 2:58 min (*SD* = 00:08 min; minimum = 2:30 min; maximum = 3:08 min). Video analysis revealed that hand posture was as instructed almost all of the time for participants included in the analysis.

As described in the method section, a threshold had to be defined for steering wheel angles indicative of takeover initiation. As drivers could not foresee takeover situations, no driver-initiated control transitions had to be taken into account. Thus, the threshold served to indicate the first driver input deviating from the system regulated steering wheel angle of 0° at TOR, while taking into account possible changes in steering wheel angle due to re-positioning of the hands on the steering wheel. Earlier work has considered steering input of 2° as onset of steering (e.g., Feldhütter et al., 2017; Louw et al., 2015). However, to better differentiate steering wheel angles for curve negotiation and lane change onset in low speed scenarios from values of straight, manual driving, a higher threshold of 4° was considered a better indicator of active steering input.

To define the relevance of the exact detection threshold, mean initiation times after hands-on and hands-off for each participant were compared with different criteria (i.e., 2°, 4° and 6°). In every takeover situation, each participant set steering wheel angles of more than 6°. Thus, none of the limits was considered unfeasible per se. Mean differences between steering wheel angles were 110 ms between 2° and 6° (60 ms for 4° compared to 2°) for intervention in the hands-off condition and 160 ms (100 ms) for the hands-on condition. The continuity in steering wheel angles set for hands-off and hands-on conditions was further reflected in the similar differences between feedback conditions in timing defined by the three thresholds (mean difference between off and on; $\Delta_{2°}$ = 339 ms, *SD* = 264; $\Delta_{4°}$ = 299 ms, *SD* = 267; $\Delta_{6°}$ = 287 ms, *SD* = 269). Analyses on takeover initiation were therefore conducted with a threshold of 4° steering wheel angle. A comparison with timing of brake pedal depression (5 % of maximum pedal travel) revealed steering initiation to be the first active driver input after TOR in all cases, as intended by scenario design.

A slightly higher number of hand posture changes before TOR occurred in the construction zone condition, as reported above for exclusion of data. Although the time available for takeover after TOR was the same for both scenarios (TLC = 3.48 s), behaviour of the lead vehicle could have given an earlier cue for takeover in the construction zone (see 5.1.1). Furthermore, the unambiguousness and decision for necessary actions could have been different between the two scenarios, as safe lane changes include the awareness of vehicles on the left lane. Takeover initiation (FDI) was therefore compared between the two scenarios to decide whether takeover complexity influenced the timing of actions and whether to aggregate samples into one dataset. For hands-on, no significant difference for mean takeover times was found between scenarios, $t(36) = 0.46$, $p = .645$ ($M_{con} = 649$ ms, $SD = 246$; $M_{curv} = 681$ ms, $SD = 184$)[9]. The same was observed for hands-off drives, $t(36) = 0.12$, $p = .904$ ($M_{con} = 961$ ms, $SD = 292$; $M_{curv} = 971$ ms, $SD = 190$). One common analysis with data from both scenarios is thus considered feasible with the type of takeover scenario not influencing intervention times.

A repeated-measures ANOVA revealed a significant effect of takeover initiation with longer intervention times after hands-off monitoring, $F(1, 37) = 45.23$, $p < .001$, $\eta_p^2 = .55$ ($M_{\text{on}} = 668$ ms, $SD = 273$; $M_{\text{off}} = 967$ ms, $SD = 312$). The main effect of duration of use was significant only by tendency with slightly later steering initiation after longer durations of use, $F(1, 37) = 3.24$, $p = .080$, $\eta_p^2 = .08$ ($M_{\text{short}} = 781$ ms, $SD = 263$; $M_{\text{long}} = 854$ ms, $SD = 322$). The interaction was not significant, $F(1, 37) = 0.01$, $p = .958$.

Takeover was successful in all cases for curve handling. For the construction zone, two lane departures were found, both occurring after hands-off driving. In one case, the driver paused his input after a timely steering initiation. In the other case, the driver carried out a disproportionate counter-steering manoeuvre at the end of the lane change leading to a lane departure. This manoeuvre, being less safety-critical than the other case, resulted from the inadequate handling of the takeover. In contrast, the first case seems to indicate mode confusion, that is, a misinterpretation of the automation’s state and the shift in responsibilities.

---

[9] Differences are analysed using t-tests instead of one 2x2 repeated-measures ANOVA due to violation of the equality of covariances.

On average, participants reported to be less surprised by the takeover when driving hands-on in the curve scenario (M_{curv} = 6.64, *SD* = 3.96; M_{con} = 8.59, *SD* = 4.98), but equally surprised in both scenarios for hands-off automation (M_{curv} = 9.05, *SD* = 2.95; M_{con} = 8.88, *SD* = 4.41). All ratings were clustered within the same verbal category (“little”), which might be related to the prior presentation of the TOR in HMI-familiarisation and system instruction. Criticality ratings were similar between the two takeover scenarios as well (z = 0.64, p = .524) and therefore analysed over the complete sample. A significant effect of hand posture on perceived criticality was found (z = 2.73, p = .006; M_{on} = 5.21, *SD* = 4.09, *MD* = 4.50; M_{off} = 7.25, *SD* = 4.04, *MD* = 7.50) with hands-off takeover resulting in increased subjective criticality compared to hands-on takeover situations. No effect of the duration of use (z = 0.483, p = .625; M_{short} = 6.32, *SD* = 3.50, *MD* = 6.25; M_{long} = 6.14, *SD* = 3.60, *MD* = 5.75) was found for the subjective criticality of control shifts. Thus, data indicate a disadvantage for missing feedback for both scenarios and durations of use regarding timing (FDI), success as well as subjective criticality.

Gaze behaviour (PRC) was compared between the two traffic jam durations (long and short traffic jam) for the overall sample as driving conditions did not differ outside the takeover scenario. Due to technical problems with the eye tracker and overall quality issues for some data sets, data of 24 participants entered analysis. Only the main effect of duration of use as well as the effect of feedback on monitoring during intervals of longer use was considered. The main effect of automation (man, on, off) was analysed using the aggregated data sample (see Chapter 6.2). No significant effect in PRC was found between the two durations of automated driving (z = -1.20, p = .230; M_{long} = 0.61, *SD* = 0.14; M_{short} = 0.63, *SD* = 0.12).

Taking a closer look at monitoring during longer traffic jams, PRC values did neither differ between feedback conditions for the first half (z = -1.57, p = .116; M_{on} = 0.64, *SD* = 0.12; M_{off} = 0.61, *SD* = 0.18) nor for the second half of use (z = -1.59, p = .112; M_{on} = 0.61, *SD* = 0.12; M_{off} = 0.56, *SD* = 0.17). This result does not support a change in monitoring strategy over time related to variations of haptic feedback. Additionally, as an equal decline over time in mean PRC was observed for unassisted driving ($M_{first\ interval}$ = 0.78, *SD* = 0.20; $M_{second\ interval}$ = 0.75, *SD* = 0.18), larger changes in gaze attribution to the forward road scene under longer automation use compared to unassisted driving cannot be assumed based on data of this study.

Post-automation driving performance after different durations of automated driving was analysed with an emphasis on the curve scenario. The construction zone was passed by drivers in on average 14 s, restricting the available interval for post-automation performance analysis due to lane changes conducted. A detailed analysis of driver performance on straight road sections after lane changes, compared to curved road sections in the other studies, is undertaken using data from Study 4 instead (see Chapter 5.6).

As all first driver responses were steering responses and occurred within the first three seconds after TOR, analysis of $steer_{max}$ in the previously defined interval (t_0 to 3 s) to describe the initial takeover was considered adequate. Differences between durations of driving were rather small (difference in median $steer_{max}$ for long versus short traffic jam; Δ_{on} = 3.99°; Δ_{off} = 1.54°; Δ_{man} = 0.35°). More apparent were differences between automation conditions with much higher $steer_{max}$ after automated driving (median of single responses per drive, i.e., two responses per participant per condition; MD_{man} = 9.90°, MD_{on} = 23.71°, MD_{off} = 26.53°).

Before analysing driver performance, post-automation mean speed was compared within the relevant driving intervals to account for speed as a possible reason for differences in steering behaviour. Mean speed (see Figure 12-1 provided in Appendix) increased over time, as the traffic jam dissolved after TOR, but was very similar between conditions (maximum difference between means: 2.59 kph), therefore justifying a comparative analysis between conditions in combination with the comparable road curvatures of each condition (see Chapter 5.2.4).

Increased variation in steering input (SDST) after automation, as compared to that of continuous manual driving, was found for the interval including takeover initiation (0-3 s; see Figure 5-7) and persisted, although becoming notably smaller thereafter, until up to 6 s after takeover, at least after hands-off monitoring. The mean of individual differences in SDST between the manual condition and post-automation fell below 1° after hands-on monitoring in both traffic jam durations within the second analysis interval, whereas similar differences were reached after hands-off monitoring one interval later (after 6-9 s of manual driving).

Furthermore, the individual variation in SDST over the four intervals was notably smaller for manual driving (mean of the individual standard deviation in SDST over the four analysis intervals; $M_{SD\ man}$ = 0.78°) than after automated driving ($M_{SD\ on}$ = 3.45°; $M_{SD\ off}$ = 4.45°). Low variation between intervals indicated a stable and continuous steering behaviour in manual driving conditions, whereas post-automation steering input underlay greater variation not only between *conditions* (i.e., higher *absolute* SDST), but also between *intervals* over time (i.e., higher *variation* of SDST).

The variation in steering wheel position (SDST), capturing fast changes in driver input, was larger than the effects observed for the lane position (SDLP), as can be seen in comparison of Figure 5-7 and Figure 5-8. The change of variation in the manual drive between analysis intervals ($M_{SD\ man}$ = 0.02 m) and within the single intervals was extremely small (mean SDLP $_{man}$ in both traffic jam conditions between 0.04 m and 0.06 m). This is likely due to the short measurement intervals in time and, at low speeds, in distance covered. Verster and Roth (2011) stated that “the reason for lower SDLP values in shorter segments is that the MLP has less opportunity to show large changes if the distance is shorter” (p. 365). Nonetheless, as was pointed out in Chapter 5.2.4, short intervals were chosen deliberately to capture fast convergence between conditions after initial differences, which seems even more justified after the analysis of SDST.

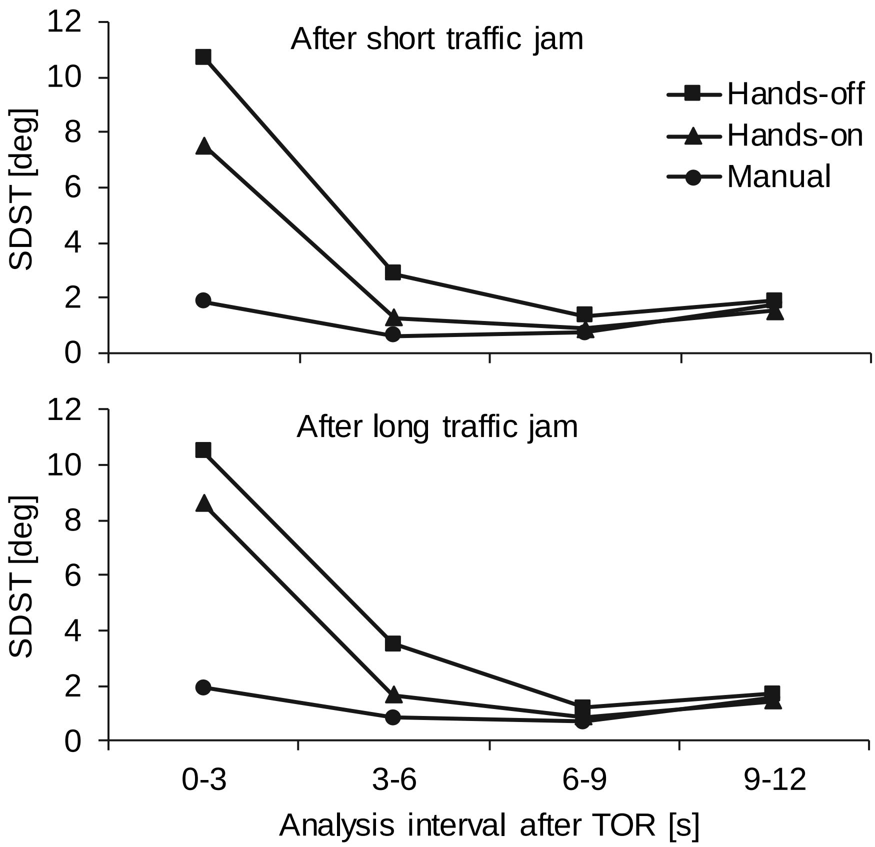

Figure 5-7: Mean SDST in the curve scenario aggregated over 3-s-intervals beginning at TOR after short (upper graph) and long (lower graph) traffic jam driving for the three automation conditions (*n* = 21).

Greater changes in lane position were observed for the first two intervals after automated driving (largest mean SDLP measured 3-6 s after long traffic jam with M_{on} = 0.15 m and M_{off} = 0.21 m) and for the mean standard deviation of SDLP between intervals ($M_{SD\ on}$ = 0.06 m; $M_{SD\ off}$ = 0.07 m). However, in light of the minimal variation in SDLP, SDST seems to be the more sensitive and adequate measure for the current analysis purpose and the timeframes considered. An analysis of longer intervals for SDLP, increasing overall variation in lane position, does not seem feasible as differences between conditions seem to be rather short-termed (see Figure 5-7).

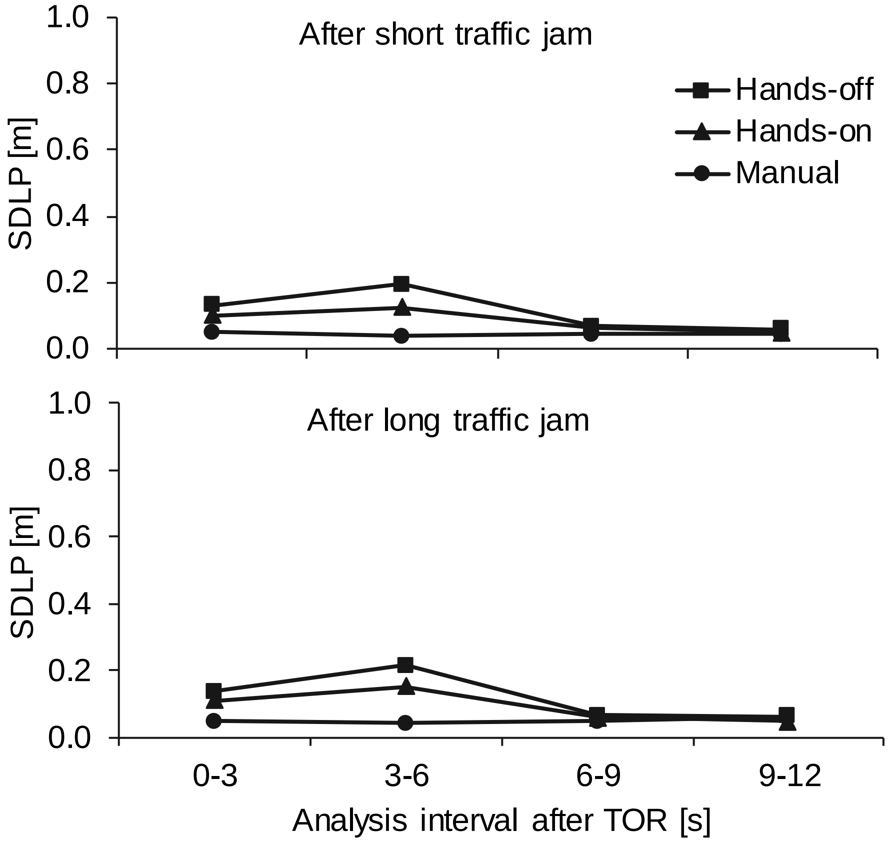

Figure 5-8: Mean SDLP in the curve scenario aggregated over 3-s-intervals beginning at TOR after short (upper graph) and long (lower graph) traffic jam driving for the three automation conditions (n = 21).

The additional analysis of MLP, reflecting the driver's strategy (Knappe et al., 2006), showed smaller changes in MLP over analysis intervals for continuous manual driving (mean of individual SD in MLP over 13 single intervals: $M_{SD\ man}$ = 0.18 m) than after automated driving ($M_{SD\ on}$ = 0.25 m; $M_{SD\ off}$ = 0.27 m). Figure 5-9, depicting the MLP in each interval and condition, suggests that the effect of automation, similar to what was already shown for SDST and SDLP, ceased descriptively after about 6 s (hands-on monitoring in short traffic jam) to 12 s (in all other conditions) of manual driving. The difference to manual driving was descriptively larger after longer use of automation. Furthermore, a general change in driver strategy as a possible carryover effect of automation, such as a more centred lane positioning, was not apparent in data. Instead, drivers seem to quickly resume their individual driving strategy after automation.

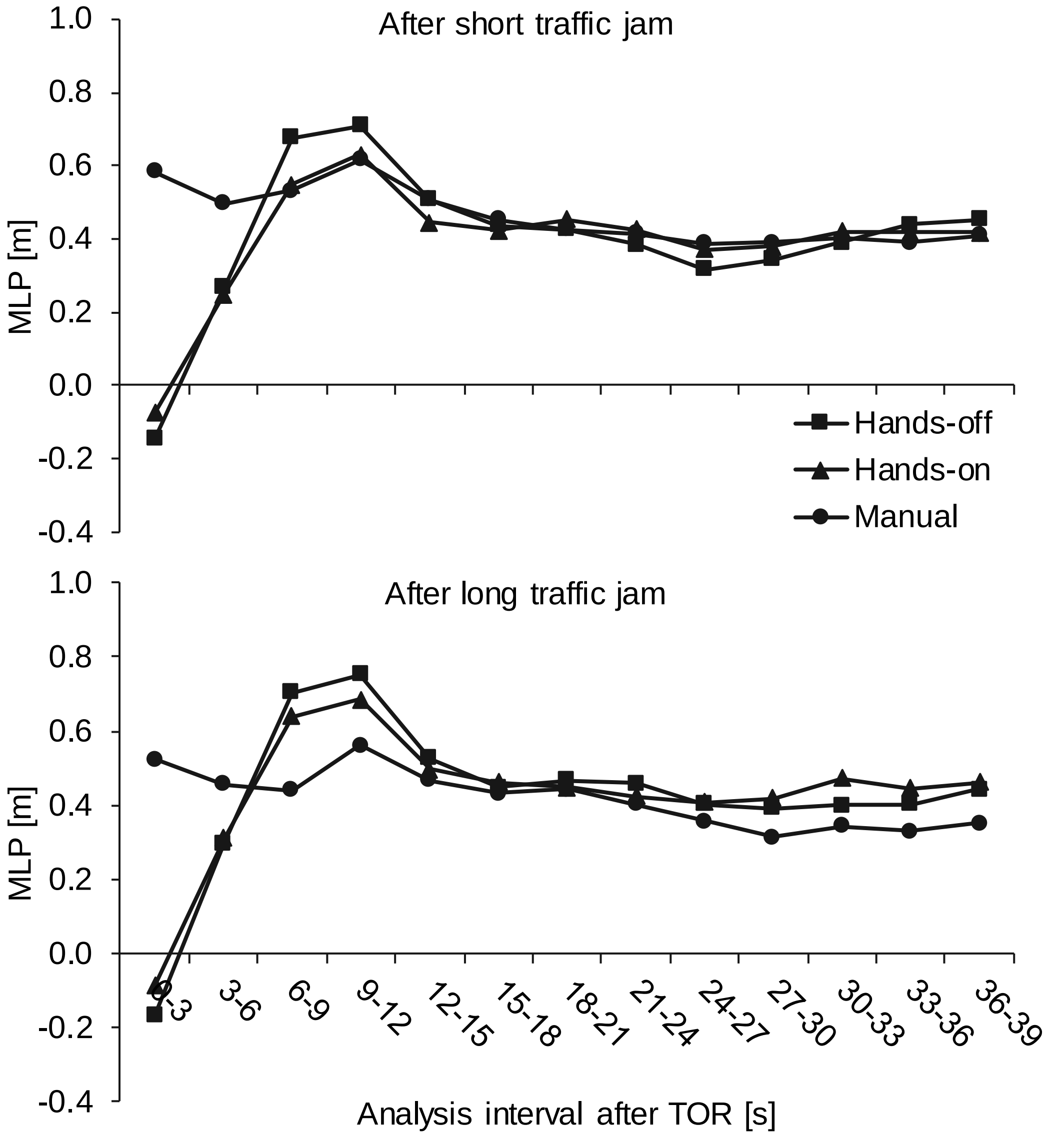

Figure 5-9: MLP aggregated over 3-s-intervals (sample means for *n* = 21) beginning at TOR in the curve scenario after short (upper graph) and long (lower graph) traffic jam driving for each of the three automation conditions.

Together with the larger maximum steering wheel angles ($steer_{max}$) found for the initial takeover, the overall larger steering variation and the more persistent differences to the baseline after hands-off monitoring indicate a less controlled takeover process than after hands-on monitoring. This trend was the same for both monitoring durations.

5.4 Study 2: Secondary Task Engagement

In the second driving simulator study[10], a self-paced secondary task was offered before (TOR *without*) and during (TOR *with*) one of the takeover situations (curve scenario) while driving with each PAD system. To define two distinctive settings of monitoring quality for analysis of takeover performance, involvement with the monitoring task had to be controlled for this study using eye-tracking measures and video footage of the driver. Next to taking-up on secondary task offers and driving with the instructed hand posture, drivers in this study were required to monitor the driving task in measurement situations without secondary task offers to be included in data analysis. Opposed to the monitoring condition, discontinued monitoring was determined when the last fixation before a TOR was one onto the tablet with the navigational task or – in case of low eye-tracking quality – video footage showed participants in interaction with the secondary task or gazing out of the side windows of the vehicle at occurrence of the TOR.

Of the 20 participants who were tested, seven were excluded from analysis of takeover performance. One of them was excluded because of the wrong hand posture before the takeover situation and another one due to technical difficulties with the secondary task. Further excluded were two participants who did not take up on offers of the secondary task during the TOR situation, one who did not monitor the system before the TOR, when he should have, and one who constantly worked against the system by steering wheel movements. For one participant, only PAD will be analysed as the secondary task offer was declined in the relevant situation of the manual drive.

Of the remaining 13 participants (one female), seven participants drove with the hands-off system first. Mean age was 29 years (*SD* = 6). Participants reported to travel on average 10423 km (*SD* = 6843) per year. Participants had on average 11 years of driving experience (*SD* = 6) and 38 % travelled on highways at least once weekly. All participants had experience with CC systems and 62 % had used ACC systems at least once. Inclination towards secondary tasks during driving was high for tasks similar to the one selected for this study. Only 7.7 % of the sample reported to never interact with their navigational system during drives (38.5 % very seldom) and 23.1 % reported to interact only very seldom with their entertainment system (all others more often).

[10] Data of this study can be found in adapted form in Josten et al. (2016).

Time of PAD use was the same for hands-on and hands-off with around 4:40 min of automated driving before each TOR. Data screening revealed no variation of expected vehicle behaviour at TOR (speed 40 kph, lateral lane offset 0 m, steering wheel angle 0°). For manual drives, small deviations to PAD in steering wheel angle (M = 3.43°) and lane position (M = 0.70 m) were observed. Vehicle speed varied with secondary task availability in manual drives, with lower speed during secondary task engagement (M_{with} = 38.72 kph, SD = 3.90; $M_{without}$ = 41.06 kph, SD = 2.08).

Interaction time with the secondary task was estimated by analysing video footage and noting beginning and end of task input by the driver as well as phases of orientation away from the task, that is, orientation away from the tablet combined with no input made, in relation to the overall duration of the offer. Overall, the involvement with the secondary task was high for the remaining 13 participants of the study. For only 10 % of all offers, participants worked on the secondary task less than 50 % of the allotted time, including the manual drives[11]. No difference in average involvement with the secondary task was found between feedback conditions with both 80 % of the available time used for interacting with the task. Encountering a TOR during an active secondary task led to slightly lowered overall engagement with the task (6.8 % shorter engagement hands-on, 12.6 % shorter engagement hands-off). No difference for engagement was found for manual drives (< 1 %), where no TOR interrupted the engagement in the measurement situation. Five participants in the hands-on drives (seven for hands-off) did cease engagement with the secondary task before the task offer ended in case of a TOR, compared to one in each condition during an offer without TOR. For the manual task, earlier disengagement from the secondary task was observed in four cases in the measurement situation and in three cases during additional offers. In two cases, the offer was completely neglected. For manual drives, secondary task involvement led to compensation behaviour in the driving task, as was already described for vehicle speed. More specifically, increased distances to lead vehicles were found for secondary task involvement compared to driving without secondary task (z = 2.98, p = .003; M_{with} = 28.79 m, SD = 9.75; $M_{without}$ = 16.07 m, SD = 4.23). Thus, although drivers also engaged in the secondary task during manual driving, they did so with consequences for the primary task (THW_{with} = 2.65 s; $THW_{without}$ = 1.41 s; $THW_{instructed}$ = 1.5 s). It is important to note, however, that the task was designed to be in line with the obligation to monitor the

[11] The high engagement with the secondary tasks in the automated drives as well as in the manual drive is not necessarily in opposition to the supervision or conduction of the driving task (see 5.1.2). The secondary task was explicitly designed to allow for self-paced secondary task input. Gaze metrics are needed to estimate the effect of secondary task engagement on the involvement with the driving task.

system and regularly checking the traffic situation was enforced by instruction and possible at all times.

All participants initiated a steering manoeuvre before any longitudinal adaptation of vehicle behaviour. Steering was initiated significantly later after hands-off monitoring than after hands-on monitoring, $F(1, 12) = 5.26$, $p = .041$, $\eta_p^2 = .305$ ($M_{on} = 750$ ms, $SD = 70$; $M_{off} = 955$ ms, $SD = 73$). No effect of secondary task involvement was found for intervention times, $F(1, 12) = 0.01$, $p = .932$ ($M_{with} = 850$ ms, $SD = 69$; $M_{without} = 855$ ms, $SD = 62$). Although hands-on driving resulted in descriptively faster takeover with secondary tasks ($\Delta_{with\text{-}without} = -62$ ms), whereas hands-off driving resulted in slower takeover with secondary tasks ($\Delta_{with\text{-}without} = 51$ ms), the interaction of secondary task involvement and hand posture was non-significant, $F(1, 12) = 0.31$, $p = .587$. All takeover situations were handled successfully in terms of the lane departure criterion.

In line with the findings on takeover interventions, subjective criticality differed little between secondary task conditions (hands-on: $M_{with} = 5.85$, $SD = 2.64$, $M_{without} = 6.31$, $SD = 3.75$; hands-off: $M_{with} = 8.00$, $SD = 3.74$, $M_{without} = 7.31$, $SD = 3.99$). The difference in criticality between feedback conditions was non-significant ($z = 0.54$, $p = .593$) albeit slightly higher criticality ratings in the hands-off condition.

As secondary tasks were offered in all traffic jams, with the only difference being an offer during takeover, no analysis for PRC during complete drives was conducted. The focus of gaze analysis lay instead on a comparison of changes in PRC with and without active secondary task offers. To this end, gaze behaviour was calculated as the mean individual PRC of two driving intervals with secondary task and two equally long intervals without secondary task with intervals being between 18 s and 50 s long. One of the two intervals ended 500 ms before a TOR was issued, preventing any, albeit unlikely, effect of scenario recognition on gaze behaviour. Due to the influence of secondary task engagement on overall data quality, data of 10 participants entered analysis of which two participants were considered with data of one interval per condition instead of two. The results showed a large effect of secondary tasks on PRC ($z = -2.80$, $p = .005$) with less gaze attributed to the road centre during secondary task engagement ($M_{with} = 0.31$, SD = 0.16; $M_{without} = 0.65$, SD = 0.16). As displayed in Figure 5-10, PRC was reduced under secondary task performance in all three conditions, but not to an equal degree (mean difference in PRC with and without secondary task; $X^2(2) = 6.20$, $p = .045$; $\Delta_{man} = 0.26$, $SD = 0.27$; $\Delta_{on} = 0.39$, $SD = 0.13$; $\Delta_{off} = 0.39$, $SD = 0.15$). In accordance with the non-significant difference in intervention time between feedback conditions, equal average changes were found however in the two automation conditions.

The steering input after TOR was again analysed using median values for $steer_{max}$ for each condition. The largest difference between variations of context occurred after hands-off driving (difference between median $steer_{max}$ with and without secondary task involvement; Δ_{man} = 1.69°; Δ_{on} = 2.51°; Δ_{off} = 6.01°). The differences between automation conditions were again much larger than those between the secondary task conditions (median of $steer_{max}$ values, with two responses considered per participant for each automation condition; MD_{man} = 12.31°, MD_{on} = 28.09°, MD_{off} = 34.02°).

The secondary offer stayed active until on average 20 s after TOR. Thus, an offer was provided for the complete analysis interval of SDST, but only for half of the intervals analysed for MLP. As shown above, the overall time attended to the secondary task suggests a continued engagement for at least some time after TOR in all conditions, although secondary task engagement decreased in combination with a TOR (i.e., in the automation conditions), whereas no change was found in manual drives. Again, mean speed was compared between conditions for the selected analysis intervals. As can be seen in Figure 12-2 (provided in Appendix), speed increased over time, but no considerable difference was found between automation conditions with a maximum speed difference between single interval means of 2.32 kph (off versus man without secondary task) and 3.76 kph (man versus on with secondary task).

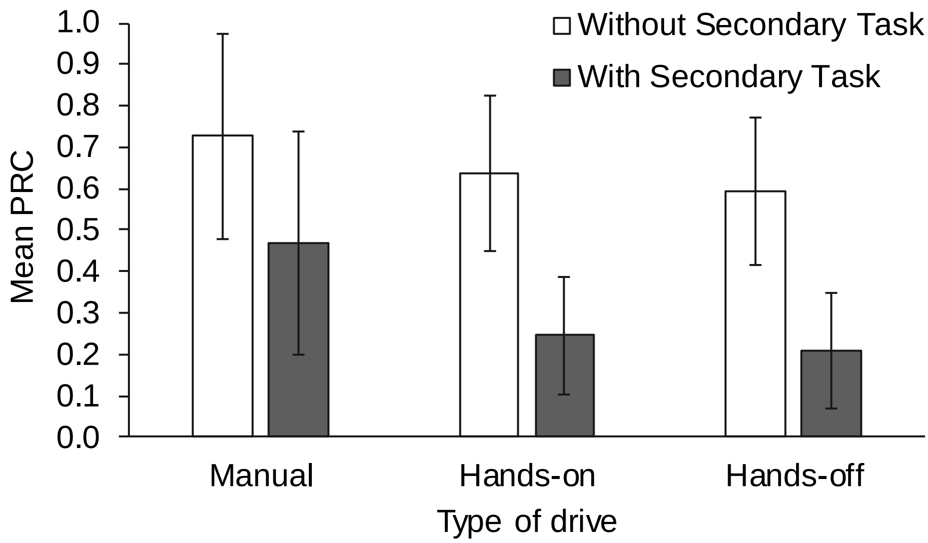

Figure 5-10: Mean PRC (and SD) for intervals with and without secondary task (n = 10) during automated (hands-on; hands-off) and unassisted driving (manual) in traffic jams. Takeover situations are not included in the data displayed.

For SDST, the same effect as described for Study 1 was observed with much higher steering activity within the first analysis interval, followed by rapid convergence towards baseline performance (see Figure 5-11). After hands-off driving, differences in mean SDST to the baseline were overall larger (mean difference to baseline in first interval; $\Delta_{\text{on-man}}$ = 6.76°; $\Delta_{\text{off-man}}$ = 10.20°) and persisted longer with secondary task engagement than without, in this case up until 6 s after TOR. The individual variation in SDST over the four intervals was again smallest for manual drives (mean of individual SD between intervals; $M_{SD\ man}$ = 1.36°; $M_{SD\ on}$ = 4.23°; $M_{SD\ off}$ = 5.70°), indicating the least variable steering behaviour over time. In combination with secondary tasks, an increase in steering variation in all analysis intervals was observed regardless of the automation condition. Visual-manual secondary tasks did thus influence steering performance, but effects were small and confined. Again, the changes were largest in magnitude and duration after hands-off monitoring.

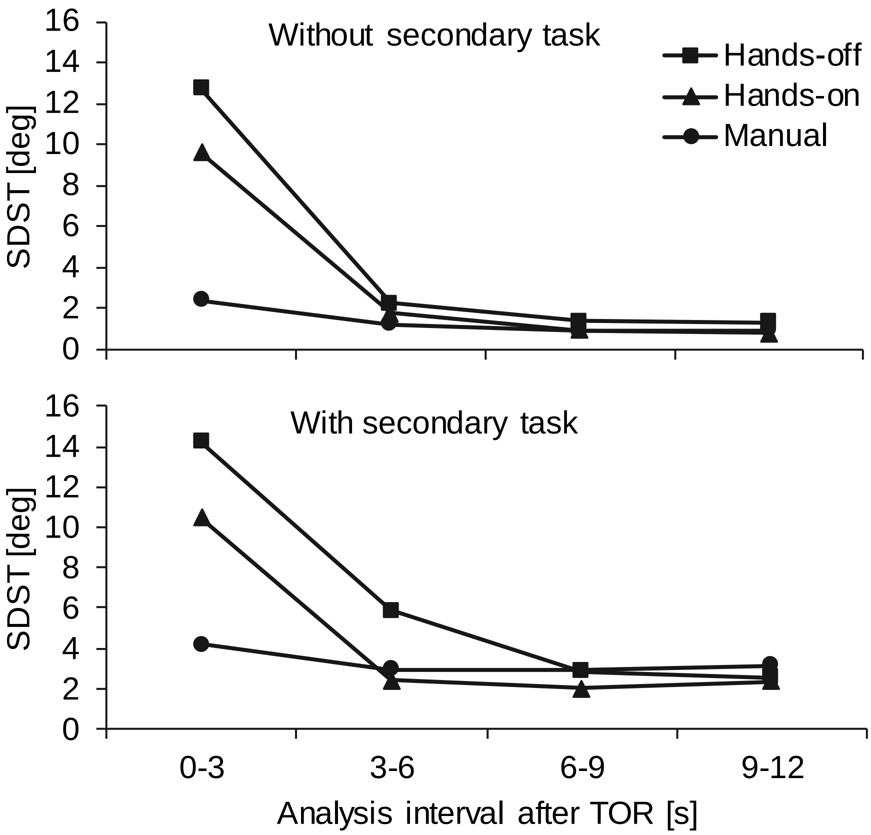

Figure 5-11: Mean SDST aggregated over 3-s-intervals beginning at TOR without (upper graph) and with (lower graph) secondary task for the three automation conditions (n = 12)

For MLP, variation over time (SD in MLP over intervals) was higher when a secondary task was offered at least part of the time, compared to driving conditions without secondary task offers. This effect was especially pronounced for the manual driving condition, were variation over time was comparatively low without a secondary task (manual: $M_{SD\ without}$ = 0.18 m; $M_{SD\ with}$ = 0.31 m; hands-on: $M_{SD\ with\text{-}}$

out = 0.20 m; *M* *SD with* = 0.27 m; hands-off: *M* *SD without* = 0.20 m; *M* *SD with* = 0.29 m). For the condition with secondary tasks, no defined point of convergence for MLP was apparent (see Figure 5-12), as MLP within the baseline condition was highly variable over time. This effect cannot be attributed to differences in road curvature that was explicitly considered as a relevant influencing factor in the design of the scenarios. Without secondary tasks, similar MLP between all conditions was observed after 12 s. Again, differences were less pronounced after hands-on monitoring.

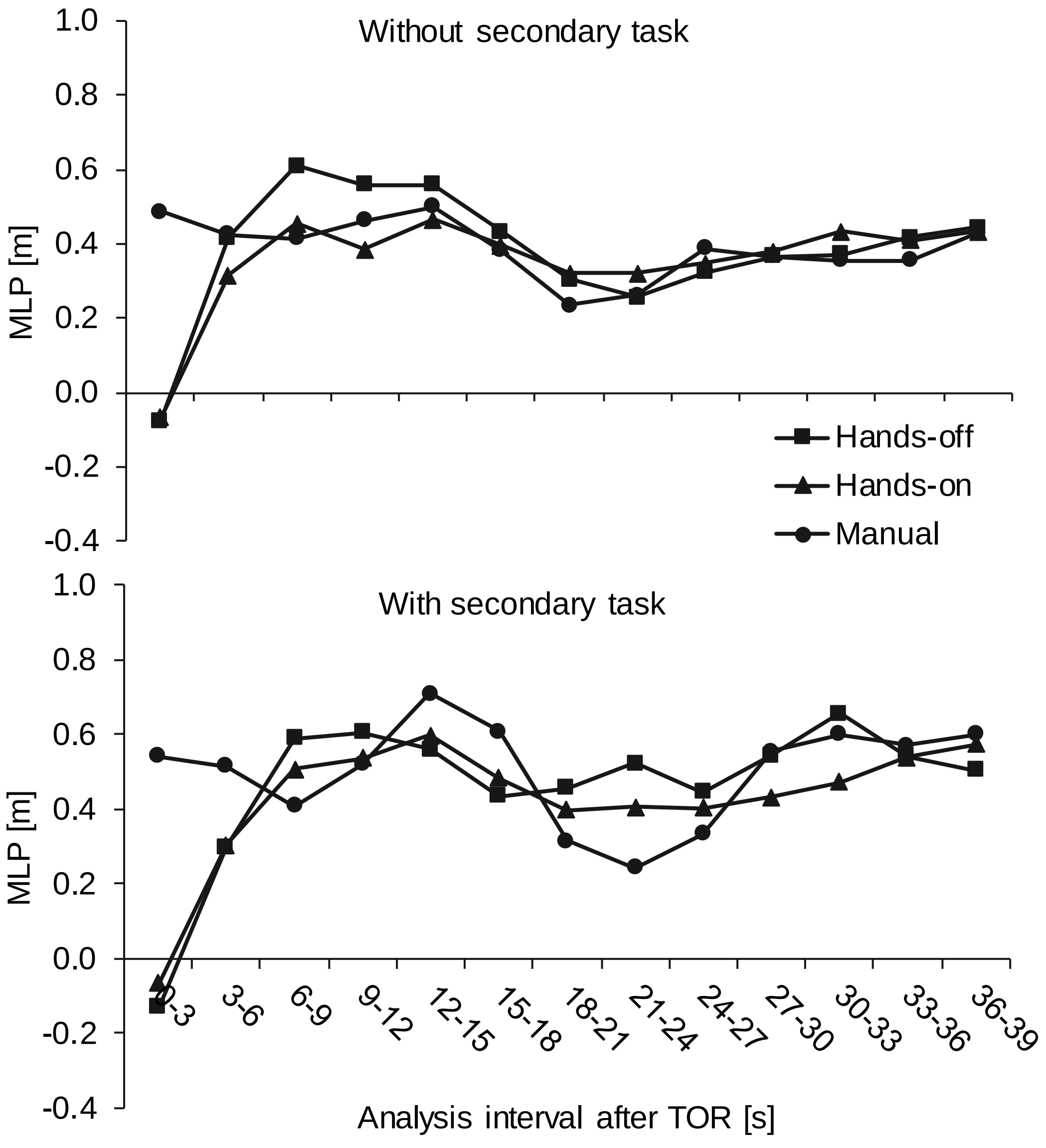

Figure 5-12: MLP for 3-s-intervals (sample means for *N* = 12) considered after TOR in the curve scenario without (upper graph) and with (lower graph) secondary task for each of the three driving conditions.

5.5 Study 3: Traffic Density

In the third driving simulator study[12], the complexity of the driving situation was varied within each drive in addition to automation (manual versus PAD) and type of PAD system (hands-*on* versus hands-*off*) between drives. Each drive featured a short traffic jam (*40* kph) which ended, as in the studies before, in a curve scenario. Additionally, a similar scenario at higher speed (*120* kph) with heavily reduced traffic density was implemented in each drive.

Of the 16 participants tested, one had to be excluded due to not following the hands-off instruction in one of the takeover situations. Hands-off was the first system to be tested for eight participants. Mean age of the sample was 27 years (SD = 2), three females took part in the study. Mean mileage per year was 10973 km (SD = 9885) and participants held their drivers licence for on average 8 years (SD = 2). All participants had driving experience with CC systems and 40 % additionally had experience with ACC systems. Of the sample, 53 % travelled on highways at least once weekly.

Duration of use was the same for hands-on and hands-off systems. The automated drive was on average about two minutes longer in the traffic jam condition than in the high-speed scenario (5:59 min compared to 4:04 min). System behaviour at TOR did not differ between hand posture conditions and corresponded to the defined specifics for the scenarios (for scenario 40 / 120: distance to lead vehicle 20 m / 52 m, steering wheel angle 0°, speed at TOR 40 kph / 119 kph, lane offset 0 m). For manual drives, deviations from system behaviour similar to those observed in prior studies occurred (mean values for scenario 40 / 120; distance to lead vehicle 28.92 m / 58.27 m, steering wheel angle 4.95° / 3.60°, lane offset 0.52 m / 0.48 m, speed 40 kph / 117 kph).

The first driver input measured after deactivation of the automated system was always a change in steering wheel angle. Steering was initiated significantly earlier in case of hands-on monitoring than for hands-off monitoring, $F(1, 14) = 23.39$, $p < .001$, $\eta_p^2 = .63$ (M_{on} = 546 ms, SD = 173; M_{off} = 848 ms, SD = 265). The traffic situation had no influence on takeover initiation, $F(1, 14) = 0.43$, $p = .524$ (M_{40} = 712 ms, SD = 45; M_{120} = 682 ms, SD = 47). Again, no interaction between context and feedback was found, $F(1, 14) = 0.79$, $p = .389$.

No lane departures occurred in the traffic jam scenario. For the high-speed situation, five lane departures were observed, four of them in the hands-off condition. Three of

[12] Data of this study can be found in adapted form in Josten et al. (2016).

the lane departures occurred in the first of the three drives per participant. Takeover in lane departure cases was on average initiated 323 ms (hands-on) and 485 ms (hands-off) later than the mean initiation in successful cases. In three of the lane departure cases, the vehicle barely entered the adjacent lane, one of them the hands-on case. The vehicle left the lane completely in the remaining two cases. All lane departures can be attributed to late steering initiation in combination with a shorter time frame for successful takeover, as defined by the lower TLC in the high-speed scenario. The five cases were excluded from the analysis of takeover quality.

In accordance with the number of lane departures, takeover at higher speed was rated significantly more critical than at lower speed ($z = 3.05$, $p = .002$; $M_{40} = 6.17$, $SD = 2.11$; $M_{120} = 9.37$, $SD = 2.63$). No difference between feedback conditions was found for criticality ratings ($z = 1.13$, $p = .258$; $M_{on} = 8.33$, $SD = 1.57$; $M_{off} = 7.20$, $SD = 3.41$).

The analysis of gaze behaviour (PRC) focused on the effect of context (high traffic density with low speed versus low traffic density with high speed; $n = 13$). The main effect of driving context was significant with higher PRC at higher speed ($z = 2.62$, $p = .009$; $M_{40} = 0.74$, $SD = 0.14$; $M_{120} = 0.78$, $SD = 0.12$). Analysing the automation conditions separately, no effect of traffic condition was found in manual drives ($z = 0.08$, $p = .937$; $M_{40} = 0.83$, $SD = 0.13$; $M_{120} = 0.85$, $SD = 0.10$). An effect of traffic condition was however found in PRC for both automated drives (hands-on: $z = 2.73$, $p = .006$, $M_{40} = 0.70$, $SD = 0.15$; $M_{120} = 0.76$, $SD = 0.14$; hands-off: $z = 2.15$, $p = .032$, $M_{40} = 0.69$, $SD = 0.18$; $M_{120} = 0.74$, $SD = 0.16$), with a higher percentage of gaze towards the road at higher speed. The significant differences between context conditions were descriptively larger than those between the two feedback conditions. The significance of influences of feedback and automation on PRC was a focus of the analysis using the aggregated data sample (see Chapter 6.3).

For the ten participants without lane departures, $steer_{max}$ was compared between contexts and automation conditions. Differences between medians for driving contexts were surprisingly small, with the largest difference again found for hands-off driving (difference between median $steer_{max}$ at low and high speed; $\Delta_{man} = 1.94°$; $\Delta_{on} = 1.92°$; $\Delta_{off} = 5.57°$). Differences between automation conditions showed the same pattern as in prior studies (median of single responses per drive, considering two responses per participant per condition; $MD_{man} = 13.32°$, $MD_{on} = 30.18°$, $MD_{off} = 33.63°$).

Data of ten participants was considered for the analysis of performance in the high-speed condition. Data of 15 participants was considered for the low-speed condition. Due to the higher speed and the accordingly shorter time until leaving the highway at the end of the drive, the analysis interval was shortened to 33 s for this study (i.e., to 11 instead of 13 analysis intervals).

For the low-speed condition, analysis of mean speed over intervals revealed no apparent differences between automation conditions (see Figure 12-3 in Appendix). For takeover at higher speed, a slight post-automation speed reduction (8 kph) was found. Relative to the overall speed, this difference was deemed negligible. Differences in mean speed between conditions became comparatively continuous after 12 s.

The analysis of SDST revealed the established pattern of takeover effects with large initial differences (see Figure 5-13). Although performance converged towards the baseline consistently for both traffic situations, convergence at high-speed took longer and resulted in larger mean differences until the last interval analysed in both feedback conditions (mean difference to baseline after 9-12 s; 40 kph: $\Delta_{\text{on-man}} = 0.12°$, *SD* = 1.40; $\Delta_{\text{off-man}} = 0.14°$, *SD* = 1.24; 120 kph: $\Delta_{\text{on-man}} = 0.80°$, *SD* = 1.03; $\Delta_{\text{off-man}} = 0.51°$, *SD* = 1.56).

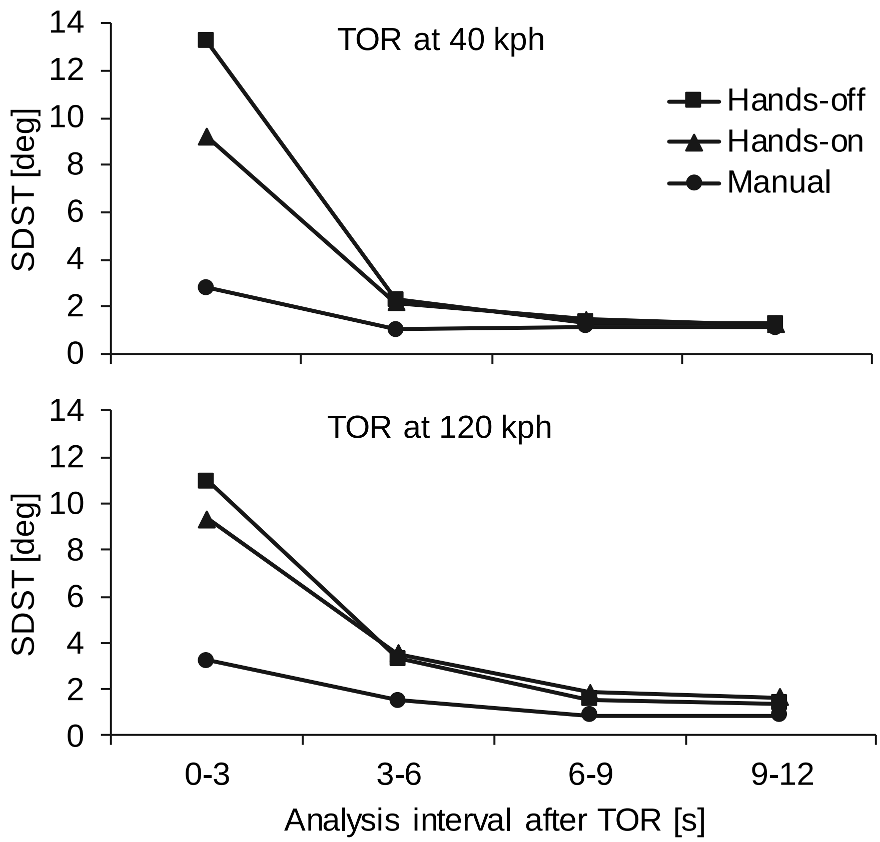

Figure 5-13: Mean SDST aggregated over 3-s-intervals beginning with TOR at low speed (upper graph; *n* = 15) and high speed (lower graph; *n* = 10) for the three automation conditions.

An effect of driving speed on initial differences in steering variation was observed only in the hands-off condition, whereas the difference to manual driving was unchanged by speed after hands-on monitoring (mean difference to baseline after 0-3 s; 40 kph: $\Delta_{\text{on-man}}$ = 6.43°, *SD* = 3.65; $\Delta_{\text{off-man}}$ = 10.50°, *SD* = 5.22; 120 kph: $\Delta_{\text{on-man}}$ = 6.10°, *SD* = 3.30; $\Delta_{\text{off-man}}$ = 7.75°, *SD* = 4.60). The specific influence of speed on SDST in the hands-off condition, in which the least controlled takeover behaviour was frequently observed, might be due to greater steering wheel forces at higher speed that hinder large absolute as well as large changes in steering wheel angles.

Due to the influence of different track designs for the two traffic scenarios, mostly by road curvature, a comparison of (absolute) MLP between low and high speed condition was not conducted. For the three driving conditions (man, on and off), the influence of road curvature on MLP was contained by driving with comparable speed on roads of similar curvature, allowing for a comparison of differences between conditions.

Although MLP in the already familiar low-speed condition showed descriptively larger and more persisting differences in this sample than in prior studies, performance in the high-speed condition was even more variable (see Figure 5-14; mean of individual SD in MLP between intervals; manual: $M_{SD\ 40}$ = 0.19 m; $M_{SD\ 120}$ = 0.24 m; hands-on: $M_{SD\ 40}$ = 0.19 m; $M_{SD\ 120}$ = 0.29 m; hands-off: $M_{SD\ 40}$ = 0.20 m; $M_{SD\ 120}$ = 0.29 m) and resulted in larger differences to the manual baseline condition. The two different samples (n = 15 and n = 10) considered did not influence this pattern, as mean variation between intervals did not change when considering the same sample as the one considered for high speed conditions (changes in means when considering n = 10 were 0.01 m or smaller in all conditions).

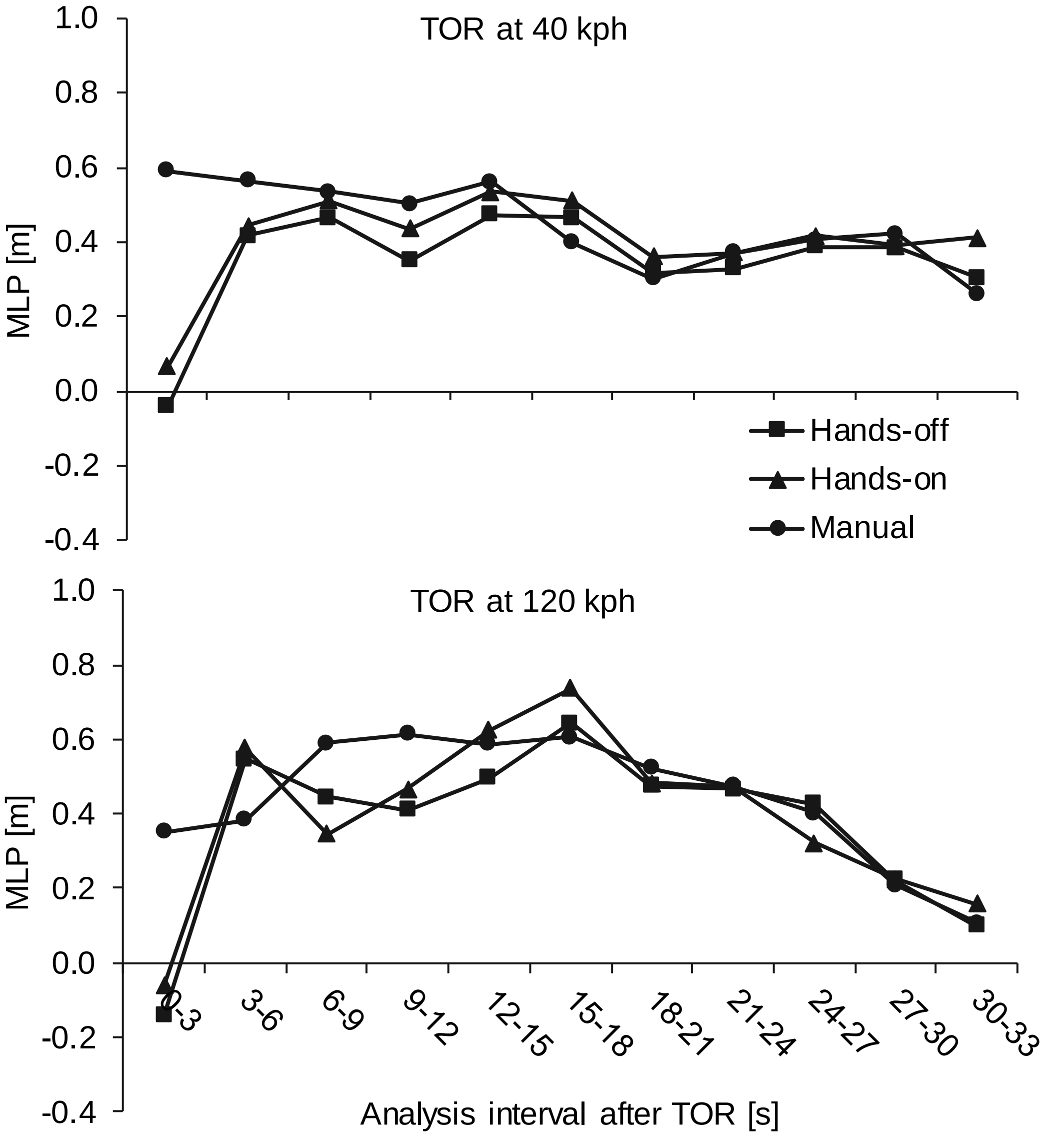

Figure 5-14: MLP aggregated for 3-s-intervals beginning at TOR for low speed (upper graph; sample means for *n* = 15) and high speed (lower graph; sample means for *n* = 10) with sample means for each of the three driving conditions.

5.6 Study 4: Planned Takeover

In contrast to the prior studies, three groups of participants were tested for control transitions before construction zones[13]. Two of the groups used automation while lead vehicles were present (referred to as the hands-*on* and hands-*off* group or as *PAD* groups). One additional group (*man*ual group) drove the same track with lead vehicles, but without automation (see Figure 5-4). Two types of scenario were

[13] Data reported in this chapter was collected in a project funded by the Profile Area ICT of RWTH Aachen University (RWTH-PH-ICT).

compared within each participant group that differed in their criticality, as defined by the TLC to the lane end when the system limit became visible (*early* versus *late* visibility, i.e., planned versus immediate takeover). The manual group further tested hands-off PAD in a comparison scenario without a lead vehicle. Manual driving performance data for post-automation comparisons was provided for each participant by a self-initiated late, but subjectively uncritical manual lane change without a lead vehicle.

A sample of 37 naïve participants took part in the study (30 male; $M = 25$ years, $SD = 4$). For the hands-off group, 13 participants were tested. For both other groups, 12 participants were tested each. As this study featured a between-subject design, the sample's characteristics within each group needed to be taken into account more closely. The three experimental groups did not differ regarding age, $F(2, 36) = 3.10$, $p = .058$, with mean differences between groups of less than one year to up to four years. The groups did not differ regarding reported annual mileage, $F(2, 36) = 0.83$, $p = .444$, although mileage varied greatly over participants ($M = 11640$ km, $SD = 24217$). Driving experience in the overall sample ranged from three to 18 years ($M = 8$ years, $SD = 4$). The hands-on group had least regular experience with ADAS, with five out of 12 participants reporting to never have used either CC, ACC, TJA or LKA more than once (for both hands-off and manual group: $n = 2$). However, mean summed ADAS experience was similar between the three groups ($M_{man} = 6.58$, $SD = 1.83$, $M_{on} = 5.42$, $SD = 2.31$, $M_{off} = 6.92$, $SD = 2.96$). Most participants consistently followed the instructed hand posture during automated driving and all followed it shortly before a TOR-situation became visible or before an unexpected TOR was issued.

Each group reported an equally correct pre-drive mental model, with minimum values of 90 % correct judgements ($MD = 100$ % for each group). The inability of the system to perform lane changes on its own was indicated correctly by all participants, whereas four participants did incorrectly state that all kinds of lane markings would be correctly detected by the system (three in the manual group). As the construction zone, the takeover scenario in this study, featured both system limits, an adequate understanding of the situation can be assumed nonetheless for all participants. Furthermore, reported certainty of when monitoring might be especially important was similarly high in each group (all $M > 5$). An analysis of trust ratings (mean of all items) revealed no differences in initial trust between the groups, $F(2, 30) = 0.18$, $p = .839$. Trust increased significantly after use of the systems, $F(1, 30) = 9.68$, $p = .004$, $\eta_p^2 = 0.24$ ($M_{pre} = 4.79$, $SD = 0.91$; $M_{post} = 5.14$, $SD = 0.99$). The interaction between group and point of measurement (before / after use) was not significant, $F(2, 30) = 0.45$, $p = .645$. Differences in the initial level of trust or the mental model can thus not account for possible group differences found in the following analysis.

For all PAD situations and manual driving situations with lead vehicle, speed at relevant points, that is, at TOR point or point of the lead vehicle changing lanes, was on average 100 kph with slightly larger speed variation in manual drives (SD_{man} = 3 kph; SD_{PAD} = 0.4 kph). Mean distance to the lead vehicle at relevant points was comparable between groups, but manual driving resulted in higher variance as expected (mean distances [SD] early / late visibility; man: 55.10 m [15.91] / 57.93 m [16.92]; PAD: 57.71 m [1.44] / 57.46 m [0.55]). For the manual comparison situation without lead vehicle, higher variations around the target speed were found for all three groups of participants (SD_{man} = 4.7 kph, SD_{on} = 3.3 kph, SD_{off} = 4.3 kph). Mean duration of automated driving was 5:00 min for the early-visibility condition and 4:13 min for the late-visibility condition. For the manual group, the hands-off automated drive without lead vehicle in the end of the study was shorter by design with on average 1:13 min. However, all durations were well below those relevant for vigilance effects and within the range of durations used for testing short-termed takeover situations in prior studies.

Three participants, two in the hands-on group, deactivated the system before the lead vehicle began to change lanes or permanently caused an offset in steering angles by working against the function. These participants, as well as two participants in the manual group who intervened at a very early point in time in their already short hands-off PAD drive, thereby heavily limiting the use of the system, were excluded from the analysis of the corresponding manoeuvre. One participant of the manual driving group did not follow instructions in the baseline drive, changing lanes not at the last uncritical point, but more than 7 s before reaching the construction zone. This participant was excluded from the analysis of post-automation driving performance.

The late-visibility situation left, by definition, less time for a lane change. Ten cases of lane departures respectively collisions with traffic cones before changing lanes were observed, compared to one lane departure in the early-visibility condition with lead vehicle. In most cases, lane departures were minor and the COG did not leave the lane. In the manual driving condition, all lane departures were minor. No differences in the frequencies of lane departure could be observed, with four cases in the manual and hands-off group each, including the only case in the early-visibility condition (hands-off), and three cases in the hands-on group. Lane departures before changing lanes, indicating delayed initiations of the necessary lane change without braking in time, were excluded from the analysis of driving quality, whereas cases with lane departures after a timely lane change were included in quality considerations as characterising the quality of the re-orientation on the new lane. Overshoot of the target lane was observed only in late-visibility conditions for six participants of both automation groups each, as well as for one participant of the manual driving group. Only in two of these cases ($n = 1$ hands-off) was the lane change initially successful, raising the overall count of unsuccessful lane change manoeuvres to 13 cases. These numbers are a first indication that lane changes were generally conducted in a more controlled manner for manual driving and when more time was provided to anticipate the lane end.

Lane changes were conducted farther from the lane end with earlier visibility (Figure 5-15). The difference between unassisted and automated driving becomes apparent by considering the range of distances to the lane end at lane change, with 75 percent of lane changes in the automation groups being conducted considerably closer to the lane end than 25 percent of the latest lane changes in the manual group (manual: $\%ile_{25}$ = 72.91 m, $\%ile_{75}$ = 109.59 m; hands-on: $\%ile_{25}$ = 38.89 m, $\%ile_{75}$ = 57.65 m; hands-off: $\%ile_{25}$ = 27.44 m, $\%ile_{75}$ = 45.35 m). For takeover situations without lead vehicle, the range for lane changes was larger, showing a mixture of critical, post-automation-like, and uncritical, manual-like lane changes (manual group driving hands-off: $\%ile_{25}$ = 34.64 m, $\%ile_{75}$ = 90.12 m).

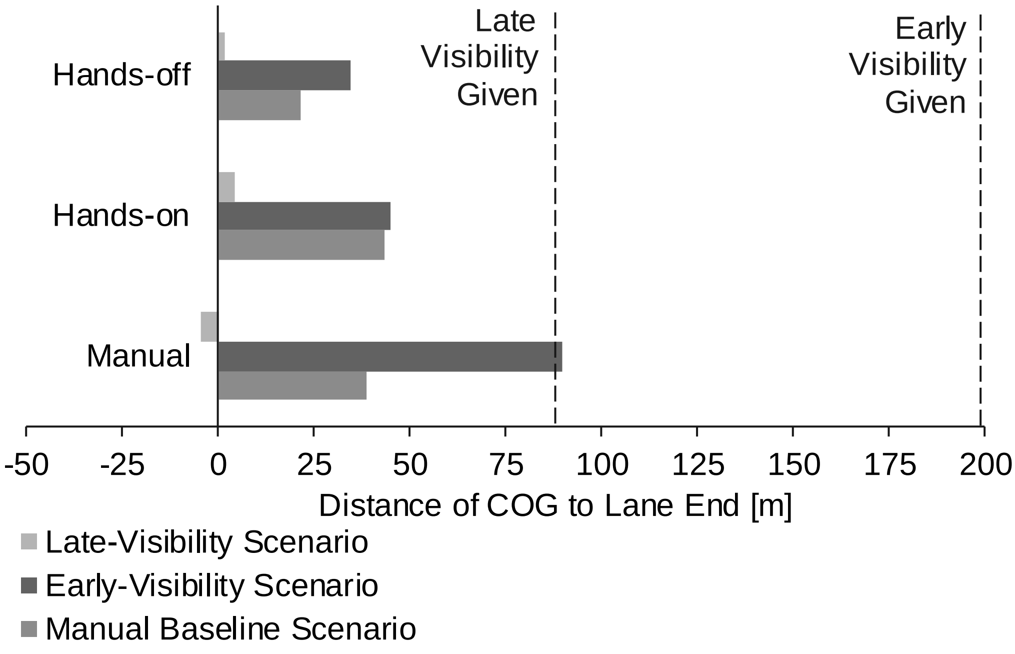

Figure 5-15: Mean distances to the lane end (of the right lane), measured when the COG crosses the marking between right and left lane in the three driving scenarios (early, late and baseline). Late visibility corresponded roughly to the point of TOR in the automated conditions. Construction lane markings were s-shaped lines of 100 m (corresponding to the distance between -50 m to 50 m), intersecting with the right lane's centre at zero metres.

Four drivers initiated a takeover before receiving a TOR in each automation condition. In the other 15 valid cases, participants waited for the system to issue a TOR and to deactivate before they regained control. Of all drivers acting in advance, only one experienced the relevant situation as the first takeover scenario, speaking for an effect of experience on the self-initiation of control transitions. The safety-critical takeover experience (i.e., late visibility) of the first situation was mentioned twice as the reason for a self-initiated intervention in the second situation. In the situation without a lead vehicle, a TOR was issued in four out of ten cases, indicating slightly more self-initiated driver actions than in situations with lead vehicle. In cases with lead vehicle, drivers intervened mostly shortly before a TOR would have been issued anyway, whereas much earlier interventions were noted without lead vehicle. The lane change was nonetheless performed with a significantly larger distance to the lane end by drivers who did not receive a TOR in the automation groups, $t(21) = 2.81$, $p = .010$ ($M_{anticipated} = 54.75$ m, $SD = 18.96$; $M_{TOR} = 30.93$ m, $SD = 19.55$). Additionally, these drivers took more time to change lanes after initiating the lane change manoeuvre, $t(21) = 2.60$, $p = .017$ ($M_{anticipated} = 3.10$ s, $SD = 0.94$; $M_{TOR} = 2.14$ s, $SD = 0.78$). Surprisingly, four of the participants who did receive a TOR instead of performing a driver-initiated takeover reported afterwards that they did not believe that this situation would result in a takeover initially. This is contrary to the correctly identified system limits in the mental model questionnaire.

Analysis of steering wheel angles applied by the automated system on a straight road section revealed maximum steering wheel angles of 0.1°. Thus, active steering input of the driver was detected at values larger than the maximum system input. The threshold was kept at a deliberately low level due to the rather high deactivation threshold (see Chapter 5.1.3) in an attempt to include drivers who initially did not steer enough to deactivate the system. Pedal input larger than 1.14 % of maximum pedal travel, excluding signal noise, was considered as active driver input for this purpose. The first threshold reached, either by steering, braking or accelerating, sets the time of first driver input. It was verified whether the unintendedly high system-deactivation thresholds influenced takeover initiation. Timings based on both approaches (FDI and LCI, see Chapter 5.2.3) were compared and found to be congruent. This indicates that the first driver action after automation is always executed either until lane change or a significant reduction of speed. All further analyses thus focus on the LCI definition, thereby enabling a direct comparison to manual driving data.

Steering was again found to be the predominant initial reaction to encountering critical situations. Only in seven cases was a braking reaction (at least 10 % of maximum pedal travel) observed before detecting a steering manoeuvre, with four braking reactions (only one of these after PAD) in the late-visibility condition. Braking was never the initial reaction in early-visibility conditions for manual driving, but once after automated driving and twice in the takeover scenario without lead vehicle. A follow-up inspection of data revealed, however, that speed was reduced more often and more heavily after initiation of the lane change manoeuvre for the time-critical scenario, but predominant braking (> 10 % of maximum pedal travel) occurred similarly often in all groups.

Comparison of visibility (*early* versus *late*) and group (*man* versus *on* versus *off*) revealed a significant main effect of visibility, $F(1, 31) = 14.43$, $p = .001$, $\eta_p^2 = 0.32$, with faster interventions in the late-visibility scenario ($M_{late} = 1.82$ s, $SD = 0.53$; $M_{early} = 2.66$ s, $SD = 1.44$; see Figure 5-16)[14]. The main effect for group, $F(2, 31) = 11.94$, $p < .001$, $\eta_p^2 = 0.44$, indicated later interventions after automated driving ($M_{man} = 1.61$ s, $SD = 0.58$; $M_{on} = 2.26$ s, $SD = 0.95$; $M_{off} = 2.85$ s, $SD = 0.90$). The interaction between scenario and group was significant as well, $F(2, 31) = 6.19$, $p = .005$, $\eta_p^2 = 0.29$, motivating the separate comparison of automation conditions for each visibility scenarios.

For the late-visibility condition, no significant difference between automation groups was found, $F(2, 32) = 1.16$, $p = .328$ (mean differences; $\Delta_{\text{off-on}} = 156$ ms; $\Delta_{\text{on-man}} = 163$ ms; $\Delta_{\text{off-man}} = 319$ ms). The equally fast LCI in the manual driving group is noteworthy, as this group did not receive a visual-acoustic notification at the lead vehicle's lane change, as did the automation groups by receiving the TOR.

For the early-visibility condition, a significant main effect was found, $F(2, 32) = 10.94$, $p < .001$, $\eta_p^2 = 0.41$, with the largest mean difference between the manual and the hands-off group ($\Delta_{man\text{-}off} = -2.16$; $p < .001$). The mean differences between the manual and the hands-on group were significant by tendency only ($\Delta_{\text{man-on}} = -1.08$; $p = .085$) as was the difference between the feedback groups ($\Delta_{\text{on-off}} = -1.07$; $p = .089$). Drivers anticipating the takeover ($n = 8$) in the early-visibility condition acted on average 1.76 s earlier than those waiting to receive the TOR ($n = 15$).

As can be seen in Figure 5-16, the type of scenario did not influence manual drives, as drivers intervened equally fast whenever the situation became apparent, $t(11) = -0.62$, $p = .549$ ($M_{late} = 1.66$ s, $SD = 0.59$; $M_{early} = 1.56$ s, $SD = 0.57$). This is not the case for the hands-off group, with drivers waiting longer to intervene when provided with more time, $t(11) = 3.92$, $p = .002$ ($M_{late} = 2.00$ s, $SD = 0.40$; $M_{early} = 3.71$ s, $SD = 1.40$). The hands-on group, although showing descriptively the

---

[14] Contrary to the other driving simulator studies, intervention times are reported in seconds for this study due to the larger timeframe available for intervention in early-visibility scenarios.

same pattern as the hands-off group, did not react significantly later when provided with more time, $t(9) = 1.81$, $p = .10$ ($M_{late} = 1.82$ s, $SD = 0.60$; $M_{early} = 2.70$ s, $SD = 1.29$). As the study was conducted using a between-subject design, intervention times in the manual baseline drive were additionally compared between groups to test for differences in the self-initiated late, but still uncritical lane change decisions. No difference between groups was found, $F(2, 33) = 1.89$, $p = .167$.

Situational assessment ratings were considered valid only if the situation was classified subjectively and objectively as either with or without TOR, excluding two participants from analysis (remaining $n = 35$). The subjective criticality of the situation was influenced by the time the upcoming lane change was visible before TOR, resulting in significantly higher criticality ratings for late-visibility conditions ($z = 4.87$, $p < .001$; $M_{\text{early}} = 7.00$, $SD = 3.82$, $M_{\text{late}} = 12.83$, $SD = 2.51$). This pattern was the same in all groups (man: $M_{\text{early}} = 7.25$, $SD = 3.91$, $M_{\text{late}} = 12.00$, $SD = 2.98$; on: $M_{\text{early}} = 6.27$, $SD = 4.13$, $M_{\text{late}} = 13.91$, $SD = 2.17$; off: $M_{\text{early}} = 7.42$, $SD = 3.68$, $M_{\text{late}} = 12.67$, $SD = 2.10$). The differences between groups were rather small in both conditions.

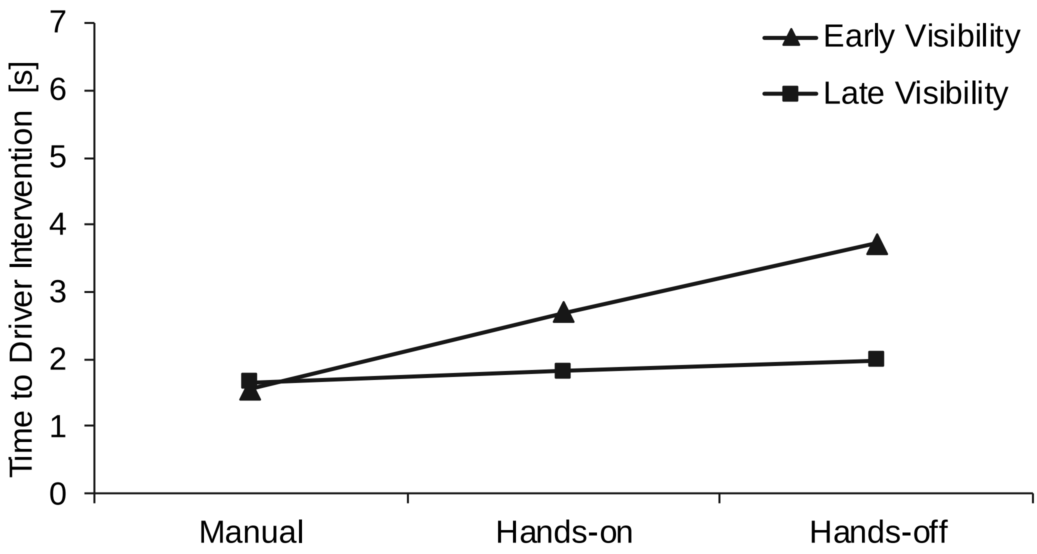

Figure 5-16: Intervention times (LCI or braking, referenced to the beginning lane change of the lead vehicle) for the two types of takeover scenario (early and late visibility) and the three automation groups (manual, hands-on, hands-off).

Of higher relevance is the question whether the subjectively perceived criticality differed between participants conducting anticipative actions and those waiting for the TOR to be issued. For this analysis, criticality ratings for the *early-visibility condition* were considered separately, including the hands-off drive without a lead vehicle (manual group) as well as participants not considered for the analysis of driving data. Overall, 17 participants did not receive a TOR due to an early, self-initiated takeover, whereas 18 participants received a TOR. Subjective criticality differed significantly between groups with and without TOR, with the group receiving a TOR rating the situation more critical although encountering the objectively same situation (z = 1.98, p = .049; M_{TOR} = 7.78, SD = 3.52; $M_{anticipated}$ = 5.18, SD = 3.19). The difference was even larger when ratings on the situation without lead vehicle were excluded (M_{TOR} = 8.14, SD = 3.66, n = 14; $M_{anticipated}$ = 4.89, SD = 3.44, n = 9). However, comparing ratings on the subjective takeover quality, no difference was found between the two groups (z = -1.56; p = .119; M_{TOR} = 9.56, SD = 3.20; $M_{anticipated}$ = 11.24, SD = 3.11). The same group comparison was conducted for the *late-visibility scenario*, testing for general differences in criticality judgements. Ratings of the two anticipation groups did not differ for this scenario (z = -1.32, p = .186; M_{TOR} = 12.17, SD = 3.03; $M_{anticipated}$ = 13.53, SD = 1.62). Results do thus not support a general underlying difference in criticality perceptions between groups.

Gaze analyses focused on changes in monitoring during the early-visibility condition and compared participants receiving a TOR to those initiating a control transition before TOR on their own (TOR versus anticipation). Data of all participants of both automation groups were included. Contrary to analyses on short-termed takeover, three different PRC intervals were analysed. The first interval (Interval I) was defined similar to the analyses of prior studies and included the 60 s of automated driving before lane change initiation of the lead vehicle. The other two intervals included a maximum of 4 s each, one aggregating the last seconds before lane change initiation of the lead vehicle (Interval II), likely most relevant for the detection of the lead vehicle's lane change early on. The third interval began with the lane change initiation of the lead vehicle and ended with the own lane change or at TOR (Interval III). As shown in Figure 5-17, the group initiating a control transfer before TOR exhibited higher PRC values in all three intervals, with the largest group difference in the interval directly preceding the lead vehicle lane change (Δ = 19.5 %; Interval II). The difference in PRC between groups in this interval was significant (U = 113, p = .036; $n_{anticipated}$ = 10, n_{TOR} = 15). Hand position was not balanced over the two groups, as detailed above. However, the tendency towards higher PRC for participants intervening before TOR was consistent in both automation conditions (difference in mean PRC for participants with and without TOR; hands-on: $\Delta_{\text{Interval I}}$ = 9.72, $\Delta_{\text{Interval II}}$ = 13.75, $\Delta_{\text{Interval III}}$ = 2.68; hands-off: $\Delta_{\text{Interval I}}$ = 6.12, $\Delta_{\text{Interval II}}$ = 25.23, $\Delta_{\text{Interval III}}$ = 9.24).

For comparison of driving performance, participants with lane exceedances before changing into the target lane were excluded from analysis (early: n = 1; late: n = 10). Additionally, one participant of the manual group was excluded from the analysis of MLP for not providing a valid baseline drive. $Steer_{max}$ was compared between automated conditions for the first three seconds after LCI. SDST was compared after lane change between samples, whereas MLP, reflecting the individual driving strategy, was compared within-drivers to the baseline condition of a self-initiated, uncritical lane change.

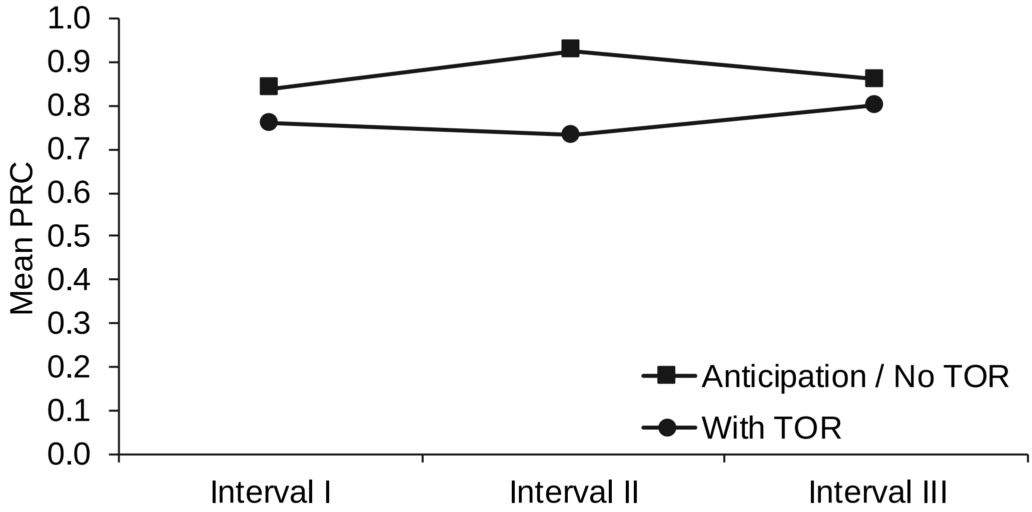

Figure 5-17: Mean PRC for participants intervening *before* TOR (Anticipation / No TOR; $n = 10$) and *after* TOR (With TOR; $n = 15$). Interval I: 60 s until limit visibility; Interval II: 4 s until visibility; Interval III: from visibility until lane change or TOR (maximum duration: 4 s).

$Steer_{max}$ was smallest in all groups for the self-initiated lane change in the baseline drive ($MD_{man} = 1.84°$; $MD_{on} = 1.98°$; $MD_{off} = 2.33°$), underlining the necessity for a separate indicator of planned driver interventions at high speed compared to prior takeover scenarios at lower speed in Studies 1-3. Equivalently, the lower urgency of the early-visibility situation was reflected in smaller differences between conditions and smaller overall median values ($MD_{man} = 3.76°$; $MD_{on} = 11.82°$; $MD_{off} = 8.16°$) as in the late-visibility situation ($MD_{man} = 7.52°$; $MD_{on} = 20.55°$; $MD_{off} = 25.93°$). The differences between conditions as well as the magnitude of overall $steer_{max}$ are thus influenced by the visibility of the situation. A comparison to the takeover without lead vehicle showed that unrestricted visibility of a takeover situation did not further reduce $steer_{max}$ ($MD_{man} = 10.17°$). This value might however be influenced by the criteria for manual disengagement of the system. The difference between drivers initiating a takeover before TOR and those waiting until TOR, thus initiating a steering manoeuvre after system off, was however minimal ($MD_{TOR} = 9.27°$; $MD_{anticipated} = 10.44°$), indicating no effect of self-initiated deactivation on $steer_{max}$. Similar effects were found for the early-visibility conditions with lead vehicle ($MD_{TOR} = 9.67°$, $MD_{anticipated} = 13.88°$).

Mean speed after lane change was analysed as a control variable. At least for late-visibility conditions, larger differences in speed were expected due to significant braking interventions. Low mean speed in the first analysis interval resulted in the exclusion of two participants from the manual driving group in the late-visibility condition with mean speeds of 20 kph and 33 kph. These additional exclusions, in combination with exclusions due to lane departures and invalid data sets, reduced the dataset to 24 participants for the late-visibility situation (six manual, nine for each automation condition). For the early-visibility condition, 33 participants were considered ($n_{man} = 11$, $n_{on} = 10$, $n_{off} = 12$). For takeover without lead vehicle, ten participants of the manual group were considered. Mean speed for the reduced dataset was above 90 kph in all intervals and conditions, with maximum mean values of 108 kph occurring in conditions without lead vehicle (see Figure 12-4 provided in Appendix).

SDST was highest in the first interval after lane change and became markedly smaller over the next two intervals in all automation and visibility conditions. The pattern observed for $steer_{max}$ with higher values in situations of restricted visibility was evident for SDST as well (see Figure 5-18). Mean SDST in the baseline drive (< 0.62° at all times) was initially exceeded in all automation conditions. For the manual group, steering input after early visibility ($M_{0\text{-}3\,s} = 0.72°$) was not discernible from self-initiated, uncritical lane changes. Differences between automation conditions and in comparison to the baseline dissolved after the first analysis interval, suggesting once again a very short effect of takeover and lane change on driver performance for those drivers who managed the initial lane change successfully. The exception was the hands-off group for time-critical takeover, where differences to the within-group baseline ($\Delta_{\text{off-baseline}} = 1.62°$) were still observable 3 s later (difference to the manual group; $\Delta_{\text{off-man}} = 1.52°$; $\Delta_{\text{on-man}} = 0.41°$). The effect of the takeover was attenuated in the early-visibility condition (see Figure 5-18).

Self-initiation of the takeover did not seem to influence steering behaviour after lane change, as a descriptive comparison of participants receiving a TOR to those initiating a takeover before TOR revealed ($M_{TOR} = 3.16°$, $SD = 1.96$; $M_{anticipated} = 2.80°$, $SD = 1.12$). However, as drivers receiving a TOR in the early-visibility condition did have an equal amount of time after intervention as in the late-visibility condition, their more controlled steering input in comparison to late visibility likely indicates preparation effects. SDST after TOR without a lead vehicle (upper part of Figure 5-18) differed initially from unassisted driving ($M_{PAD} = 1.77°$, $SD = 1.34$, $M_{baseline} = 0.56°$, $SD = 0.35$), but comparatively little only, when considering the hands-off group with lead vehicle in comparison (i.e., with limited time to intervene; $M_{late} = 10.85°$, $SD = 7.68$; $M_{early} = 2.46°$, $SD = 1.12$). Although the effect of automation was thus still observable after maximum preparation for takeover, its effect was heavily attenuated.

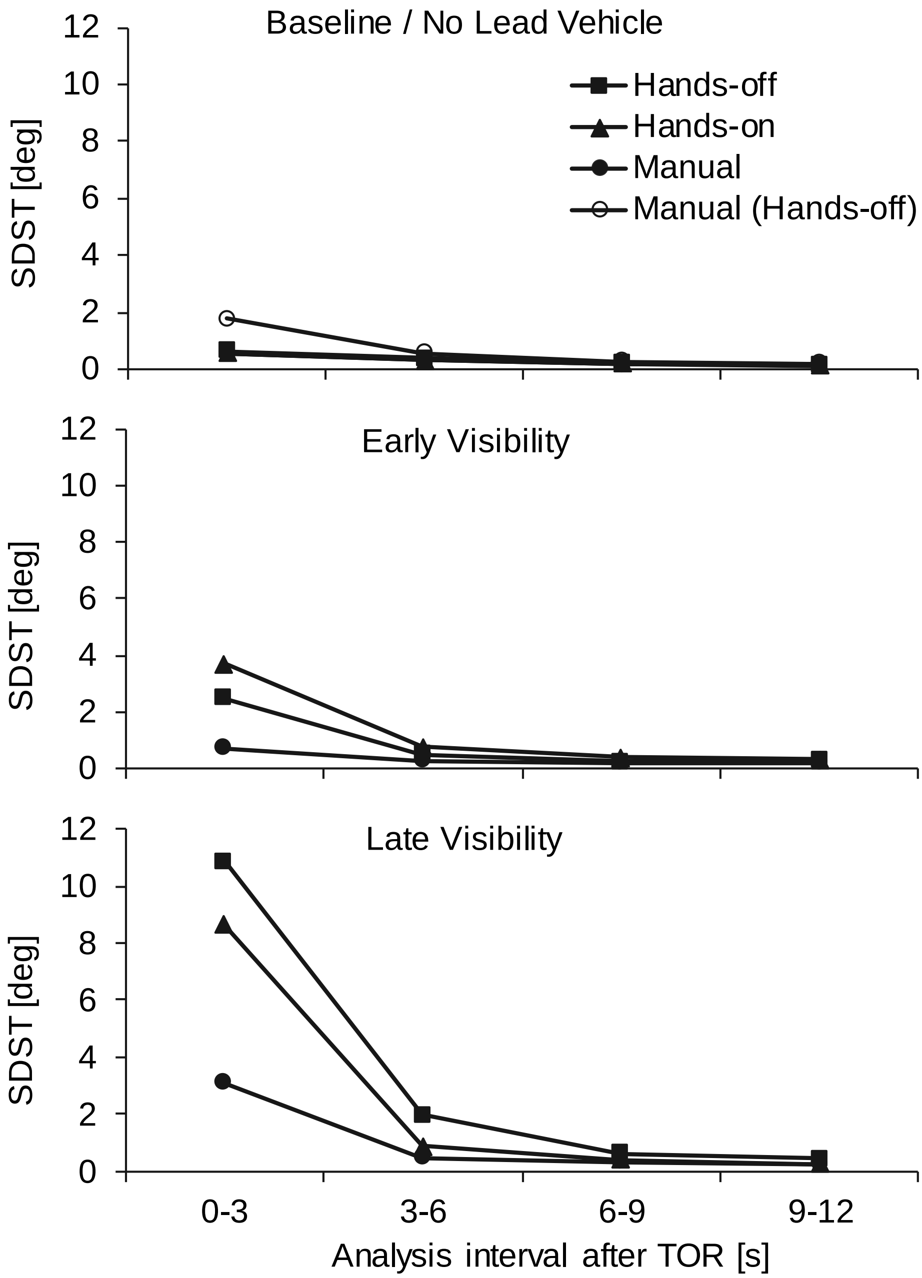

Figure 5-18: SDST aggregated over 3-s-intervals, starting at COG entering the new lane. Only valid data sets were included in sample means for each group and condition. Upper graph: Data for the baseline drive of each group and the additional automated drive without lead vehicle of the manual group. Middle graph: Early visibility condition (TLC > 6 s). Lower graph: Late visibility condition (TLC > 2 s).

As for SDST, the largest change in MLP (see Figure 12-5 in Appendix) was observed between the first two intervals, reflecting the re-positioning on the new lane.

Afterwards, MLP can be described as relatively stable for all conditions. MLP for the first interval in the late-visibility condition indicated a larger lateral distance to the construction zone than after less critical lane changes. This is in line with the higher average initial overshoot (cf. Happee et al., 2017) for PAD after late visibility (initial overshoot in lane offset; positive values indicate a larger lateral distance to the construction zone markings; early visibility: $M_{man} = 0.30$ m, $SD = 0.17$, $M_{on} = -0.19$ m, $SD = .60$, $M_{off} = 0.05$ m, $SD = 0.28$; late visibility: $M_{man} = 0.30$ m $SD = 0.53$, $M_{on} = 0.27$ m, $SD = 0.66$, $M_{off} = 0.50$ m, $SD = 0.79$). No general effect of automation conditions was visible in MLP. However, the effect of visibility was observed in each group with descriptively larger differences in the late-visibility condition for at least 6 s.

5.7 Overview on Effects of the Context of Use

Four driving simulator studies were conducted with the primary aim of investigating the effects of reduced haptic feedback (i.e., hands-off driving) on monitoring behaviour, takeover initiation and on post-automation driving performance. Each study investigated a specific variation of context, addressing factors that have been connected to either deteriorated takeover performance or monitoring behaviour in prior research. In Study 1, short and long durations of use were compared in two types of takeover scenarios. In Study 2, a visual-manual secondary task was offered to drivers, targeting drivers' intent to monitor under the variation of haptic feedback. In Study 3, use of automation was investigated in two different traffic situations, namely in dense, low speed traffic and light traffic at high-speed. Finally, Study 4 considered differences in the anticipation of system limits (i.e., hazard perception) as another way to quantify changes in the engagement with the driving task. The analyses presented in the preceding chapters focused on behavioural measures and the effect of context variations.

For interventions after short-termed takeover, a consistent disadvantage in the timing of takeover initiation was found after hands-off monitoring. Overall, intervention times in all conditions were fast with mean intervention times after TOR below one second. A detailed analysis using timing distributions was conducted using the aggregated data set with a larger sample (see Chapter 6.3). Comparing response times after unassisted and automated driving, differences in lane change initiation could not be shown for time-critical driving situations (TLC < 3 s) in Study 4, but only for conditions in which more time was provided between the visibility of the system limit and the system limit itself (TLC > 6 s). An equal number of self-initiated control transitions before TOR was found, regardless of haptic feedback provided during automation use. However, lane changes were initiated later after hands-off monitoring, aligning with the results from the other simulator studies (Study 1-3). All other variations of context had no effect on the difference in timing of actions between feedback conditions or on takeover timing in general. Contrary to the stable effect of later interventions, subjective ratings did not continuously indicate a higher criticality of takeover situations after hands-off monitoring, with significant effects found only for Study 1.

Monitoring behaviour did align with the non-significant intervention time differences between context variations in that only a negligible decline in gaze attribution to the road centre was found over longer durations of PAD (Study 1). In accordance with the high secondary task engagement observed in all automation conditions, an influence on gaze behaviour with an overall significantly lower PRC during secondary task offers was noted in Study 2. The change in attention attributed to the road centre during secondary task engagement was however not influenced by levels of haptic feedback. For different traffic situations (Study 3), an effect on monitoring behaviour was found. Contrary to the hypothesis that higher situational complexity, introduced by higher traffic density, increases monitoring efforts, more gazes were attributed to the PRC in light traffic at higher speed. Finally, drivers initiating a lane change before a system-triggered TOR (Study 4) showed an overall higher attribution of gazes towards the forward road scene as those waiting to receive a TOR for an immediately necessary driver intervention. This effect was again observed in both feedback conditions.

As intended by design, the selected takeover scenarios resulted in steering interventions as first driver input for all short-termed takeover studies. For Study 4, a small number of braking manoeuvres was observed in the late-visibility scenario. Higher initial steering movements ($steer_{max}$) were found consistently after hands-off monitoring. The higher, albeit overall small, number of lane departures in the hands-off condition aligns with the other metrics analysed.

SDLP analysis was discharged as an insightful metric for very short analysis intervals (i.e., interval lengths of 3 s) used in the descriptive analysis of post-automation driving performance. The analysis of SDST, reflecting fast changes in steering behaviour, and MLP, reflecting the general strategy of a driver, revealed fast dissolving effects of automation on driving performance. However, in congruence with mean takeover interventions after hands-off monitoring occurring between on average 205 ms (Study 2) and 302 ms (Study 3) later, effects of automation on driving performance were not only consistently larger, but often apparent for at least one analysis interval longer compared to hands-on monitoring. This effect was observed regardless of context variations or roadway design (curved or straight sections). A comprehensive discussion of the effects found is provided under consideration of all analyses conducted in Chapter 8.

6 Stability of Effects Over Methods and User Characteristics

Changes in interaction under variation of feedback in different contexts of use have been shown in the previous chapter. Open questions on the behavioural level concern the influence of driver experience, of repeated exposure to TOR (practice), as well as the abstraction level of interaction (see Figure 6-1). Furthermore, results by Kerschbaum et al. (2014), finding that the effects of a decoupled steering wheel were "highly depending on the individual participant" (p. 1690), warrant a closer look at the range of performance differences under variation of feedback. To conduct these analyses, an aggregated, larger sample is needed.

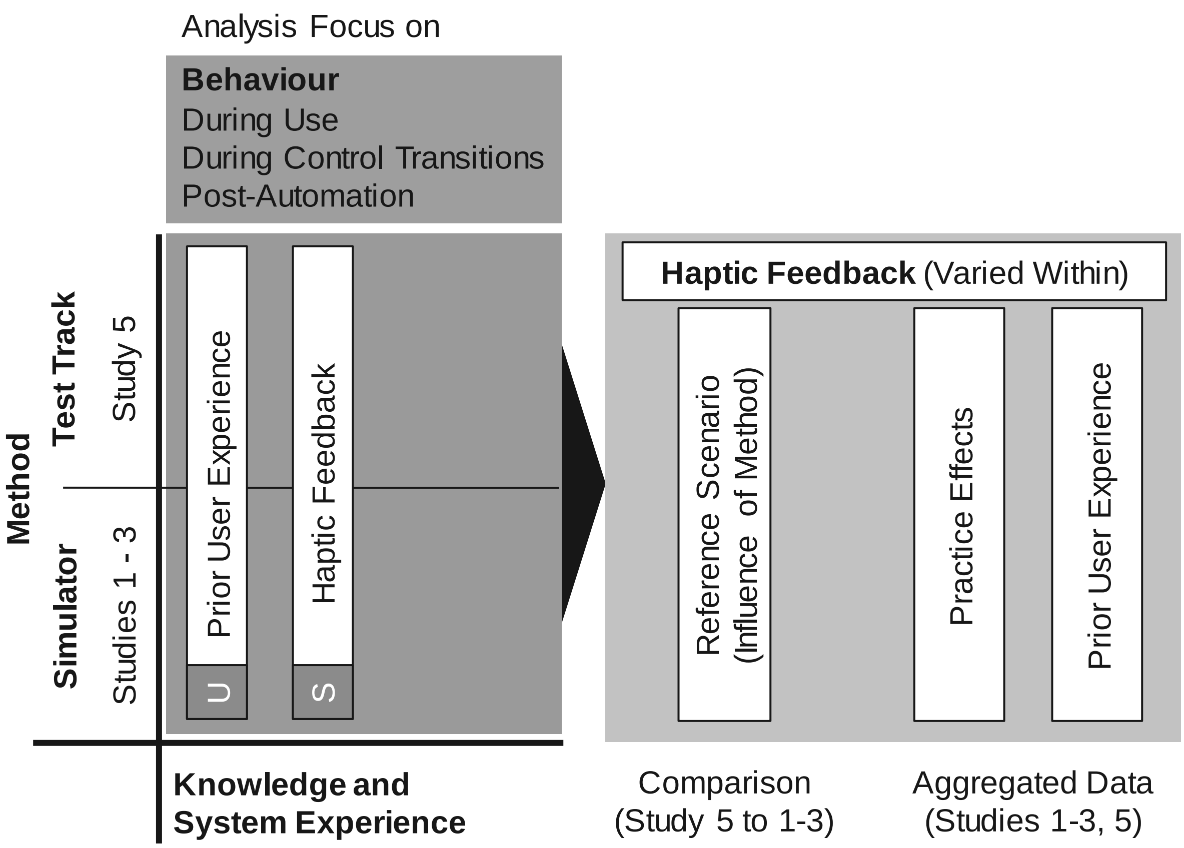

Figure 6-1: Overview on user (U) and PAD system (S) factors being investigated on the behavioural level for different phases of interaction (i.e., during use, during a control transition and after automation use; left hand side), and their consideration in the analysis (right hand side).

6.1 Study 5: System Use on Test Track

Controlled experimental studies are an abstraction of reality underlying methodological constraints through method choice. Driving simulator studies have proven to be a valid method for investigating the interaction of users with assistance systems (e.g., Kaptein et al., 1996; Abendroth et al., 2012). However, the missing lateral and longitudinal acceleration of a fixed-based driving simulator as well as the reduced risk associated with late or insufficient takeover actions might influence the size and occurrence of automation effects (cf. Stanton & Young, 1998). Especially for the analysis of driving quality, the insufficient feedback of driving dynamics caused by large or abrupt steering interventions after takeover might be a relevant factor in the definition of the end of corrections after takeover. To validate the results of the simulator studies, a test track study[15] was conducted.

6.1.1 Setup of Test Track Study

The scenario of the test track study was designed to match the reference scenario tested in the simulator studies (see Chapter 5.1.2). The general procedure of the study, including system instructions and questionnaires, was the same as in the driving simulator studies. The test vehicle used was a Volkswagen CC. The implemented system was designed under consideration of the system used in the driving simulator, including HMI and activation logic. Vehicle data (sample rate: 75 Hz), gaze data and hand posture (by camera) were captured equivalently to the simulator studies.

Two surrounding vehicles restricted available space to the rightmost two lanes of a three-lane oval on which testing took place (see Figure 6-2). Similar to the driving simulator studies, a TOR was issued shortly before entering a curve with 113.5 m radius while following another vehicle with 40 kph. A third vehicle was following and attempting to overtake the lead vehicles in the test situation and at other occasions during the drive to decouple the overtaking manoeuvre from the test situation. The overtaking manoeuvre was included to narrow the available space for corrective actions to the rightmost lane after the system had transferred control back to the driver. To prevent a direct re-activation of the function after TOR, the lead vehicle accelerated shortly after the TOR, dissolving the traffic jam situation and leaving the functional range of the automated system.

The HMI used in the test vehicle featured the same relevant aspects as the HMI implemented in the driving simulator, specifically colour-coded lane markings in green and orange and written descriptions of the current status (“System ready”, in German: “System bereit”; “Take over! System off”, in German: “Übernehmen! System aus”) as well as system sounds. The offer to activate the system was triggered at

[15] Data of this study and a comparison to the simulator studies can be found in adapted form in Josten et al. (2016).

40 kph, if a lead vehicle was detected. The system, once activated, followed the lead vehicle with a THW of 1.5 s. Driving speed of the lead vehicle was between 15 kph and 50 kph. The correct target speed (measured CAN-speed of own vehicle: 40 kph) was controlled by ACC activation of the lead vehicle when approaching the TOR scenario. Due to the given radius of the curve, the TOR was issued not on the last straight road metre but a few metres inside the curve to achieve TLC values comparable to the simulator studies. This resulted in steering wheel angles of less than 5° being set by the system at the time of the TOR (see Chapter 6.1.2). To account for this, active driver input was detected at a difference of 4° compared to the steering angle at TOR instead at an absolute steering wheel angle of 4° as in prior studies. Longitudinal driver input was detected at 5 % of maximum pedal travel.

Participants were able to experience the system including its activation and a driver-initiated deactivation through steering input in an introduction drive. One instructor sat on the passenger seat and another instructor sat behind the driver to handle the measurement equipment. The TOR was triggered at a defined GPS-position, which had been passed successfully by the system multiple times beforehand during the drives. Overall, each automated drive lasted for about five minutes before the TOR occurred. Only one TOR was issued for each feedback condition. The procedure for the manual drive was the same as in the automated drives with the exception that passing the GPS-position did not trigger any specific event.

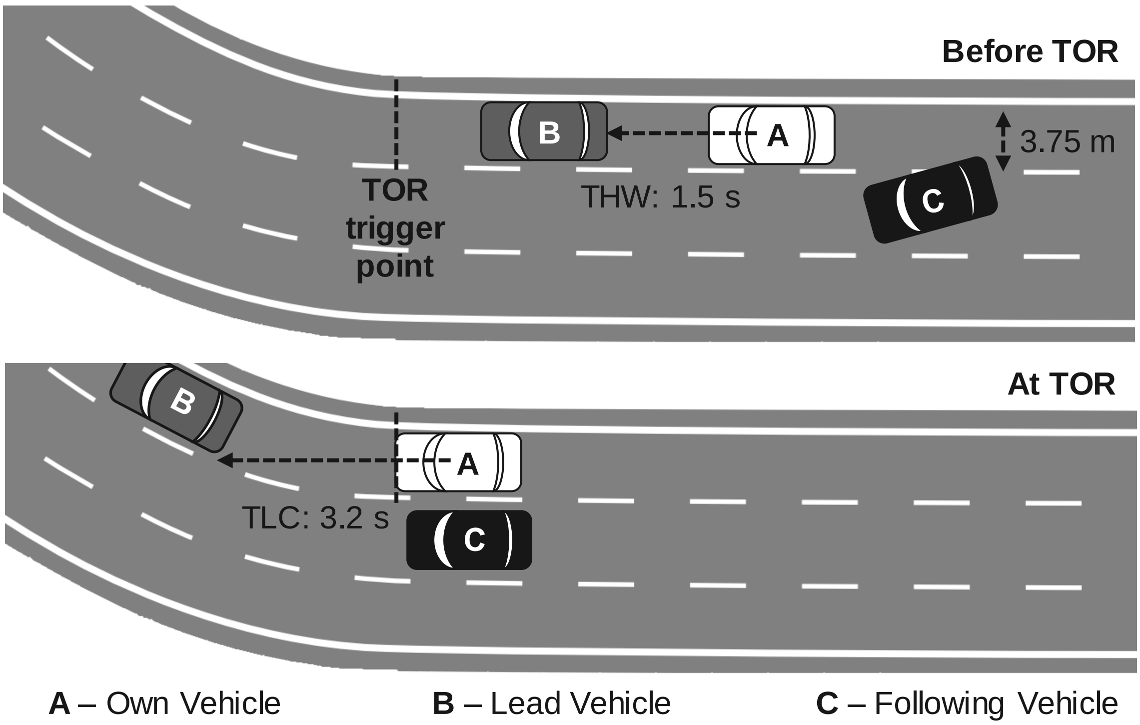

Figure 6-2: Schematic set-up of the takeover scenario on the test track. TLC was calculated from the COG of the own vehicle to the lane marking of the rightmost lane.

6.1.2 Results

Sixteen participants took part in the study. Data of three participants had to be discarded due to technical difficulties with data recording. One participant did not react to the TOR in the first automated drive (hands-off). This takeover was not considered valid for the analysis of performance, as the instructor had to tell the driver to intervene, but taken into consideration for the assessment of automation risks. Data was checked for speed at the trigger point (target value: 40 kph; mean deviation from target speed for manual driving: 2 kph; mean deviation for PAD: 0.4 kph), distance to the lead vehicle (mean in all conditions: 17 m) and steering wheel angle (absolute values; $M_{man} = 3.6°$; $M_{PAD} = 2.5°$).

Of the remaining 12 participants (two female), five participants drove with the hands-off system first. Mean age was 43 years ($SD = 10$) and mean mileage per year was 18083 km ($SD = 5071$). All but one participant had experience with CC systems (over 60 % regular CC experience), but only 25 % had experience with ACC systems. Only one participant travelled on highways less often than once a week. On average, participants were in possession of a driver’s licence for 25 years ($SD = 10$).

For PAD, the type and timing of first input after TOR was compared. In five measurement cases, three drivers reacted by accelerating before taking over lateral guidance. In all other cases, drivers reacted by steering first. For analysis, FDI, exceeding either 5 % of the maximum pedal travel or a difference of 4° steering wheel angle to the value at trigger point, defined intervention time. The first input after monitoring hands-on was carried out on average 110 ms earlier in comparison to hands-off monitoring ($M_{on} = 1.02$ s, $SD = 0.26$ s; $M_{off} = 1.13$ s; $SD = 0.21$ s). This difference is not significant, $t(11) = -1.12$, $p = .285$. Three participants (five measurement cases) took over longitudinal guidance before lateral guidance by accelerating. In congruence with driving simulator data, where steering input was the first active input in all cases, steering input was analysed additionally, but found to result in a similar mean difference between hand postures (100 ms) than FDI. The difference between hand postures in the controlled field was thus of the expected direction, but less pronounced.

To follow-up on the much smaller effect of hand posture in this study, steering intervention times were compared to those in the reference scenario of the driving simulator studies ($n = 50$), thereby establishing maximal congruency between the two test environments. Differences between test track and simulator were larger for the hands-on condition ($\Delta_{on} = 443$ ms) than for the hands-off condition ($\Delta_{off} = 262$ ms) with consistently earlier takeover in the simulator studies. This difference in timing was significant for both hand postures, $t(60) = -5.26$, $p < .001$ (hands-on) and $t(60) = -3.00$, $p = .004$ (hands-off)[16]. Data for the reference scenario in the simulator showed a significant difference between hand postures in contrast to the non-significant effect reported above for the test track, $t(49) = -6.03$, $p < .001$ (M_{on} = 651 ms, SD = 245; M_{off} = 931 ms, SD = 277).

Although criticality of the takeover situation on the test track was rated 'very low' for both automation drives, ratings were slightly higher for the hands-off conditions ($M_{on} = 2.75$, $SD = 2.77$; $M_{off} = 3.58$, $SD = 3.22$). The low subjective criticality together with the five cases of acceleration as first driver input suggest the driving situation was perceived as less demanding than the scenario in the driving simulator, although the TLC at TOR was comparable between studies. The mean subjective criticality ratings for the reference scenario in Study 1-3 (driving simulator), indicated a slightly higher subjective situational criticality ('very low' versus 'low', see Figure 5-6; $M_{on} = 6.01$, $SD = 3.39$; $M_{off} = 6.49$, $SD = 3.81$).

For a comparison of monitoring behaviour between the two test environments, only driving situations comparable to the test track study were included, to be more specific, data from the traffic jam preceding the reference scenario without secondary task offers. This excludes participants from the study on secondary task engagement. Data from overall 46 participants was used. For the test track study, complete data sets were recorded for ten participants, of which one data set needed to be excluded due to extremely low data quality (resulting samples: $n = 9$ from test track study and $n = 37$ from the two simulator studies). A tendency for lesser PRC in automated drives was apparent in both test environments (see Figure 6-3). Mean overall PRC was very similar with 68 % ($SD = 13$) in the test track study and 70 % ($SD = 15$) in the simulator studies. The largest difference in PRC between samples was observed for manual drives (Δ_{man} = 7 %). Of higher interest were however differences in PRC during supervision of the automated systems, which amounted to less than 1 % in both feedback conditions (hands-on: $z = 0.49$, $p = .628$, $M_{test\ track}$ = 66 %, $SD = 14$, $M_{simulator}$ = 67 %, $SD = 12$; hands-off: $z = -0.37$, $p = .708$, $M_{test\ track}$ = 63 %, $SD = 12$, $M_{simulator}$ = 64 %, $SD = 16$). These results support the hypothesis that participants monitored automation similarly in both test environments. Note, however, that test track as well as simulator are both controlled test environments and a diversion of attention to self-selected secondary tasks was prohibited.

[16] Following Bortz & Schuster (2010), the t-Test can be expected to perform robustly with uneven sample sizes.

Mean speed after TOR for the test track study was comparable between conditions, with differences between means below 2.6 kph in all conditions (see Figure 12-6 in Appendix). Speed was more consistent over the different intervals on the test track than in the simulator, with the lead vehicle increasing speed after the TOR situation only slightly.

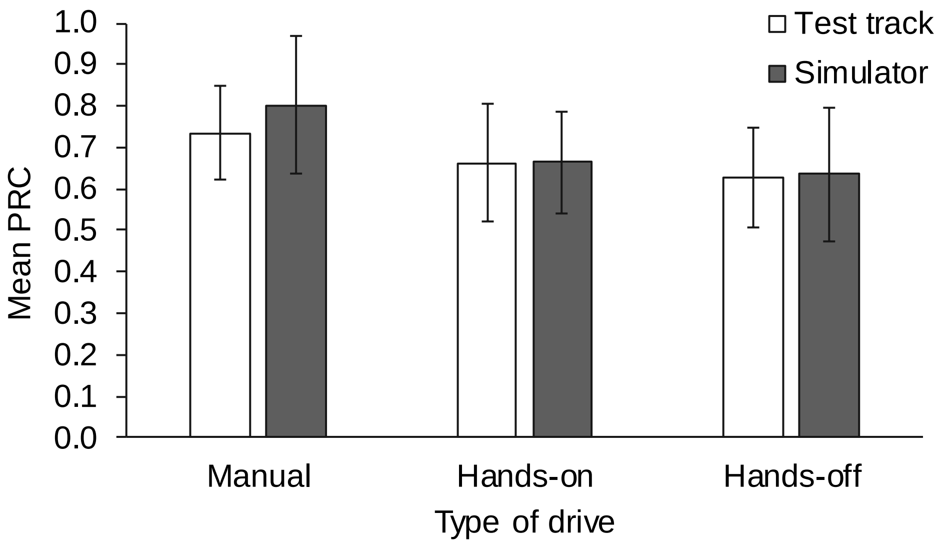

Figure 6-3: Mean PRC (and SD) during automation use and during the same interval of the unassisted, manual drive for the study conducted on the test track (*n* = 9) and for comparable scenarios of the simulator studies (*n* = 37).

$Steer_{max}$ values were consistently lower on the test track than in the reference scenario in the simulator (hands-on: $MD_{test\ track} = 14.85°$, $MD_{simulator} = 26.05°$; hands-off: $MD_{test\ track} = 16.28°$, $MD_{simulator} = 32.94°$). In Figure 6-4, SDST in the test track study is plotted in comparison to SDST in the reference scenario of the driving simulator studies. Similar to the non-significant difference in intervention times, differences to the baseline condition were much smaller on the test track, although still apparent in the first analysis interval. Whereas SDST was unaffected by method in unassisted driving (means for first interval; $M_{test\ track} = 2.87°$, $SD = 0.72$; $M_{simulator} = 2.27°$, $SD = 0.92$), less variation in steering input was observed after automation use on the test track compared to the simulator studies (hands-on: $M_{test\ track} = 4.60°$, $SD = 0.76$; $M_{simulator} = 8.57°$, $SD = 3.57$; hands-off: $M_{test\ track} = 6.19°$, $SD = 1.69$; $M_{simulator} = 12.00°$, $SD = 4.75$). Similar in comparison of methods was the fast convergence between conditions after the first analysis interval in the reference scenario.

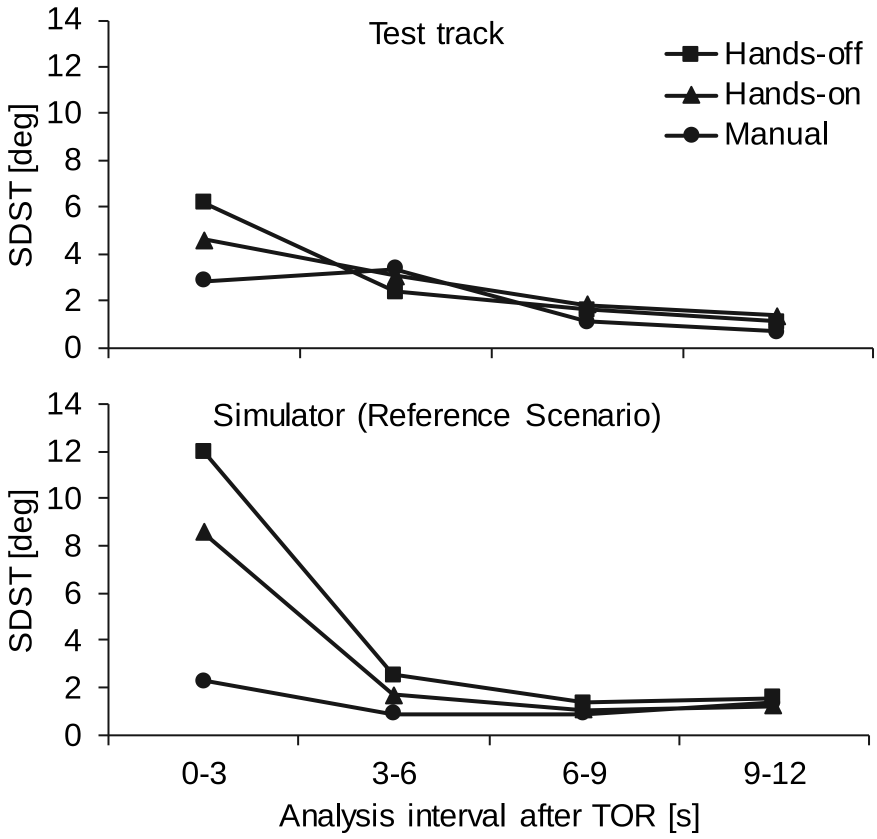

Figure 6-4: SDST aggregated over 3-s-intervals beginning at TOR for the test track study (upper graph; *N* = 12) and for the comparable reference scenario of the driving simulator studies (lower graph; *N* = 50).

For MLP (Figure 6-5), a similar pattern as for SDST became apparent with smaller initial differences between conditions (mean difference to unassisted driving in first interval; on: $\Delta_{test\ track}$ = 0.11 m; $\Delta_{simulator}$ = 0.59 m; off: $\Delta_{test\ track}$ = 0.19 m; $\Delta_{simulator}$ = 0.66 m). However, differences in the simulator stayed minimal after convergence, whereas the largest differences on the test track was observed after around 12 s of manual driving with anew convergence between conditions after 21 s. No such pattern was observed in any of the driving simulator studies. Likely due to the design of the test track (i.e., curvature changes), MLP varied more over time on the test track as in the simulator studies. Means in all three driving conditions indicated a change in absolute lane position over driving time, whereas the differences between automation conditions stayed negligible after 20 s. Overall, however, results on quality pointed in the same direction as data on takeover initiation. The test track data exhibited the same trend as data collected in the driving simulator, while indicating a smaller initial effect of automation on performance in short-termed takeover situations.

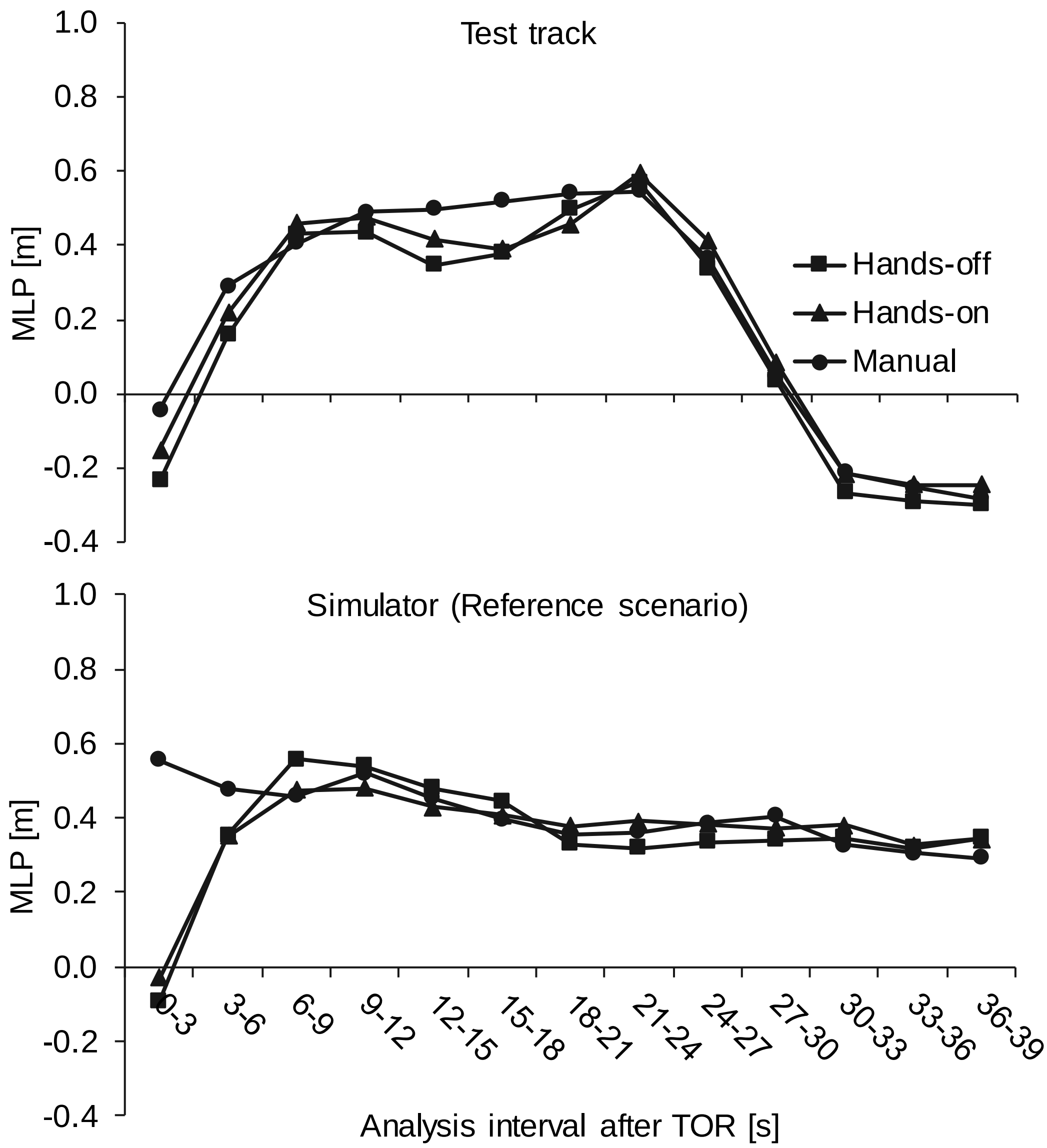

Figure 6-5: MLP aggregated over 3-s-intervals beginning at TOR for the test track study (upper graph; $N = 12$) and the comparable reference scenario of the driving simulator studies (lower graph; $N = 50$).

6.2 User Characteristics

For the following analyses regarding user characteristics, practice effects (see Chapter 6.3) and attitudes before and after use (see Chapter 7), data was aggregated over studies manipulating hand posture within subjects, thus excluding data from Study 4. Overall, a sample of 78 participants was considered. Smaller sample sizes were considered for some analyses, such as for comparisons concerning post-automation driving performance in the reference scenario (excluding the construction zone due to roadway design) and for eye tracking data (due to issues with data quality). A description of the overall sample ($N = 78$) can be found in Table 6-1. In addition to analyses of short-termed takeover scenarios, interaction strategies for different levels of prior ADAS experience were analysed for early-visibility TOR scenarios (*N* = 25 of Study 4).

When not limiting analyses to the reference scenario, driver-performance measures for the within-subject studies were averaged for each driver over all situations encountered with the same hand posture before calculating sample means for manual, hands-on and hands-off driving each. For some of the analyses, only the reference scenario (as defined in Chapter 5.1) was considered, for example for the analysis of practice effects, where averaging data over multiple takeover situations would have been inappropriate.

Table 6-1

Sample description of the aggregated data set.

Variable	Descriptive statistics
First automated drive[a]	Hands-on: n = 40 Hands-off: n = 38
Age (in years)	M = 36, SD = 12, MD = 30 Range: 25 - 64
Gender	n_{male} = 67, n_{female} = 11
Driving experience (average km driven per year)	M = 16738, SD = 11407, MD = 15000 Range: 300 - 60000
ADAS experience[b]	M = 7.81, SD = 2.35, MD = 7 Range: 3 - 15

Note. Sample (N = 78) aggregated from all studies manipulating hand posture within subjects.
[a]Either as first or as second experimental drive, if the first drive was the manual drive.
[b]ADAS experience reported as summed experience points (0-16) for prior use of ACC, CC, LKA and TJA on scale as printed in Appendix 12.5.

The influence of user characteristics on system interaction was analysed with a focus on prior user experience. Experience with technology moderates the attitudes towards it (Bhattacherjee & Premkumar, 2004; Gold et al., 2015; Jensen, 2014) and experience with similar ADAS is thus expected to influence interaction behaviour, as was shown for practice effects by Zeeb et al. (2016). Special emphasis is put on monitoring behaviour, as ADAS users should, for example, be more aware of technological limitations relating to the need to supervise technology.

The sample was split in two subsamples for comparison between feedback conditions. Using the median of the overall sample (see Table 6-1), the split resulted in 43 participants with little experience (i.e., seven or less experience points; $M_{inexp} = 6.1$) and 35 experienced ADAS users (i.e., eight or more experience points; $M_{exp} = 9.9$). As was already shown for the sample of the online survey (Chapter 4), driving experience in general, defined here in kilometres driven per year, was correlated with prior ADAS use ($\rho = .40$, $p < .001$) and age ($\rho = .53$, $p < .001$). When using the median of general driving experience instead of prior ADAS experience to split the sample (N = 78) into two groups, 66.7 % of the sample were classified congruently as having little or much experience. Of the remaining 33.3 % being sorted differently for a split by driving experience than by ADAS experience, 17.9 % were classified as having much driving experience, but little ADAS experience. Due to the high concordance of the two variables, the literature reviewed and results of the survey (see Chapter 4), only prior ADAS experience was considered for further analyses.

To analyse whether prior experience – mostly with Level 1 automation (i.e., ACC, CC, LKA; SAE, 2018) – influences the monitoring behaviour when interacting with higher degrees of automation, PRC was compared between ADAS experience groups ($N = 56$; inexperienced users: $n = 25$; experienced users: $n = 31$) separately for manual, hands-on and hands-off drives. The descriptive tendency over systems was similar for both experience groups with the smallest PRC in the hands-off condition, as was the difference between experience groups in automated and manual drives. For automated drives, ADAS experience did not influence PRC (hands-on: $z = 1.42$, $p = .156$, $M_{exp} = 0.67$, $SD = 0.13$; $M_{inexp} = 0.62$, $SD = 0.15$; hands-off: $z = 1.57$, $p = .117$; $M_{exp} = 0.65$, $SD = 0.18$; $M_{inexp} = 0.57$, $SD = 0.14$). PRC in the manual drives differed by tendency in unassisted driving between experienced and inexperienced ADAS users ($z = 1.77$, $p = .078$), with a mean increase in PRC of 9 % for experienced users ($M_{exp} = 0.81$, $SD = 0.12$; $M_{inexp} = 0.72$, $SD = 0.21$). Gaze behaviour during automation use and unassisted driving was thus not dependent on ADAS experience in the studies conducted.

In accordance with the non-significant differences in monitoring behaviour, a similar timing of interventions was observed for both ADAS experience groups in short-termed takeover scenarios (hands-on: $M_{exp} = 710$ ms, $SD = 280$ ms, $M_{inexp} = 740$ ms, $SD = 290$ ms; hands-off: $M_{exp} = 930$ ms, $SD = 210$ ms, $M_{inexp} = 1020$ ms, $SD = 270$ ms). Although inexperienced users showed a slightly higher hands-off disadvantage, no significant difference between user groups was found, $t(76) = 0.95$, $p = .346$ (difference in FDI between hands-off and hands-on condition; $\Delta_{inexp} = 282$ ms, $\Delta_{exp} = 218$ ms).

For the comparison of performance between ADAS experience groups in short-termed takeover, the reference scenario was used (N = 60). Data from the test track study was excluded from the analysis of MLP due to the heavy influence of track design on this indicator, but was included in the analysis of SDST. Both experience groups showed the same descriptive tendency with large differences between automation conditions as well as overall large SDST in the first interval and fast convergence between conditions in the second interval (see Figure 6-6). Inexperienced users did not exhibit higher differences to their according baseline condition, indicating a similarly controlled takeover. The same pattern can be observed for MLP (i.e., convergence after around 6 s; see Figure 12-7 in Appendix).

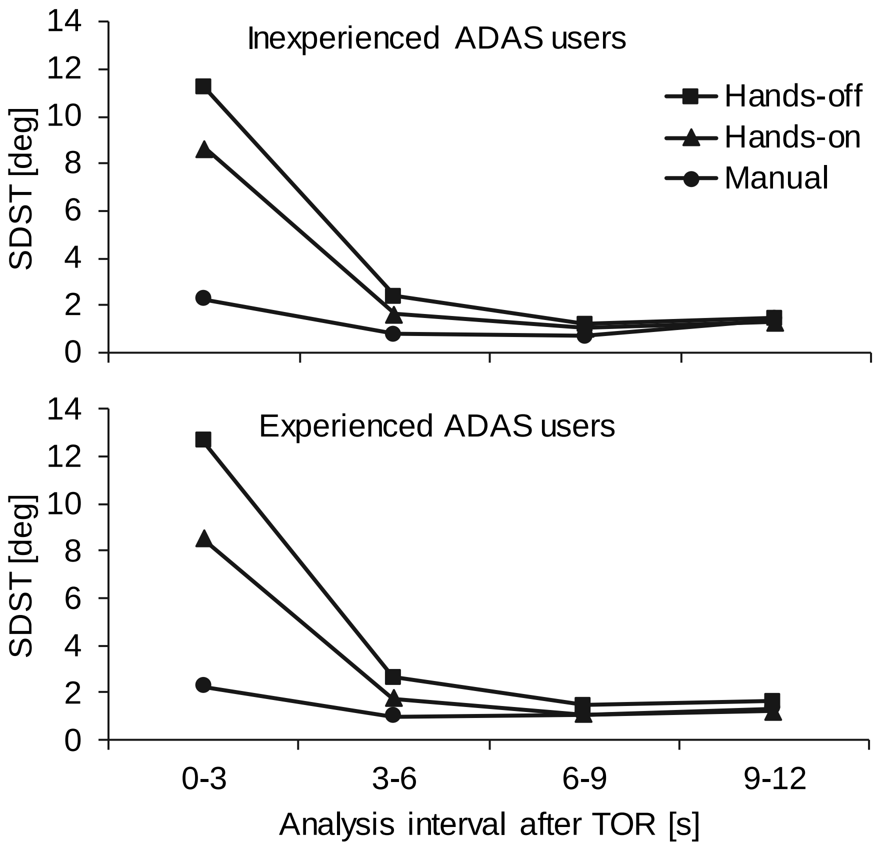

Figure 6-6: Mean SDST in the reference scenario, aggregated over 3-s-intervals for inexperienced (upper graph; n = 29) and experienced (lower graph; n = 31) users.

As short-termed takeover situations do not leave much room for applying different handling strategies, the sample of the study on early-visibility takeover (N = 25; hands-on and hands-off groups of Study 4, see Chapter 5.6) was used to test for different strategies when approaching a visible system limit. Summed ADAS experience (range of experience: 0-13 on ADAS experience scale from 0-16) correlated with the timing of driver interventions (ρ = .45, p = .025), indicating later driver initiated interventions with increased ADAS experience. Contrary to the correlation of ADAS experience and timing of interventions, SDST in the first interval after lane change, as a measure of controlled takeover behaviour, did not correlate with ADAS experience (ρ = .003, p = .988, n = 22). Additionally, analysis revealed a significant relationship between ADAS experience and a priori trust (ρ = .435, p = .038, n = 23), with higher levels of trust for participants with more extended prior ADAS experience.

6.3 Practice Effects and Performance Range

The larger, aggregated sample (Table 6-1) was used to analyse the stability of the effect of haptic feedback over levels of practice with PAD systems. Further, the main effect of haptic feedback on gaze and quality metrics was assessed. Finally, the range in driver responses to TOR was addressed to complement the assessment of haptic feedback effects by mean differences. In most of the analyses in this chapter, performance in the reference scenario, that is, for curve negotiation at low speed with no secondary task after a short automated drive (see Figure 5-2), was compared between subjects.

6.3.1 Practice Effects

Practice effects were addressed between subjects by comparison of performance in the reference scenario (*feedback*: on versus off) for different TOR positions (*position*: first, second, third or fourth TOR) within the experimental design, beginning with an analysis of the first encountered automation failure. All participants encountering the reference scenario as the first takeover situation were considered. The first takeover situation was either experienced in the first or the second of the three experimental drives, depending on the sequence of automated and manual driving conditions. Overall, 39 participants were selected for comparison (n_{on} = 18). The mean difference in intervention time between the two groups (on versus off) was 175 ms. Thus, under the combined effect of surprise and lack of practice with takeover handling, the difference between feedback conditions was significant only by tendency, $t(37)$ = -1.75, p = .089 (M_{on} = 938 ms, SD = 351; M_{off} = 1113 ms, SD = 272).

Furthermore, interventions in the reference scenario were compared for participants of the simulator studies (N = 50) according to the reference scenario's position within the three experimental drives. As each participant drove the reference twice (once hands-on, once hands-off) and handled two takeover situations per drive, the reference scenario could be encountered as the first, second, third or fourth takeover situation. For the third or fourth TOR-position (on or off), one encounter with the reference has prior been handled by each participant when using the other automated system (off or on). Note that in manual drives, the curve scenario did not stand out as a measurement situation and provided no practice in handling the occurrence of a TOR.

As shown in Figure 6-7, the difference between hands-on and hands-off was apparent between subjects for each position in the sequence of takeover situations. Practice was analysed for each feedback condition separately. Although interventions were descriptively faster with increased practice, no learning effects were found for hands-on takeover interventions, $F(3, 46) = 2{,}06$, $p = .119$. The effect of position was significant by tendency for hands-off automation use, $F(3, 46) = 2.44$, $p = .076$, but position did not reduce intervention times in a linear fashion over the number of exposures.

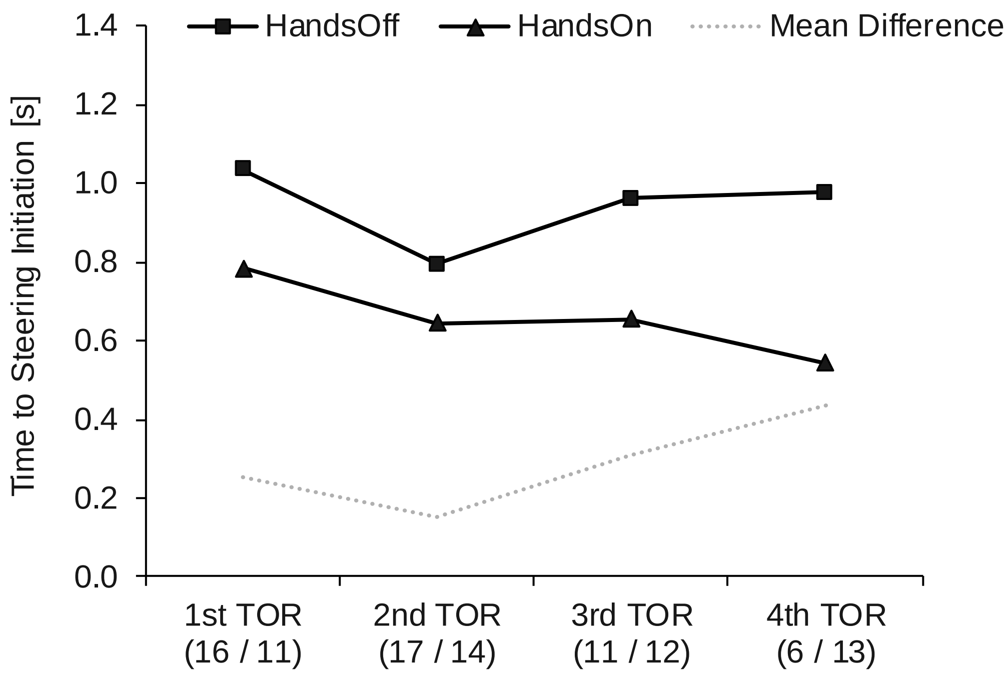

Figure 6-7: Mean intervention times in the reference scenario (curve scenario; $N = 50$ from simulator studies) for TOR positions within the sequence of the four takeover scenarios encountered during two automated drives. The number of participants considered for each condition is displayed in brackets under the x-axis (hands-off / hands-on). Each participant was considered once in each condition. The mean between-subject difference is displayed as a grey dotted line.

Monitoring behaviour did not differ depending on the sequence of automated drives (between-subject comparison of PRC_{on} and PRC_{off} in the first and second automated drive). The difference between means of the first and second automated drive was 3 % for both PAD systems. Comparing the differences between systems over the three experimental drives revealed similar changes in gaze behaviour over time. No systematic effect of time-on-task was found (mean differences in PRC between systems in the first / second / third drive for $n = 16$ to $n = 21$ participants interacting with each system; $\Delta_{\text{man-on}} = 0.08 / 0.15 / 0.13$; $\Delta_{\text{man-off}} = 0.14 / 0.15 / 0.17$; $\Delta_{\text{on-off}} = 0.06 / 0.00 / 0.04$).

The influence of practice on driver performance, approximated by the number of TOR situations encountered, was considered for SDST only (see Figure 6-8), as the driver dependency of MLP does not allow for a valid comparison between subjects. Again, only driving simulator data of the reference scenario taken from the aggregated sample for quality comparisons was used (N = 48), as the test track study included only two takeover situations instead of four.

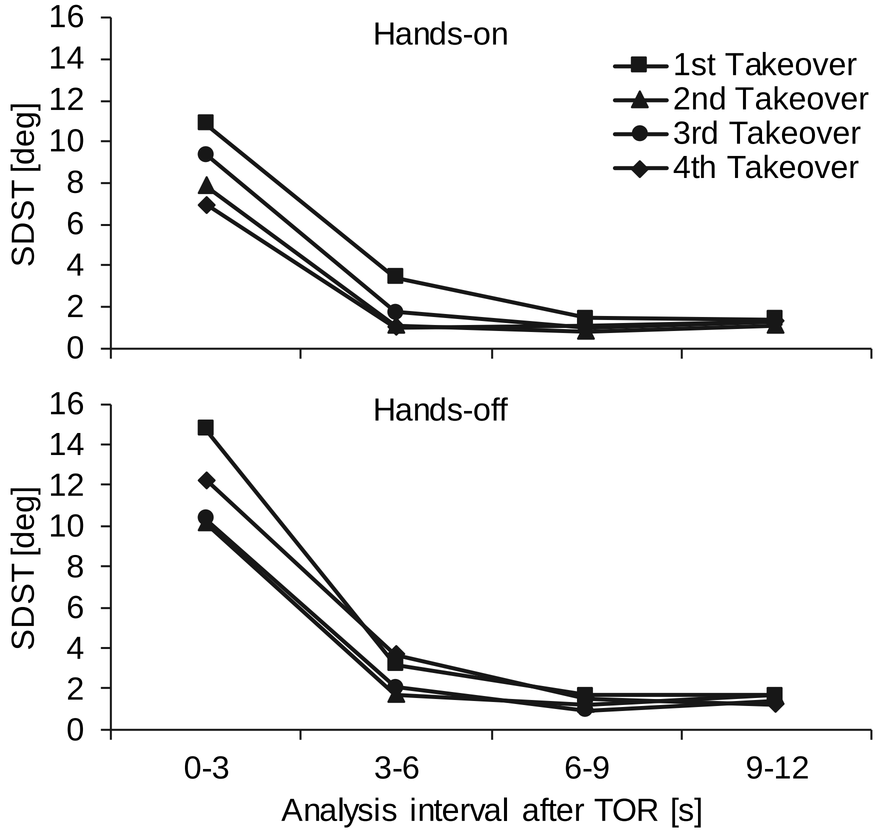

Figure 6-8: SDST aggregated over 3-s-intervals. Means provided for subsamples encountering the reference scenario at different positions during the trials. Each participant was considered in both feedback conditions. *Upper graph*: Hands-on reference scenario as 1st ($n = 16$), 2nd ($n = 19$), 3rd ($n = 12$) and 4th ($n = 13$) takeover scenario. *Lower graph*: Hands-off reference scenario as 1st ($n = 21$), 2nd ($n = 23$), 3rd ($n = 10$) and 4th ($n = 6$) takeover scenario.

A similar pattern with fast changes in SDST between the first and second analysis interval was found regardless of position for both automation conditions, with initially higher overall SDST after hands-off monitoring. SDST did not decrease with increasing practice, although the highest SDST in both automation conditions was observed for the group encountering the reference scenario as the first takeover situation. Regardless of position, little change was again observed in SDST after the initial effect of takeover had ceased (mean differences in each group between third and fourth interval between 0.02° and 0.45° compared to mean differences of 5.94° and 11.59° between the first and second interval). Neither performance nor timing did thus profit markedly from practice in short-termed takeover situations.

6.3.2 Effect of Feedback for the Aggregated Sample

The aggregated larger sample was used to analyse individual differences and the range in performance in each feedback condition. The type of the first automated drive (hands-on or hands-off) correlated with the resulting individual difference in intervention timings ($r_{pb} = .47$, $p < .001$), with larger differences between the two feedback conditions if the first automated drive was conducted with the hands-off condition. Further analysis revealed that all participants driving hands-off first took longer to initiate steering in this first drive than they took hands-on in the second drive. This subsample ($n = 38$; see Table 6-1) thus demonstrated the established hands-off disadvantage in timing ($\Delta_{\text{off-on}} = 390$ ms), in congruence with the main effect of feedback found in all of the simulator studies.

However, of the 40 participants driving hands-on first, $n = 13$ revealed the opposite pattern with faster interventions after hands-off monitoring, resulting in a significant *hands-on disadvantage* instead, $t(12) = 3.43$, $p = .005$ ($\Delta_{\text{off-on}} = -172$ ms). The remaining $n = 27$ participants driving hands-on first, even when combining hands-off monitoring with the maximum of practice in the repeated-measurements setting, initiated steering significantly later after hands-off monitoring than after hands-on monitoring, $t(26) = -5.30$, $p < .001$ (i.e., showing the established hands-off disadvantage; $\Delta_{\text{off-on}} = 259$ ms).

When compared to participants driving hands-on first and exhibiting a hands-off disadvantage ($n = 27$), the hands-on disadvantage group ($n = 13$) exhibited prolonged hands-on interventions, $t(38) = 3.23$, $p = .003$ ($\Delta_{\text{on}} = 277$ ms), but similar hands-off interventions, $t(38) = -1.83$, $p = .076$ ($\Delta_{\text{off}} = 154$ ms). The hands-on disadvantage found in the subsample ($n = 13$) can thus be attributed to hands-on performance differences.

For the complete sample ($N = 78$), hands-on interventions ranged from very fast, reflexive steering interventions (5th percentile of hands-on interventions; aggregated participant-wise for all TOR of the automated drive: 385 ms / 5th percentile for non-aggregated, single event intervention times: 330 ms) to times equalling those of around one second usually found after unexpected auditory warning signals for brake reactions (perception-reaction-time, as reviewed in Olson, 1989; 95th percentile for aggregated times: 1343 ms / single: 1340 ms). The range in intervention times was smaller for hands-off reactions, lacking very fast, reflexive steering interventions (5^{th} - 95^{th} percentile; aggregated interventions: 607 ms - 1388 ms / single intervention times: 503 ms - 1568 ms).

Based on the findings for the hands-on disadvantage group and the different ranges of intervention times, the effect of feedback was analysed over different levels of individual performance. Participants were sorted into four groups according to their performance after hands-on supervision. As responses were notably slower in the test track study (see 6.1.2), (mean) intervention times in the hands-on condition were split into quartiles for each experiment separately. Afterwards, the study-dependent quartiles (Q_1 – Q_4) were aggregated, with the fastest hands-on interventions of each study being aggregated into Q_1 and the slowest hands-on interventions being aggregated into Q_4. Hands-off intervention times were aggregated for the same subsamples to calculate the effect of feedback for each hands-on performance group (see Figure 6-9).

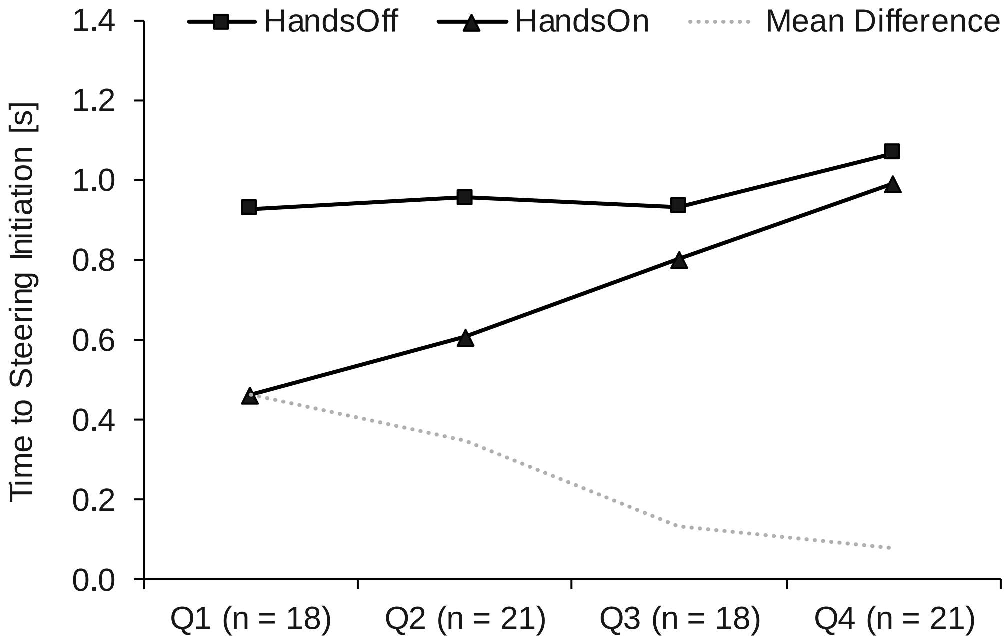

Figure 6-9: Mean intervention times for hands-on quartile-groups aggregated from participant-wise (mean) hands-on intervention times. The hands-off means were calculated for the same four subsamples to illustrate the resulting hands-off disadvantage in each quartile group. Grey dotted line: Mean difference between the automation conditions (hands-off disadvantage).

The four quartile-groups differed significantly in their hands-off disadvantage, $F(3, 74) = 9{,}73$, $p < .001$, $\eta_p^2 = .28$ (mean difference between hands-off and hands-on condition; M_{Q1} = 466 ms, M_{Q2} = 350 ms, M_{Q3} = 133 ms, M_{Q4} = 78 ms). With increasing time until FDI in the hands-on condition, the difference between automation conditions became descriptively smaller until becoming non-significant for Q_4, $t(20) = -1{,}13$, $p = .274$ (for all other quartile groups: $p < .05$; see Figure 6-9). In conclusion, participants with comparatively slow intervention times after hands-on monitoring (Q_4) showed no further significant impairment by hands-off monitoring.

Following the same procedure as described above for the definition of hands-on quartile-groups, participants were split additionally according to their hands-off performance (Figure 6-10). In comparison to the hands-on performance of each group, the hands-off disadvantage, although descriptively visible in each of the hands-off performance groups, differed again significantly between groups, $F(3, 74) = 12.46$, $p < .001$, similar to the results for the sample split based on hands-on performance.

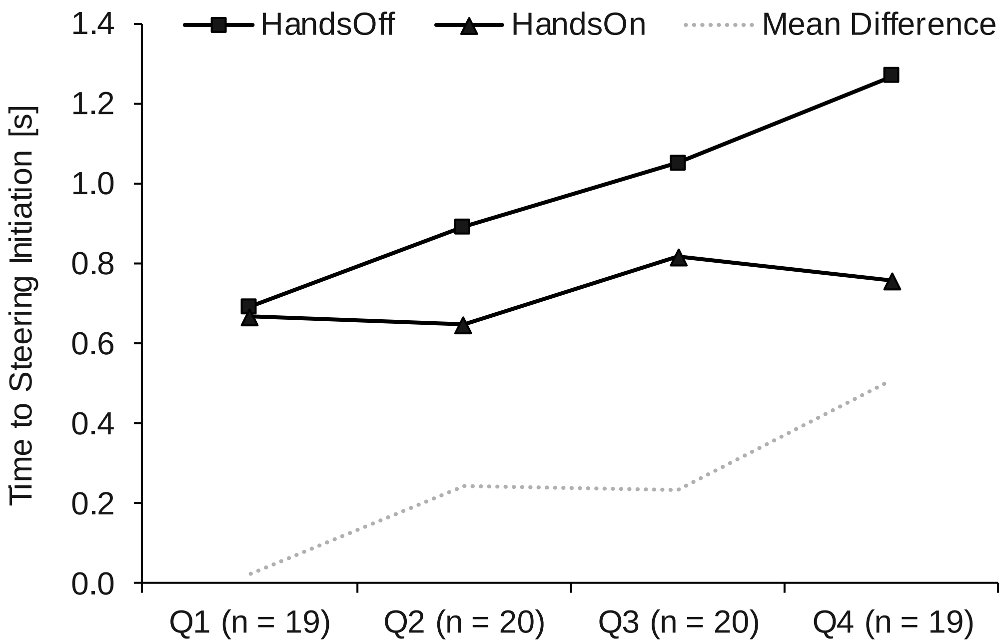

Figure 6-10: Mean intervention times for hands-off quartile-groups aggregated from participant-wise (mean) hands-off intervention times. The hands-on means were calculated for the same four subsamples to illustrate the resulting hands-off disadvantage in each quartile group. Grey dotted line: Mean difference between the automation conditions (hands-off disadvantage).

For the subsample with the fastest hands-off intervention times, the hands-off disadvantage was non-significant, $t(18) = -0.48$, $p = .635$ (Q_1; for all other hands-off quartile groups: $p < .01$). Participants with comparatively fast intervention times after hands-off monitoring (Q_1) showed thus no significant improvement in takeover interventions when monitoring hands-on (see Figure 6-10).

Note that the groups without significant hands-off disadvantage (hands-on split: Q_4, $n = 21$; hands-off split: Q_1, $n = 19$) as well as the hands-on disadvantage group ($n = 13$) were not composed of the same participants. Only 17 % of the aggregated sample ($n = 13$ of $N = 78$) intervened faster after hands-off than after hands-on monitoring. The difference in intervention times between feedback conditions was reversed only for this subsample with first use of automation in the hands-on condition, whereas hands-off monitoring always resulted in later interventions for the remaining 83 % of the aggregated sample. However, the hands-off disadvantage dissolved in case of comparatively slow hands-on or comparatively fast hands-off interventions. An assignment of cause to this effect, as for the hands-on disadvantage group, is however not possible. Nonetheless, it seems that monitoring hands-on is in itself no universal guarantee for faster interventions to TOR in PAD.

In addition to the analysis of intervention times as an indicator of takeover performance, the general effect of feedback on gaze data during automation use and post-automation driving performance was assessed. This was done for the reference scenario only, thereby excluding any influence of context manipulations or track design on the results. Using the aggregated data sample for gaze data (reference drive for both takeover scenarios in simulator and test track; $N = 56$), monitoring behaviour during automated drives and in the corresponding section in manual drives was compared between systems overall. The main effect of automation was significant, $X^2(2) = 65.17$, $p < .001$, with highest PRC for manual drives ($M_{man} = 0.77$, $MD_{man} = 0.82$, $SD = 0.17$; $M_{on} = 0.65$, $MD_{on} = 0.67$, $SD = 0.14$; $M_{off} = 0.62$, $MD_{off} = 0.64$, $SD = 0.16$). Pairwise tests between the three conditions indicated significant differences in PRC between all conditions (man - on: $z = -5.59$, $p < .001$; man - off: $z = -6.06$, $p < .001$; on - off: $z = -2.45$, $p = .014$). However, the mean difference between automation conditions, although significant, was very small compared to mean differences to unassisted driving (mean differences; $\Delta_{\text{man-on}} = 0.12$; $\Delta_{\text{man-off}} = 0.15$; $\Delta_{\text{on-off}} = 0.03$).

Initial $steer_{max}$ values in the reference scenario were significantly higher after hands-off monitoring ($z = 4.26$, $p < .001$; $N = 62$; $MD_{on} = 21.92°$, $MD_{off} = 29.63°$). Next to later interventions, hands-off monitoring went thus along with higher maximum steering wheel angles.

Using the available larger number of data sets in the aggregated sample, differences in SDST in the reference scenario ($N = 60$) were tested for significance, revealing a significant effect of automation type in the first analysis interval, $X^2(2) = 100.80$, $p < .001$ (all pairwise comparisons with $p < .001$). SDST was largest after hands-off monitoring (M_{off} = 10.84°, *SD* = 4.90) and smallest after manual driving (M_{man} = 2.39 °, *SD* = 0.91; M_{on} = 7.78°, *SD* = 3.58). A test of significance for the last interval analysed (9-12 s after TOR) revealed persistent differences between conditions, $X^2(2) = 7.43$, $p = .024$, with post-hoc pairwise comparisons showing significant differences only between the hands-off and manual condition ($p = .011$; M_{off} = 1.47°, *SD* = 0.75; M_{man} = 1.22°, *SD* = 0.88). However, the differences in steering behaviour between conditions, albeit still significant, became smaller with time on average, as was already described for the descriptive analysis of the studies on context variation.

A similar analysis of MLP differences between conditions (man, on, off; without data from the test track study) revealed again significant differences in the first interval, $X^2(2) = 74.38$, $p < .001$ (all pairwise comparisons with $p < .001$). Contrary to the analysis of SDST, no difference between conditions was found for the fourth interval, $X^2(2) = 0.50$, $p = .779$.

6.4 Overview on the Stability of Feedback Effects

Data from studies on short-termed takeover situations was aggregated to investigate the stability of effects over different user groups and methods. To validate the results of the simulator studies, a test track study was conducted. For the analysis of user experience, either in the sense of prior ADAS use or practice with specific PAD systems, data from the four studies on short-termed takeover situations (Studies 1-3 and Study 5) was used.

The test track study, albeit showing the same tendencies as the simulator studies, was characterised by smaller differences between the feedback conditions for all metrics alike. A descriptive comparison of monitoring behaviour in the two test environments revealed a similar pattern in overall PRC size as well as differences between conditions. Response times on the test track were longer, even more so in the hands-on condition which was characterised by fast and reflexive responses in the simulator studies. Steering behaviour was further characterised by less variance in all conditions. Albeit the implementation of an objectively similar scenario on the test track, a significant effect of haptic feedback was found only in the simulator studies. However, no contradictory effects between the methodologically different studies were found.

Overall, the range of intervention times was larger after hands-on monitoring. Contrary to the rather stable advantage found when monitoring with haptic feedback over studies, a minority of participants exhibited an individual hands-on disadvantage. Hands-on interventions were significantly slower for this subsample, likely due to the disproportionate effect of unexpected control transitions during first automation use on performance. Additionally, participants with comparatively slow intervention times in the hands-on condition or with comparatively fast intervention times in the hands-off condition did not show any effect of haptic feedback.

Prior ADAS experience did not change the timing of takeover interventions in short-termed takeover situations. In early-visibility takeover scenarios, more ADAS experience was associated with later steering interventions. No apparent relation of experience to the quality of steering behaviour was found. Although means were higher in all conditions for experienced users, no significant difference in monitoring behaviour during automation existed between experienced and inexperienced ADAS users in congruence with results on takeover initiation times. During unassisted driving, a tendency for more attention being attributed to the forward road scene was found with more ADAS experience. Recent experience with specific PAD systems (i.e., practice) and its relation to different levels of haptic feedback were analysed by comparing the driver interventions in a reference scenario implemented in all studies as well as in the first takeover experienced. The difference between feedback conditions was reduced to a tendency for the first takeover, but a general effect of practice over position in the study was neither found for any of the feedback conditions nor metrics used.

A lack of haptic feedback resulted in significantly reduced gaze attribution to the forward road scene not only in comparison to manual, but also to hands-on use. Furthermore, takeover intervention after hands-off monitoring was associated with higher and more persisting changes in steering input in comparison to unassisted driving.

7 Attitudes After System Use

The influence of system use on attitudes and system preferences was analysed using the aggregated sample described in Chapter 6.2. The focus of the following analysis is on changes in attitudes after use in comparison to before use (acceptance and attractiveness of the two systems; see Figure 7-1)[17]. Intention to use and comfort of use were compared between both systems as well as between two ADAS experience groups (inexperienced versus experienced ADAS users). Correlation coefficients of user characteristics (age, driving experience and ADAS experience), can be found in Chapter 6.2. For a detailed sample description, see Table 6-1.

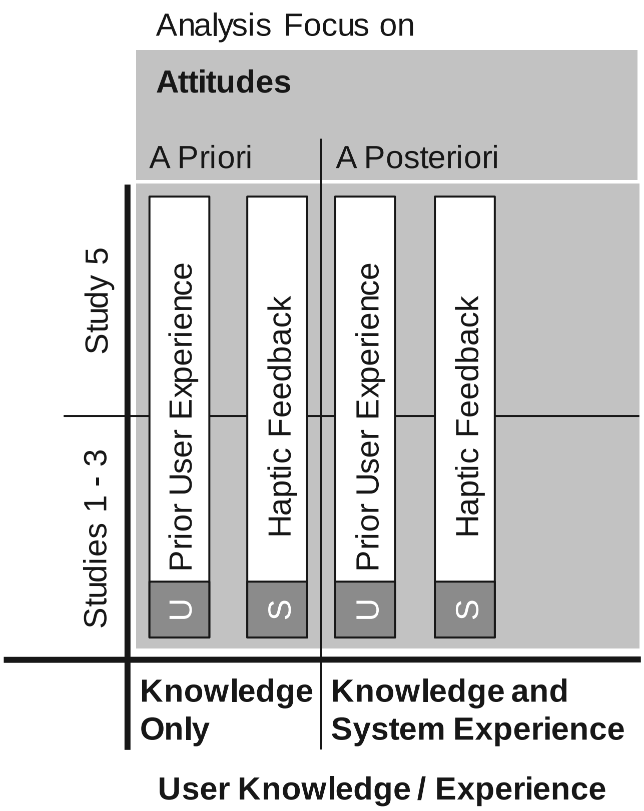

Figure 7-1: User (U) and PAD system (S) characteristics considered for changes in attitudes before (a priori) and after system use (a posteriori), analysed with aggregated data from simulator and test track studies.

Whereas attractiveness was measured on the same scale as usefulness and satisfaction (Van der Laan et al., 1997), comfort, intention to use, effort and drowsiness were measured on a scale allowing for a finer differentiation between

[17] Results reported in this chapter can be found in adapted form in Josten et al. (2016).

conditions. Overall, the observed ratings covered the full range provided by the scale (16-point scale with verbal anchors, see Figure 5-6).

Ratings of attractiveness were affirmative for both systems with mean ratings well above the scale mean (i.e., zero) before and after use. For both ADAS groups, ratings increased slightly after use, but more so in the experienced group (mean changes; inexperienced group: $M = 0.05$ for both on and off; experienced group: $M_{on} = 0.54$; $M_{off} = 0.12$). Overall ratings were similar after use ($M_{off} = 1.02$, $SD = 0.96$; $M_{on} = 1.09$, $SD = 0.88$), with only slightly higher ratings in the experienced user group (on: $M_{exp} = 1.06$, $SD = 1.08$; $M_{inexp} = 0.98$, $SD = 0.83$; off: $M_{exp} = 1.14$, $SD = 0.81$; $M_{inexp} = 1.05$, $SD = 0.95$).

For both sub-scales of acceptance (i.e., usefulness and satisfying), ratings increased after use, but more so for the inexperienced group (see Figure 7-2). This pattern was found for both automated systems alike. The effect of use on ratings was tested for each user group separately with the factors *time of measurement* (a priori versus a posteriori) and *system* (on versus off; 2x2 repeated measurements design), using the mean of the two subscales. The mean change in ratings was significant for inexperienced ADAS users, $F(1, 40) = 8.46$, $p = .006$, $\eta_p^2 = 0.175$ ($M_{pre} = 0.91$, $SD = 0.59$; $M_{post} = 1.04$, $SD = 0.57$), but not for experienced ADAS users, $F(1, 33) = 0.77$, $p = .783$ ($M_{pre} = 1.16$, $SD = 0.48$; $M_{post} = 1.17$, $SD = 0.51$). The difference between systems (*on* versus *off*) was non-significant in both user groups, $F(1, 40) = 0.69$, $p = .411$ (inexperienced; $\mathrm{M}_{on} = 1.00$, $\mathrm{SD} = 0.56$; $M_{off} = 0.95$, $SD = 0.59$) and $F(1, 33) = 1.20$, $p = .282$ (experienced; $M_{on} = 1.13$, $SD = 0.48$; $M_{off} = 1.21$, $SD = 0.51$). No significant interaction between time of measurement and system was found, $F(1, 40) = 1.02$, $p = .320$ (inexperienced) and $F(1, 33) = 0.15$, $p = .704$ (experienced). Additionally, ratings after use were compared for both systems (*on* versus *off*; see analyses for separate samples above for main effect of system) between ADAS experience groups, focusing on effects of the user group (*inexperienced* versus *experienced*). A significant effect was neither found for the main effect of ADAS group, $F(1, 74) = 1.82$, $p = .181$, nor for the interaction between system and ADAS group, $F(1, 74) = 0.45$, $p = .505$.

Comfort of use was rated after use only. Higher comfort ratings were assigned to the hands-off system as compared to the hands-on system ($M_{on} = 11.50$, $SD = 2.87$; $M_{off} = 12.29$, $SD = 2.20$). The difference between systems was significant ($z = -2.16$, $p = .031$). Experienced ADAS users did not rate the systems as more comfortable than inexperienced ADAS users ($z = -0.35$, $p = .724$; $M_{inexp} = 12.01$, $SD = 2.39$; $M_{exp} = 11.76$, $SD = 2.69$).

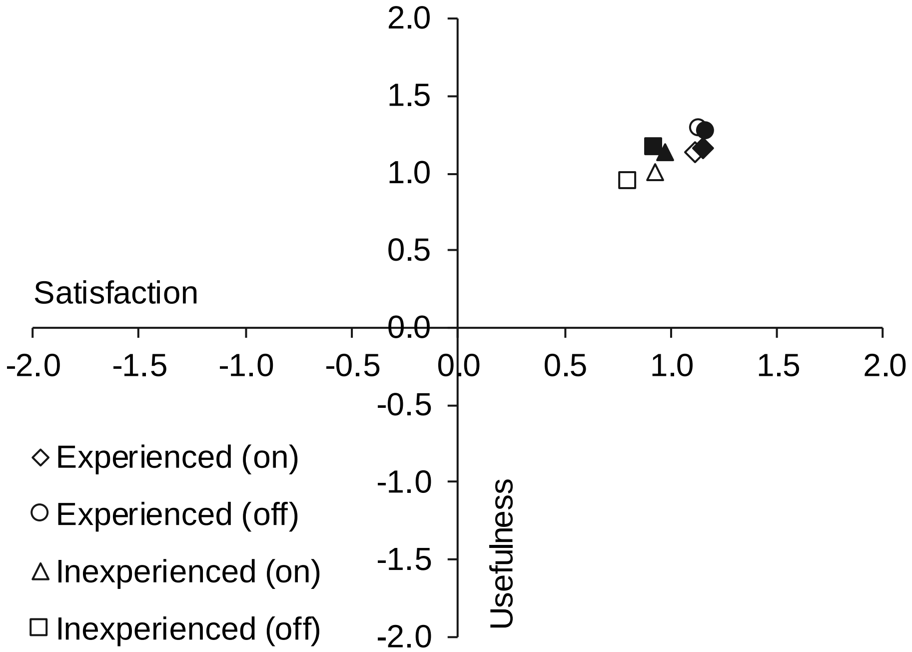

Figure 7-2: Acceptance (Van der Laan et al., 1997) before and after use of the hands-on and the hands-off PAD system for inexperienced and experienced ADAS users. Ratings before use are displayed as unfilled shapes.

Intention to use, respectively the intended frequency of use in case the system were available to users, was the same for both systems (z = -0.95, p = .342; M_{on} = 11.28, SD = 3.39; M_{off} = 11.38, SD = 3.05), albeit the significant difference in comfort of use. The influence of feedback was small regardless of prior experience (mean difference between off and on; Δ_{inexp} = 0.23; Δ_{exp} = -0.06). However, the overall intention to use any of the automated systems was significantly higher for experienced ADAS users (z = 2.35, p = .019; M_{inexp} = 10.56, SD = 3.71; M_{exp} = 12.29, SD = 2.88).

Effort related with and drowsiness induced by the drive were collected to compare the experience of automated driving to unassisted driving. Ratings on effort did differ between manual drives and both automated drives (man - on: z = -3.44, p = .001; manual - hands-off: z = -3.70, p < .001; M_{man} = 7.33, SD = 3.46; M_{on} = 5.35, SD = 3.00; M_{off} = 5.42, SD = 3.30), but not between automated drives (on - off: z = 0.38, p = .701). The same effects were observed for drowsiness induced by the drive (man - on: z = 2.53, p = .012; man - off: z = 2.51, p = .012; on - off: z = 1.40, p = .162; M_{man} = 5.81, SD = 3.59; M_{on} = 6.72, SD = 3.41; M_{off} = 7.00, SD = 3.93). The effects found illustrate the main effect of partial automation with lower effort induced by delegating the driving task to the vehicle at the price of higher drowsiness associated with the passive monitoring task.

8 How Users Cope with Partial Automation

The introduction of automation into the driving task provides, amongst other benefits, a means for increased user comfort (Stanton & Marsden, 1996) by relieving the driver at least partially of a complex and dynamic task. Current levels of vehicle automation (Level 1 and 2; SAE, 2018), which leave the supervision of system and driving environment to the user, automate primarily the motoric aspects of driving, but not the cognitive task of processing and understanding the driving situation. Since the user is required to stay involved in the driving task, explanatory concepts for driver performance in unassisted driving, such as hazard perception, vigilance or driver distraction, remain important in the context of partial automation. Insufficient monitoring of the system, a lack of system feedback or system transparency as well as inadequate mental models have been identified as problematic for the interaction of users with automation (Endsley & Kiris, 1995; Manzey, 2008). A comprehensive analysis of users' interaction with partial automation from initial contact over the repeated handling of takeover situations to the analysis of post-automation driving performance was pursued in this work under variation of haptic feedback from the system. Additionally, an analysis of user's attitudes towards two PAD systems, asking for different levels of haptic involvement during use, was conducted, as attitudes form the basis of behaviour (Ajzen & Fishbein, 1977; Lee & See, 2004) and have been connected to monitoring behaviour when using higher levels of automation (e.g., Beggiato et al., 2015).

The manipulation of haptic system feedback during partial automation, here the allowance to monitor with the hands on or off the steering wheel, constituted a focus in the studies conducted. Additionally, several contexts of use and different user groups were considered to derive a better assessment of the generalisability of feedback effects found. The focus on haptic feedback was chosen based on prior research showing that hands-off monitoring increases comfort and decreases workload in comparison to hands-on monitoring (Damböck et al., 2012; Othersen, 2016), providing a benefit to users that might thus help to establish such systems. However, a continuous hands-free use is currently both not admissible (Gasser et al., 2012; UNECE, 2017b) and its effect has further not been investigated to sufficient extent for monitoring during PAD. As feedback has been associated with a positive impact on the involvement with automated tasks (e.g., Parasuraman & Riley, 1997; see also Damböck et al., 2013), a lack thereof could result in an inadequate withdrawal from the driving task with negative effects on interaction quality. A degraded interaction quality could be expected either because of the lesser involvement with the driving task, resulting directly from the lack of haptic feedback, or from additional time consistently needed to establish motor readiness before takeover, but also from changes in the driver's intent to monitor based on assumed system capabilities (for a similar reasoning, see Gold, Damböck, Bengler, et al., 2013; Louw et al., 2015; Othersen, 2016; Parasuraman & Riley, 1997).

A research model was introduced (Chapter 3) to structure open questions on the interaction of users with partial automation. An online survey, four simulator studies and a test track study were conducted to investigate whether the lack of haptic feedback through hands-off monitoring has negative effects on interaction quality when using partial automation, specifically on takeover performance. By variation of contexts of use, levels of user experience, abstraction levels of the driving task and separate analyses for different interaction phases with automation, the relevance of haptic feedback for the interaction with partial automation and the stability of its influence were assessed. Specific variations in the context of use were chosen based on their assumed link to establishing or maintaining good SA during automation use. More precisely, the duration of use, the availability of secondary tasks and the complexity of the traffic situation were manipulated. Additionally, the impact of haptic feedback on planned, early visible control transitions was assessed.

In summary, an effect of haptic feedback on takeover performance was found in timing and takeover success as well as in the amplitude of driver interventions for nearly all analyses. Haptic feedback further influenced monitoring quality during automation use, defined by the amount of visual attention attributed to the forward road. This hands-off disadvantage could be shown over different levels of practice or prior experience with relevant ADAS.

Takeover Performance

The difference in intervention times found between feedback conditions was however small. Interventions were performed on average around 200 ms to 300 ms later without haptic feedback. This disadvantage was found rather consistently over different studies, contexts of use and ADAS experience levels. The magnitude of the difference between hand posture conditions resembled that reported in prior research on monitoring for automation failures with and without the need to position the hands on the steering wheel (Gold, Damböck, Bengler, et al., 2013).

The disadvantageous effect of hands-off monitoring was further apparent in takeover success. Overall, only 17 of the 390 control transitions analysed in this work resulted in collisions or lane departures (< 5 % of all cases). Of these, 12 occurred after hands-off monitoring. Although the number of unsuccessful control transitions is much lower than in studies focussing on unannounced system limits (i.e., 28 - 30 %; Victor et al., 2018), this distribution further illustrates the disadvantageous effect of hands-off monitoring. Finally, 13 of the unsuccessful control transitions, including the four cases in the hands-on condition, occurred for the first system failure encountered by the user (first failure effect, as discussed, e.g., in Strand et al., 2014). Less time for intervention, as in the high speed condition of Study 3 or the late-visibility condition of Study 4, resulted in higher numbers of lane departures, which is in line with the beneficial effects reported for increased time budgets for takeover in higher levels of automation (Gold, Damböck, Lorenz, et al., 2013).

Although different metrics congruently indicated the disadvantageous effect of lesser haptic involvement into the driving task, its specific effect on timing is significant only in most, but not in all of the studies and analyses conducted. The test track study revealed higher intervention times in comparison to the simulator studies, especially after hands-on monitoring, thereby decreasing the difference between feedback conditions. Furthermore, a minority of participants showed a reversed effect of hand posture on intervention times, with later interventions after hands-on monitoring in the first automated drive compared to performance after hands-off monitoring in later drives. Further, takeover timings of subsamples demonstrating either comparatively slow hands-on or fast hands-off intervention times were not influenced by haptic feedback. Although the majority of participants intervened slower after hands-off monitoring, the hands-off disadvantage can thus be overcome under certain conditions. Conversely, hands-on monitoring without distraction by secondary tasks does thus not produce distinctly optimal responses under all conditions, although the general pattern of response times and takeover success is in favour of hands-on supervision.

The intervention times found are generally similar to those reported in a study on warning signals in rear-end-collision scenarios (TTC = 3 s; Scott & Gray, 2008). An analysis of the range of intervention times after hands-on and hands-off monitoring revealed a larger range of response times after hands-on monitoring, due to the occurrence of very fast changes in steering wheel angles that were not observed after hands-off monitoring. This is likely due to the additional establishment of motor readiness necessary after hands-off monitoring. The overall fast changes in steering wheel angles indicate reflexive driver actions. Neukum and Krüger (2003) found reflexive steering interventions after approximately 150 ms in case of steering failures. Additionally, Schlick et al. (2010) reported comparable minimal reaction time values to visual, acoustic or haptic signals with the addition that these reaction times can only be achieved without further processing of the signal received.

The reflexive nature of interventions after TOR might explain why no effect of practice on intervention times was found, similar to results by Strand et al. (2014). Problematic for the analysis of practice effects might further have been the between-subject comparison necessitated by the variation of context and system factors within participants. Using small changes in steering wheel angle for detecting interventions might have increased the likelihood of measuring reflexive actions without prior cognitive processing of the situation. When defining a cut-off value for steering interventions, no relevant differences in timing when using different angles (e.g., 2° and 6°) or different indicators for lane change initiations were found, speaking for the validity of the metrics used. Further, the first intervention after the TOR is usually not interrupted again. Following these results, the difference to hands-off monitoring might be attributed to the time required to establish motor readiness before takeover (Gold et al., 2016; Zeeb et al., 2015), that is, to additional movement time (as also hypothesized by, e.g., Othersen, 2016). Green (2000), reviewing studies on brake reactions, assigned 200 ms to 300 ms of the overall response time to movement alone. The time needed to “move the foot from the accelerator to the brake pedal” (Green, 2000, p. 213) corresponds to the observed mean differences in steering onsets after hands-on and hands-off monitoring in this work.

Further, the stability of differences between hand posture conditions during monitoring over different contexts of use, as found in this work, speaks further in favour of a main attribution of the hands-off disadvantage to movement time (confirming a similar conclusion by Gold, Damböck, Bengler, et al., 2013). The findings therefore align with the context-independence in establishing motor readiness that has been observed in studies investigating executional aspects of the takeover process (Gold et al., 2016; Louw et al., 2015; Zeeb et al., 2015). The hands-off disadvantage was not influenced by prolonged durations of use (i.e., vigilance), the engagement in visual-manual secondary tasks or situations of varying complexity during use or during takeover. However, no main effects of the context variations on takeover initiation could be shown in the current studies. The conclusion of a context-independent hands-off disadvantage based on the current data alone might thus be premature as it cannot be assumed that the chosen context variations reliably changed the driver’s involvement with the driving task. Rather, the context variations chosen and the design of the takeover situation in particular might have prevented significant performance differences due to reduced monitoring efforts or levels of SA.

Monitoring Behaviour

Contrastingly to the timing of driver interventions speaking for an attribution of effects to movement time, the difference in monitoring behaviour outside of takeover situations implicates that monitoring hands-free should be considered to be more than just an additional step in the takeover process, that is, the establishment of motor readiness. Rather, the allowance of lesser haptic involvement with the driving task seems to invite a lesser degree of visual monitoring, which is in contrast with the increased subjective importance of monitoring reported in the a priori setting. However, the difference between feedback conditions was again comparatively small and could not be proven significant for a subsample analysed separately with regard to effects of longer durations of use (Study 1; Chapter 5.3), albeit similar descriptive tendencies. This lack in effect is likely related to lowered statistical power for the smaller subsample. Overall, the results on monitoring behaviour are in line with those reported by Othersen (2016) who also found a tendency for a higher attribution of gazes to the road when driving hands-on. Caution against a sole attribution of hands-off effects to movement time seems further indicated by other studies that found differences in the gaze orientation towards the road, but no significant effects in responses to critical events after hands-off monitoring (Gold, Damböck, Bengler, et al., 2013).

A significant decrease in PRC was found not only between hand postures, but also in comparison to manual driving, illustrating an established, disadvantageous effect of automation on driver attention in general (Damböck et al., 2013; Jamson et al., 2013; Morando et al., 2019). This effect on PRC is much larger than the difference between feedback conditions for automation, calling for a general need to motivate the driver to monitor the surroundings during PAD and the insufficiency of assuming the driver is monitoring the automated system simply based on established motor readiness, that is, when monitoring hands-on.

PRC values found for non-distracted unassisted driving in this work confirm findings from literature with values around 80 % (Victor et al., 2005). The percentages found for automated driving (62 % for hands-off and 65 % for hands-on) were similar to those observed in prior studies on PAD (Othersen, 2016) and to those for unassisted driving when attending to low difficulty visual secondary tasks, but notably higher than those observed for difficult secondary tasks in field and simulator testing during unassisted driving (Victor et al., 2005). Albeit a significant reduction of attention to the forward road during automation use, it might thus be argued that no critical level of distraction was achieved solely by introducing automation into the driving task. Whereas values for unassisted driving while engaging in secondary tasks in this work are comparable to those found by Victor et al. (2005) for the engagement with difficult tasks, the amount of attention attributed to the road ahead when offered a visual secondary task during automated driving is extremely low in comparison. Drivers do thus prioritise their attention differently during automation use in combination with voluntary secondary tasks, although no immediately negative effect on takeover performance could be shown in this work (cf. Study 2).

Similar to the effect on intervention times, haptic feedback (hands-on monitoring) provided a significant benefit in driver monitoring performance, but only in non-distracted conditions. Contrary to intentions reported a priori, drivers did not strategically counter the lack of haptic feedback with increased monitoring of the forward road when driving hands-off.

An indication that monitoring behaviour and takeover initiation are related to some degree is provided by the results of Study 4 (Chapter 5.6). Drivers who intervened ahead of a TOR did show a larger attribution of attention to the road centre beforehand, thereby indicating a higher involvement with the driving task and increasing the likelihood for a detection of relevant cues in the driving environment. However, as Victor et al. (2018) showed, even the combination of eyes on the road and hands on the steering wheel does not present an infallible countermeasure against unsuccessful control transitions during use of supervised automation. As discussed by the authors, safe automation use might not exclusively depend on high involvement in the driving task by visual supervision and haptic feedback alone. In this work, this was confirmed by the small number of unsuccessful control transitions after monitoring hands-on without additional visual distraction.

The analysis of gaze data in this work was restricted to one of the most safety-relevant directions of gaze during driving, that is, to on-path gazes to the forward road (Seppelt et al., 2017), using PRC as a frequently used and demand sensitive metric (e.g., Victor et al., 2005). Other gaze metrics, indicative of distraction or drowsiness, such as blink frequency (Körber et al., 2015), or a combination with off-road gazes as in Seppelt et al. (2017), might provide a more detailed insight into general or safety-related changes in monitoring behaviour without haptic feedback and should supplement the PRC-metric in future analyses.

The significance of the much lower levels of PRC during automation in general and under visual distraction in specific is challenged by the fast intervention times observed in short-termed takeover situations, indicating that visual attention to the roadway might not be a necessary precondition for a successful takeover in case of a visual-acoustic TOR alerting drivers to the takeover need. This is in line with the conclusion by Zeeb et al. (2015), stating for takeover after acoustic TOR that “there seems to be no influence of visual driver distraction on the time at which the driver establishes motor readiness to take over the vehicle" (p. 220). The significance of visual attention alone was further questioned by Victor et al. (2018), as stated above, for system limits encountered without a system announcement.

Feedback-Independent Effects of Automation

Intervention times for short-termed takeover were compared mostly between the two automation conditions, as any changes in driving performance during unassisted driving cannot be related to a specific event such as the TOR. Study 4 (Chapter 5.6), assessing lane change initiations during automated as well as unassisted driving, replicated a common adverse effect of automation. In comparison to manual driving, later interventions were found after automated driving in situations where an anticipation of the necessary lane change, accompanied by a change in automation status (i.e., system limit), was possible. Drivers intervened faster if provided with less time between visibility of the event and the event itself, but only so after hands-off monitoring (for a similar connection between time-criticality and intervention times, see Gold, Damböck, Lorenz, et al., 2013; for unassisted driving, see Scott & Gray 2008). The effect of situation visibility was not present in the unassisted drives, where equally fast interventions were observed regardless of the time provided, and also non-significant after hands-on monitoring. This indicates that situation-aware drivers perceived the early-visibility condition as urgent enough for immediate action, thereby validating the scenario design chosen to investigate driver interventions in view of system limits. Following Gold, Damböck, Lorenz, et al. (2013), the use of available time, found here most notably after hands-off monitoring, may be interpreted in the sense that drivers used the available time to prepare for a more controlled, safer takeover when being OOTL after hands-off monitoring.

For short-termed interventions in Study 4, no difference between automation conditions or to unassisted driving was found. The lack of effect between feedback conditions, in opposition to the other simulator studies and albeit similar descriptive tendencies, might be related to a lack in statistical power for the between-subject design of this study. The non-significant difference to manual driving for time-critical lane changes, surprising at first, might further be related to the TOR alerting the automation users to the situation (cf. Thiffault & Bergeron, 2003), whereas no advantageous warning was given to the manual group. Likely, automation effects are partly countered by the alerting effect of the TOR, whereas manual drivers, being provided with no warning, need time to classify the situation as requiring immediate action, thereby reducing differences between manual and automated conditions. Warning signals were found to reduce intervention times in manual driving when compared to a no-warning condition (Scott & Gray, 2008; see also Baumann et al., 2008, on predictable events). This considered, the equally fast interventions in the manual group in the time-critical scenario appear to reflect the high level of attention attributed to the driving situation.

However, Scott and Gray (2008) found the effects of warning signals in rear-end collisions to be significant only for uncritical driving scenarios (TTC of 5 s), but not for more urgent situations (TTC of 3 s). “This finding makes sense because drivers have little time to react to late warnings, [. . .] and most likely reflexively release the accelerator and quickly transition to the brake” (Scott & Gray, 2008, p. 272). In Study 4, the adverse effect of automation (with TOR as warning) on intervention times in comparison to manual driving (without warning) was found for uncritical intervals (early visibility, TLC > 6 s) only, but not for more urgent situations (late visibility; TLC > 2 s). As the TOR was issued at the same TLC in both scenarios, it cannot account for the effects found here. Rather, fast responses to the situation were found in the manual drive in both conditions, whereas drivers using automation hands-off did either wait for the TOR to be issued out of a lack of understanding for the need of action (Victor et al., 2018) or took longer to assess the situation (Gold, Damböck, Lorenz, et al., 2013). Haptic feedback, being additionally associated with a slightly higher visual involvement in the driving task, seems to discourage very late interventions with more time available. However, the same descriptive tendencies were found in both automation conditions.

Considerations of the TOR as a Driver Alert

Apart from influencing the difference in intervention times after assisted and unassisted driving in urgent, late-visibility situations, the decision to investigate mostly situations with an alerting TOR might be relevant for a number of (non-significant) results observed. Especially context variations targeting drivers' levels of SA by influencing their monitoring behaviour might have been affected. In PAD, a warning is provided when the system classifies the current situation as outside the system's functional limits or capabilities, such as in case of missing lane markings (Ford Motor Company, 2019; for steering assistance in general, see also UNECE, 2017b). The TOR as a notification of system status presents thus a valid and common use case. Due to the alerting character of the TOR, this is likely not the most challenging way to encounter a system limit from the user's perspective. The TOR might have obliterated effects of SA loss and resulted, as discussed above, in reflexive driver interventions. The announcement, alerting drivers to the situation, might explain why no effects of vigilance or additional visual distraction on interventions were found. As Thiffault and Bergeron (2003) hypothesised, "ruptures of monotony should have a positive effect on driving performance by allowing a temporary restoration of alertness and vigilance" (p. 385). Similarly, the reflexive nature of establishing motor readiness after acoustic TOR was noted by Zeeb et al. (2015), as quoted earlier.

Without a TOR provided, haptic feedback might further serve as an indicator for a change in system status, being beneficial to detect a mismatch between a required change in steering angle and the actual steering wheel angle. Results by Kerschbaum et al. (2014) for HAD indicate that the *visual* perception of steering wheel movements may not be a relevant feedback channel when using automation. The difference between hand postures might thus increase if an unannounced system failure or limit needs to be detected during curve driving. In this case, hands-on users might be able to detect the missing steering wheel movement that would not immediately be detected visually by hands-off drivers, even when disregarding the lesser overall visual involvement found. For the current studies, the steering wheel, being in a straight position at TOR, provided no additional haptic takeover cue during disengagement for hands-on supervision. Both feedback groups had therefore to rely on the acoustic TOR alone and feedback can only be considered as a beneficial factor for keeping the driver in the loop and maintaining SA *during* use of automation in the current studies.

As discussed by Damböck (2013), steering wheel movements occurring during highway driving are rather small and might not induce a better involvement with or feedback from the driving task. The higher amount of gazes attributed to the road found in this work might however indicate that being forced to keep the hands on a largely static steering wheel might provide at least slight advantages for the responsible supervision of such systems. An experimental investigation of different magnitudes of steering wheel movement and on the resulting monitoring quality in future research might be insightful in this regard.

Contexts of Use

Apart from the effects of the alerting TOR on the context manipulations applied in the studies, the lack of changes in monitoring performance during prolonged use of automation indicates that this specific manipulation might not have induced an effect on drivers, as stated earlier. Other studies found effects on the detection of takeover need (Othersen et al., 2014) or on gaze behaviour (Feldhütter et al., 2017) after similar driving durations. Slight, but unpredictable adaptations in driving speed, meant to increase the realism of the traffic jam scenario, might have reduced monotony (or predictability; cf. Thiffault & Bergeron, 2003). Another possible reason for the lack of effect is that the high traffic density in the traffic jam conditions might have countered the effects of prolonged driving in the otherwise monotonous driving environment. Körber et al. (2015) suggested to use decreased traffic densities for the investigation of time-on-task effects. In their study, automation use of more than 40 minutes did not result in visible changes in a reaction time task, in line with results by Feldhütter et al. (2017), but in significant effects on measures of passive fatigue, such as blink frequency, blink duration and pupil diameter, as well as in increased mind wandering (Körber et al., 2015).

A possibility is thus that the intervention time metric and the gaze metrics used were not sensitive to the effect in question. Longer durations of unassisted driving have also been shown to influence the quality of driving, but not hazard perception before (as reviewed by, e.g., Matthews, Saxby, Funke, Emo, & Desmond, 2011). Goncalves, Happee and Bengler (2016) found effects of drowsiness after use of conditional automation only in form of degraded lateral control during takeover, but not for gaze metrics or other driving performance metrics. The descriptive analysis of steering input and lateral position conducted in Study 1 of this work (Chapter 5.3) suggests however no general effect of time-on-task on lateral control, although differences between the conditions were slightly larger after longer durations of automation use.

The voluntary engagement with intermittently offered secondary tasks enabled the investigation of driver monitoring strategies when using automated systems. On the one hand, the prioritisation of the driving task was visible in a lowered engagement with the secondary task when encountering a TOR. On the other hand, the challenges of visual driver distraction became apparent in changed driving strategies, namely increased distances to the lead vehicle, during unassisted driving and in unstable MLP values observed while engaging in a secondary task, as well as in largely decreased monitoring efforts during automation use. These effects of visual secondary tasks are in line with prior findings (e.g., Cnossen, Meijman, & Rothengatter, 2004). The motoric involvement with the driving task, binding at least one hand to the steering wheel, does not prevent drivers more effectively from a continuous and adverse engagement in visual-motoric tasks during automation. A focus for future studies might be the use of continuous instead of intermittent secondary tasks, with the research focus shifted from takeover performance to the inclination to engage in secondary tasks during hands-off driving, while measuring performance for self-detected system limits. Likewise, the level of trust in the system should be controlled for in future research, to provide a possible explanation for differences observed in monitoring behaviour and secondary task activity (Beggiato et al., 2015; Muir & Moray, 1996).

Changes in traffic density as a means to increase the demands of maintaining SA have been implemented in prior studies on automation, with differences more apparent the larger the differences in the number of vehicles present in the experimental conditions (e.g., Körber et al., 2016; Jamson et al., 2013; Radlmayer et al., 2014). For the current work, traffic density was manipulated in the wake of the manipulation of driving speed to establish two realistic use cases, one being the reference scenario of traffic jam driving. Takeover performance was thus compared for two different driving situations instead of two traffic density conditions. The manipulation led to a higher number of unsuccessful control transitions at higher speed, as interventions were not performed faster after TOR albeit a lower TLC in this condition, which might be associated to the overall fast level of interventions observed. Monitoring efforts during automation use where higher in the high-speed condition, albeit a largely reduced traffic density in this scenario. The higher driving speed thus countered the expected effect of more attention to the forward road under higher traffic density which was found in prior studies (Jamson et al., 2013). Further, the presence of a lead vehicle, necessary to obscure the takeover scenario, is known to increase PRC values (Morando et al., 2019) and might have largely reduced any effect of lowered traffic densities on the neighbouring lanes. In accordance with the study by Jamson et al. (2013), gaze data during unassisted driving did not appear to be influenced by the traffic situation, whereas the traffic situation influenced the monitoring behaviour of automation users. For unassisted driving, a ceiling effect might also be assumed based on the very high PRC values in all conditions, preventing a further increase when driving at higher speeds.

Post-Automation Driving Performance

The effects of haptic feedback on intervention times and gaze behaviour are generally in line with prior research (Damböck, 2013; Gold, Damböck, Bengler, et al., 2013; Othersen, 2016). Its influence on post-automation driving performance, although of general interest in the context of automation (de Winter et al., 2014), has however not been in the focus of prior studies. The (descriptive) analysis of post-automation driving performance focused on three indicators of lateral control, specifically on SDST, SDLP and MLP, with SDST being the most sensitive indicator by directly reflecting driver input. As for timing, the performance after hands-off monitoring indicated a slightly less smooth return to unassisted driving performance levels. Compared to hands-on automation use, differences to baseline performance (unassisted driving) were observable for longer after takeover.

Regarding takeover quality, higher initial steering movements were found consistently after hands-off monitoring, an effect that might also be attributed to the higher time-criticality due to later interventions when monitoring hands-off. This link between available time and resulting quality is fostered by the results of Study 4, showing more controlled driver actions with more time provided for takeover (see also Gold, Damböck, Lorenz, et al., 2013). Overall, the analysis of driving performance did not reveal any tendencies that were not observable in other metrics considered.

Zeeb et al. (2016), using similar metrics, namely the average deviation from the lane centre during 10 s after takeover, and scenarios, that is, curve driving with faded lane markings and wind gusts, found effects of secondary tasks on quality, but not on timing. In the current work, quality was analysed mainly descriptively and with a slightly different focus, that is, regarding the return to individual driving performance (i.e., the end of correction; Eckstein, 2015) instead of an overall statistical comparison of indicators between conditions. It is further possible that effects of context variations on driving performance were not immediately apparent in the mainly descriptive analysis.

Pinpointing the decay of post-transition effects, following a theoretical model by Eckstein (2015), proved difficult as effects on post-automation driving performance were generally short. The fast alignment with unassisted driving performance was the same after different driving manoeuvres (i.e., curve negotiation and lane changes). Early assimilation of performance aligns with results by Morgan et al. (2018), who found the most persistent effects after automation for longitudinal accelerations and pedal input, but near-immediate alignment in SDLP and steering input. Longitudinal metrics were not considered feasible in the current work due to the behaviour of surrounding traffic that was not being controlled and might have influenced driving behaviour. Future research should however address context and user effects on speed, accelerations and distance, not only for re-gaining manual performance, but also to test for behavioural adaptations after prolonged use of PAD systems (e.g., Skottke et al., 2014). However, post-automation effects have been found as well for lateral performance metrics (Louw et al., 2015; Merat et al., 2014). Automation use has been associated with more variation in driver input in the first ten seconds after control resumption in case of lead vehicle braking manoeuvres (Louw et al., 2015). After HAD, differences between experimental conditions have even been found for up to 40 s (Merat et al., 2014).

SDLP, as a measure highly stable within drivers (Verster & Roth, 2011), constitutes a promising candidate for the assessment of a return to manual driving performance. Problematic for the use of SDLP values was however the short duration of calculation intervals in this work (cf. Verster & Roth, 2011), as the theoretically observable size of changes in lane position is a function of time and driving speed. Due to the overall short duration of effects, longer intervals might allow for a more insightful use of this metric, but for no better insights into the point where reference performance is re-established.

The measures considered for post-automation driving performance are dependent on road curvature (Knappe et al., 2006). This was addressed by assessing the influence of haptic feedback through differences to a manual baseline, which was conducted with equivalent conditions such as speed and curvature, instead of using absolute values for comparison of different experimental conditions. Further, speed was tested for differences between conditions after takeover and found to be similar between conditions in the different analysis intervals.

Attitudes and User Factors

The subjective ratings align with the small effect of haptic feedback on timing and monitoring performance in that the direction of differences in takeover criticality, significant only in Study 1, pointed in the same direction as all other metrics. Overall, the subjective impression of the two feedback options resulted in little difference for acceptance or the intention to use. General automation effects found, indicating lesser effort and higher drowsiness reported for automation use in monotonous contexts, align with prior research (see, e.g., the review by de Winter et al., 2014), but ratings did again not differ between the two feedback conditions. However, a higher rating of comfort during hands-off supervision was established in line with findings by Damböck et al. (2012) and Othersen (2016).

Prior experience was a relevant factor for a priori attitudes towards automation as well as for the intention to use automation after gathering experience with the system at hand, in line with the hypothesis that past behaviour might be a good indicator for future behaviour (Kyriakidis et al., 2015). Experience with PAD, addressed by comparing a priori and a posteriori ratings, the former being based on information and prior ADAS experience, the latter including actual experience with the system at hand, changed ratings only for inexperienced ADAS users. This is in line with other authors describing notable changes in attitude towards technology to take place early on in the interaction process (cf. Bhattacherjee & Premkumar, 2004; Feldhütter et al., 2016). Rating changes, occurring only in the inexperienced ADAS group, could indicate that practical experience with a system is influential primarily for attitudes in *first contact*, as stated by Szajna and Scamell (cited after Bhattacherjee & Premkumar, 2004). The ADAS experienced user group likely based initial ratings on expectations build on prior use, leading to a closer match between initial expectations and current experience.

Experienced drivers and ADAS users, with the two variables being highly correlated, monitored the forward road as one of the most safety-relevant areas for visual attention (Seppelt et al., 2017) more thoroughly during unassisted driving, but only descriptively so during automation. As already stated for the discussion on PRC values in general, both user groups tended to exhibit rather high levels of attention to the road ahead, although less than during manual driving. Driving experience has been found to influence gaze patterns, more precisely, fixation sequences, during manual driving in prior studies (Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003), but enhanced hazard perception of experienced drivers has been attributed to the more elaborate mental models of experts (Horswill & McKenna, 2004). Expertise in this work was defined by prior ADAS experience, but different ADAS systems are not quite the same when it comes to system limits and behaviour. This might be the reason for the descriptively largest difference found in unassisted driving, due to the higher general driving experience of experts, but not during automated driving, due to a similarity in mental models for the current systems being new to both ADAS experts and novices. Experience with similar systems, following Hoff and Bashir (2015), influences initial learned trust, but not dynamic learned trust. As discussed by the authors, the influence of prior experience on automation reliance is dependent on the match between experiences made and the specific system at hand. Establishing a sufficient mental model for takeover situations with ADAS has been reported to take up to three weeks (Weinberger et al., 2001). The lack of difference between ADAS-experience groups during automation use might thus be due to the fact that the interaction time with the systems was short during the studies conducted, leaving little room to verify assumptions based on prior ADAS-experience.

In line with findings of Scott and Gray (2008), no relation between experience and intervention times to a warning signal was found. ADAS experience was however related to later interventions in situations that allowed the self-detection of takeover need. Increased ADAS experience of participants was further related to a higher level of initial trust, which has in turn been connected to later interventions in prior research (Payre et al., 2016). The gaze data collected did unfortunately not allow for a more specific analysis on the detection of visual takeover cues in addition to takeover timing and monitoring in general, which might have helped to interpret the effect found. Other studies found enhanced scanning for road hazards with 6 s TOR time and greater driving experience (Wright et al., 2016). Importantly, later interventions with higher expertise did not lead to a change in takeover quality, confirming that later interventions alone do not allow for the conclusion of a less safe interaction with automation (cf. Gold, Damböck, Lorenz, et al., 2013).

Another relevant user factor, addressed in the a priori setting, is gender. The finding that women tend to be more sceptical towards new technology (e.g., Kyriakidis et al., 2015; Payre et al., 2014, 2016) was replicated in this work. This finding might be related to the fact that women in the sample drove less and had lesser prior experience with ADAS, preventing them from enhancing their mental model with updated information on the use of ADAS, potential benefits and challenges of use and thereby from building trust. The finding that trust increases by use if experience does not contradict the initial mental model (Beggiato & Krems, 2013) was replicated in this work (Study 4 in Chapter 5.6; see also Rudin-Brown et al., 2003).

In the current studies, the same initial information was presented to all users. Sceptical user groups might however need a higher degree of information to trust automation and experienced users might be interested in specific information not requested initially by inexperienced users. The high amount of detail assumedly needed for a useful, system-specific mental model (cf. Victor et al., 2018) might also diminish the influence of prior ADAS experience on takeover performance, especially for unannounced system limits. User-specific information or the offer of supervised practice or training when buying vehicles capable of PAD might help in resolving potential barriers for initial use by building trust and adequate mental models.

In summary, the influence of haptic feedback was rather small and similar over different user groups investigated in this work regarding attitudes before and after use, as well as behavioural measures. Future work should however consider unannounced system limits to investigate the influence of different mental models for user groups, contexts and systems on monitoring behaviour and takeover performance.

Choice of Method and Experimental Procedure

The aim of analyses, the metrics used and the design of the scenarios tested, such as the accelerations perceived by the driver, should be taken into account when selecting the appropriate methodological approach. For this work, relative comparisons between conditions were in the focus of analyses, allowing for the valid use of driving simulators (Kaptein et al., 1996). The test track study resulted in overall higher intervention times than those observed in the simulator, especially after hands-on monitoring. Consequently, the reference scenario resulted in significant differences between feedback conditions in the simulator, but not on the test track. The test track available for the study featured banked corners, with free-force driving at 34 kph on the target lane. Although driving speed in the study was slightly higher (40 kph), the track design might have influenced both subjective criticality, found to be very low on the test track, and steering behaviour of drivers. The subjective assessment of the situation might also have been influenced by the lower complexity of the traffic situation, with less elements to be considered on the test track albeit a similar ‘technical’ setup of scenario and system. Additionally, the test track study featured a much smaller sample, making the detection of small-sized effects difficult.

A test track still constitutes an abstract situation of use, albeit a higher absolute validity of vehicle data compared to a driving simulator. Important as the testing on the track might be for a valid assessment of interaction safety, the relative assessment of changes in user behaviour and ratings might be achieved by easier means and, in the case of traffic jam assistance, even with a higher immersion by using a driving simulator equipped with a vehicle mock-up. The results for gaze analyses support the hypothesis that drivers took the driving task equally serious in both environments. However, for an absolute assessment of the quality of post-automation driving performance instead of relative assessments as conducted in this work, the higher validity of test track studies and field operational tests is essential. The more controlled performance after takeover on the test track compared to the simulator studies provides a first indication that the chosen method influenced the absolute size of effects found in steering behaviour after automation.

In contrast to a similar approach on the role of haptic feedback by Louw et al. (2015) for HAD, where visual disengagement was strictly controlled by the amount of information displayed in the simulation, the variations applied in this work enabled the additional investigation of effects on voluntary monitoring behaviour, that is, on drivers’ strategies to counter the higher physical disengagement from the driving task. Instead of visual information, hand posture was strictly controlled to allow for the experimentally controlled analysis of this variable. In unassisted driving, "the assessment of hand positions on the steering wheel has been gaining validity for the measurement of drivers' perceived risk and mental workload" (Bianchi Piccinini et al., 2014, p. 208), with strategic changes in hand posture indicating changes in the situational assessment of the current driving environment, although the study cited failed to show an effect during ACC use. For PAD, Naujoks et al. (2015) found drivers to place their hands on the steering wheel even before a TOR was issued. Although resulting in a stable disadvantage under experimentally controlled conditions, monitoring hands-off might not present a problem for interaction quality as long as drivers stay attentive to their environment and understand their responsibility during PAD. This prerequisite might however at least be questioned based on the lower attention to the road centre found in this work and the later interventions observed in case of early-visibility scenarios.

Drivers were found to adapt their visual monitoring behaviour to the driving context, both positively and negatively, with increased efforts for higher speeds and with reduced efforts during secondary task engagement. Driver state monitoring and a reminder to stay attentive seem necessary for PAD. Different approaches have been suggested to motivate the driver to monitor the system more thoroughly, for example by displaying system uncertainty (Beller et al., 2013). Another possibility might be a frequent or state-dependent issuance of supervision reminders (Victor et al., 2018).

This work did not provide insights into adaptive changes in hand posture that might benefit takeover performance or into voluntary and adaptive changes in hand posture after undergoing training on system capabilities. However, even the combination of driver state monitoring in terms of visual attention and haptic involvement as well as extensive driver training to enhance drivers' mental models does not necessarily result in perfectly safe control transitions in case of unannounced system limits (Victor et al., 2018). Training in the study by Victor et al. (2018) was however only theoretical. Changes in the attitudes of inexperienced users and a descriptively higher occurrence of unsuccessful control transitions for the first system failure encountered indicate that practical use with a system might be needed instead to achieve a match between trust, according supervision strategy and system capabilities, as also discussed by Victor et al. (2018). This is underlined by the distinction in levels of trust in automation introduced by Hoff and Bashir (2015) who differentiated between initial learned trust before use and dynamic learned trust during use.

In this work, no training, but necessary information was provided before use, as the relevance of correct mental models has also been shown in prior research (Beggatio

& Krems, 2013). As shown in the a priori survey, initial information does influence drivers' attitudes and their intended interaction behaviour. Mental models, pre-existing knowledge, and trust need to be considered for partial automation as they define automation reliance (Hoff & Bashir, 2015), the more detailed – and likely the less theoretical – the better. "A key component of driver engagement is related to cognitive control (understanding the need for action), rather than purely visual (looking at the threat) or having hands on the wheel" (Victor et al., 2018, p. 1113). Future studies should thus target the cognitive component of system-enforced haptic involvement and its influence on trust and context-specific mental models in more detail. Indeed, the effect of haptic feedback seems, all else being equal, to be an influencing factor on users' attitudes, their monitoring engagement and their takeover performance, but its effect is minor throughout.

9 Summary and Outlook

Automating the driving task, at least in specific driving situations such as traffic jams, is widely considered a promising approach to make driving more comfortable. However, automation risks, as is also well known from other contexts, arise when responsibilities are shared between human and automation (see, e.g., Manzey, 2008). For partial automation, being in the focus of this work, the driver is relieved solely of the motoric aspects of the driving task, but left with the responsibility to supervise the system in its current environment (SAE, 2018). This requirement imposes the maintenance of continuously high levels of SA on the driver. For the sake of a safe interaction of users with partial automation, it is important to consider how the design of such systems and the context of automation use influence the interaction of users with automation. Relevant for an assessment of the consequences of introducing partial automation into traffic are users' attitudes, their execution of the monitoring task and, ultimately, the consequences that arise for interaction behaviour in case of control transitions between automation and human driver.

This work considered the role of haptic feedback during PAD, with hands-off supervision being a currently restricted comfort feature, in different contexts and phases of interaction for selected user groups. Starting from intended behaviour based on initial information over experiencing multiple control transitions to the re-establishment of manual driving performance, the relevance of haptic feedback in PAD was assessed by a combination of experimental methods. Three major research questions with a focus on haptic feedback in interaction with PAD systems were addressed in this work:

RQ 1: Does the lack of haptic feedback during hands-off monitoring influence the assessment of or the interaction with partial automation?

The option to remove the hands from the steering wheel during monitoring had little effect on the subjective assessment of automation before use. Except for higher comfort ratings found for hands-off monitoring, no a posteriori differences were associated with haptic feedback. All in all, attitudes were positive regardless of haptic feedback. Results from four simulator studies showed that hands-off monitoring leads to later and less controlled takeover interventions than hands-on monitoring.

RQ 2: Is the influence of haptic feedback stable over different user groups, driving contexts, methodologies and phases of interaction?

The effect found for interaction behaviour, albeit small, was found over different contexts and phases of interaction. It was further independent of the user characteristics investigated. Metrics for timing and quality of interventions as well as gaze data, albeit a higher initial importance to monitor reported by users, agree on the disadvantageous effect of hands-off monitoring on performance. At least the stable hands-off disadvantage found in the timing of takeover interventions motivates an attribution to the additional movement time needed. The already small effect of haptic feedback found in the driving simulator was further reduced in a data set collected on a test track, challenging the benefit of continuous hands-on monitoring in trade-off to a gain in comfort of use.

RQ 3: Does PAD influence the quality of post-automation driving performance?

As for timing, the descriptive analysis of performance after hands-off monitoring indicated a slightly less smooth return to normal, unassisted driving performance. Overall, the after-effects on driving performance were rather contained, with noteworthy effects lasting seldom for longer than a few seconds before converging towards manual performance levels. However, the higher variation observed in steering input immediately after control transitions, especially after hands-off monitoring, might in individual cases present an increased risk for lane departures.

Especially prior experience and gender were connected to attitudes towards automation before use. A consistent influence of prior ADAS experience could however not be shown in behavioural data. Actual use of automation was found to be less relevant for attitudes when prior experience with similar systems existed, possibly related to the stability of mental models being based on experience instead solely on information. In general, the effects of automation in comparison to manual driving outweighed the effects of levels of haptic feedback during automation use for subjective, physiological and performance metrics.

Based on the results presented in this thesis, it remains to be defined whether hands-off monitoring during PAD constitutes a “risk[ ] that remain[s] ‘below the danger threshold’ and which can therefore be justified” (Gasser, 2016, p. 532). As hands-off driving provided indeed an increase in comfort for users, the necessity for continuous hands-on driving might be questioned from a user perspective. Overall, the physical contact with the steering wheel in itself might not be essential for successful takeover behaviour in cases where a TOR is issued by the system. This conclusion is based on the predominantly fast interventions observed in this work in response to announced system limits and on the reflexive, context-independent establishment of motor readiness observed by other authors (Gold et al., 2016; Zeeb et al., 2015). The slightly lower PRC values observed during hands-off monitoring even in the controlled experimental environment of the studies conducted might indicate a lesser effort invested into the monitoring task, which could be problematic whenever the driver needs to detect any ill-adapted system behaviour by himself. However, the differences found in PRC are considered too small to be of an immediate relevance for monitoring quality and were further not supported by subjective ratings, with similar effort associated with automation use regardless of haptic feedback.

Early-visibility of a control transition enhances the quality of transitions, but planning for a takeover requires drivers to understand the system's capabilities in current and evolving driving situations. However, drivers who monitored automation hands-off waited longer to intervene, even when the need for action should have been clear based on the situation and the mental model invoked by instruction. Although the current results cannot rule out an alternative explanation, seeing later interventions as an indicator for better preparation (Gold, Damböck, Bengler, et al., 2013), and although late intervention albeit earlier visibility resulted in a collision only once, considering mental models seems essential for PAD. This is also stressed by two participants not or not adequately reacting to the TOR, although stating to have noted the acoustic signal and having received clear instructions for encountering a TOR. These cases occurred although only a very basic mental model was required in the experimental setting for connecting the instructed TOR to the need to act, whereas rather complex interdependencies of situation variables and system capabilities can be expected in real world driving contexts in comparison.

Albeit being found in different metrics and contexts of use, the hands-off disadvantage can be overcome under certain conditions. Apart from the disproportionate influence of surprise on hands-on performance, which likely exceeded the additional movement time needed after hands-off monitoring, the preconditions for a hands-on-like performance after hands-off monitoring would have to be investigated further to be fully understood. One task for future research might therefore be a more detailed consideration of different hand postures in terms of distance to the steering wheel or a classification of the readiness to intervene to account for comparatively fast hands-off interventions.

The additional manipulation of context did not yield any general effect on takeover behaviour after PAD. This might be due to the decision to investigate only announced control transitions, thus providing a TOR as feedback of the change in system status. Reacting to an auditory cue might be largely reflexive and less indicative for low levels of SA as a self-detection of takeover need (Louw et al., 2015; Zeeb et al., 2016; see also Scott & Gray, 2008). Future work should focus on driver-detected intervention needs to quantify any effects of lesser monitoring associated with hands-off supervision on takeover safety. This approach might also be indicated for investigating factors connected to a loss of SA, such as vigilance effects (Othersen et al., 2014).

Even without prior notification, a majority of control transitions in partial automation can be performed successfully (Victor et al., 2018). In this work, with notification by TOR, an even higher number of more than 95 % successful takeovers was found. Supervision reminders and a more extensive driver training, as applied by Victor et al. (2018), might enhance this percentage even further. Overall, these numbers suggest that the majority of interactions with partial automation can indeed be successful, at least in experimental contexts and under the assumption of an adequate mental model, albeit hands-off monitoring accounting for 12 of the 17 unsuccessful takeover attempts observed. However, this percentage should be contrasted with results on driver behaviour in less artificial, less controlled contexts of investigation.

This work focussed on an experimentally controlled approach to define the influence of haptic feedback and context factors on user interaction with partial automation. This approach benefits the clear assignment of cause to the effects found, but might be questioned regarding the validity for user behaviour observed in daily traffic. Naturalistic driving studies, using, for example, a safety driver as a fall-back option, would thus be needed to validly assess adaptations in visual monitoring behaviour, hand postures (see also Victor et al., 2018) and, if methodologically possible, voluntary driver distraction during automation use in different contexts and for user groups with different levels of trust and prior ADAS experience. The voluntary choice for hands-off driving and the engagement in naturalistic secondary tasks might provide insights into the anticipative establishing of motor readiness or into monitoring strategies based on users' mental models. Optional, driver controlled automation use has similarly been investigated in relation to fatigue and stress in monotonous driving situations (Neubauer, Matthews, Langheim, et al., 2012).

In summary, user characteristics, variations in context and haptic feedback proved to be only of minor influence in case of a correctly processed TOR. Achieving a sufficient understanding of the system, specifically a mental model being detailed enough to be of use in complex traffic scenarios, might thus to be a key challenge for the safe use of partial automation, in agreement with Victor et al. (2018). Context-dependent, adequate monitoring behaviour and anticipative changes in hand posture, based on the user's mental model of the system's capabilities, together with system feedback provided on changes in system status, should assure that almost all control transitions can be handled successfully. How and if such a degree of strategic user behaviour and system reliability can be achieved for partial automation is a challenge for future research. None of the factors investigated in this work, however, opposes in itself the responsible and safe use of partial automation.

10 Formula Symbols and Indices

ACC	Adaptive cruise control (SAE Level 1 automation)
AD	Automated driving
ADAS	Advanced driver assistance system (SAE Level 1 automation)
X^2	Chi-square (test statistic)
CAN	Controller area network
CC	Cruise control (SAE Level 1 automation)
COG	Centre of gravity
Δ	Delta (e.g., difference between sample means of conditions)
η_p^2	(Partial) Eta square (effect size metric)
F	F-value (test statistic)
FDI	First driver input
HAD	Highly automated driving (SAE Level 3 automation)
IQR	Interquartile range
LCI	Lane change initiation
LKA	Lane keeping assist (SAE Level 1 automation)
M	Mean
MD	Median
MLP	Mean lateral position (in own lane)
N	Number of participants (overall)
n	Number of participants (in a sub-sample)
OOTL	Out-of-the-loop
p	p-value (level of significance)
PAD	Partially automated driving (SAE Level 2 automation)

PRC	Percent road centre (percentage of valid gaze data points in the road centre which is defined as a circle of 8° radius)
Q_x	$\text{Quartile}_{1\text{-}4}$
RQ	Research question
SA	Situation awareness
SD	Standard deviation
SDLP	Standard deviation of lane position
t	t-value (test statistic)
THW	Time headway
TJA	Traffic jam assist (combination of ACC Stop and Go and LKA; SAE Level 2 automation)
TLC	Time to line crossing
TOR	Takeover request
TTC	Time to collision
z	z-standardised value / test statistic

11 References

Abendroth B., & Bruder, R. (2009). Die Leistungsfähigkeit des Menschen für die Fahrzeugführung [Human performance capability for vehicle guidance]. In H. Winner, S. Hakuli, G. Wolf (Eds.), *Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (pp. 4-14). Wiesbaden: Vieweg+Teubner.

Abendroth, B., Schreiber, M., Bruder, R., Maul, S., & Maul, D. (2012). Neue Ansätze zur Beurteilung der Fahrsimulatorvalidität [New approaches for evaluating the validity of driving simulators]. *Zeitschrift für Arbeitswissenschaft, 66(1)*, 1-11.

Ahlstrom, C., Kircher, K., & Kircher, A. (2009). Considerations when calculating percent road centre from eye movement data. *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 132-139. doi:10.17077/drivingassess-ment.1313

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84(5)*, 888-918.

Alessandretti, G., Amditis, A., Metzner, S., Johannson, E., & Fahrenkrog, F. (2014). *InteractIVe. Accident avoidance by active intervention for intelligent vehicles* (Deliverable D1.9. Final Report. Grand Agreement No. 246587). Retrieved from http://www.interactive-ip.eu/index.dhtml/docs/interact-IVe_SP1_20140506v1.2-D19-Final_Report.pdf

Arbeitsgruppe Straßenentwurf (2008). *Richtlinien für die Anlage von Autobahnen (RAA)* [Rules of reference for the design of motorways]. Köln, Germany: Forschungsgesellschaft für Straßen- und Verkehrswesen e.V.

Baumann, M. R. K., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments. Critical Issues in Driver Interactions with Intelligent Transport Systems* (pp. 253-265). Heidelberg, Germany: Springer.

Baumann, M. R. K., Petzoldt, T., Groenewoud, C., Hogema, J., & Krems, J. F. (2008). The effect of cognitive tasks on predicting events in traffic. In C. Brusque (Ed.), *Proceedings of the European Conference on Human Interface Design for Intelligent Transport Systems* (pp. 3-13). Lyon, France: Humanist Publications.

Baumann, M. R. K., Rösler, D., & Krems, J. F. (2007). Situation awareness and secondary task performance while driving. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics, HCII 2007* (pp. 256-263). Heidelberg, Germany: Springer.

Beggiato, M. (2015). *Changes in motivational and higher level cognitive processes when interacting with in-vehicle automation* (Doctoral dissertation, Technische Universität Chemnitz, Germany). Retrieved from https://nbn-resolving.org/urn:nbn:de:bsz:ch1-qucosa-167333

Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J. F., Othersen, I., & Petermann-Stock, I. (2015). What would drivers like to know during automated driving? Information needs at different levels of automation. *7. Tagung Fahrerassistenz*. doi:10.13140/RG.2.1.2462.6007

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour, 18*, 47-57.

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving driver automation interaction: An approach using automation uncertainty. *Human Factors, 55(6)*, 1130-1141.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly, 25(3)*, 351-370.

Bhattacherjee, A., & Premkumar, G. (2004). Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test. *MIS Quarterly, 28(2)*, 229-254.

Bianchi Piccinini, G. F., Prati, G., Pietrantoni, L., Manzini, C., Rodrigues, C. M., & Leitao, M. (2014). Drivers' hand positions on steering wheel while using adaptive cruise control (ACC) and driving without the system. In D. de Waard et al. (Eds.), *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2013 Annual Conference*. Retrieved from https://www.hfes-europe.org/wp-content/uploads/2014/06/Piccinini.pdf

Bolstad, C. A. (2001). Situation Awareness: Does it change with age? *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, 272-276. doi:10.1177/154193120104500401

Bortz, J., & Schuster, C. (2010). *Statistik für Sozial- und Humanwissenschaftler* (7th ed.) [Statistics for social and human sciences]. Berlin, Germany: Springer.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effect of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23(4)*, 309-316.

Brookhuis, K. A., de Waard, D., & Fairclough, S. H. (2003). Criteria for driver impairment. *Ergonomics, 46(5)*, 433-445.

Brookhuis, K. A., van Driel, J. G., Hof, T., van Arem, B., & Hoedemaeker, M. (2009). Driving with a congestion assistant; mental workload and acceptance. *Applied Ergonomics, 40(6)*, 1019-1025.

Buld, S., Krüger, H.-P., Hoffmann, S., Kaussner, A., Tietze, H., & Totzke, I. (2002). *Wirkungen von Assistenz und Automation auf Fahrerzustand und Fahrsicherheit* [Effects of assistance and automation on driver state and driving safety]. (Abschlussbericht Projekt EMPHASIS: Effort-Management und Performance Handling in sicherheitsrelevanten Situationen. Project 19 S 9812 7). Würzburg, Germany: Interdisziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg (IZVW).

Cnossen, F., Meijman, T., & Rothengatter, T. (2004). Adaptive strategy changes as a function of task demands: a study of car drivers. *Ergonomics, 47(2)*, 218-236.

Damböck, D. (2013). *Automationseffekte im Fahrzeug - von der Reaktion zur Übernahme* [Automation effects in the vehicle – from response to takeover] (Doctoral dissertation, Technische Universität München, Germany). Retrieved from https://mediatum.ub.tum.de/doc/1144567/document.pdf

Damböck, D., Farid, M., Tönert, L., & Bengler, K. (2012). Übernahmezeiten beim hochautomatisierten Fahren [Takeover times in highly automated driving]. *5. Tagung Fahrerassistenz*. Retrieved from https://mediatum.ub.tum.de/doc/1142102/1142102.pdf

Damböck, D., Weißgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance. In *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013)*, 1656-1661. doi:10.1109/ITSC.2013.6728467

Desmond, P. A., & Matthews, G. (1997). Implications of task-induced fatigue effects for in-vehicle countermeasures to driver fatigue. *Accident Analysis and Prevention, 29(4)*, 515-523.

de Winter, J. 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: Traffic Psychology and Behaviour, 27*, 196-217.

Dickie, D. A., & Boyle, L. N. (2009). Drivers' understanding of adaptive cruise limitations. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, 1806-1810. doi:10.1177/154193120905302313

Dogan, E., Deborne, R., Delhomme, P., Kemeny, A., & Jonville, P. (2014, March). Evaluating the shift of control between driver and vehicle at high automation at low speed: The role of anticipation. *Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment*. Paris, France.

Donges, E. (2009). Fahrerverhaltensmodelle [Driver behaviour models]. In H. Winner, S. Hakuli, & G. Wolf (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 15-23). Wiesbaden, Germany: Vieweg+Teubner.

Duncan, J., Williams, P., & Brown, I. (1991). Components of driving skill: experience does not mean expertise. *Ergonomics, 34(7)*, 919-937.

Eckstein, L. (2015, July). *Fundamentals and experiments on handing over the driving task*. TRB Conference on Automated Driving. Ann Arbor, MI.

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

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37(2)*, 381-394.

Engeln, A., & Vratil, B. (2008). Fahrkomfort und Fahrgenuss durch den Einsatz von Fahrerassistenzsystemen [Driving comfort and driving pleasure through driver assistance systems]. In J. Schade & A. Engeln (Eds.), *Fortschritte der Verkehrspsychologie. Beiträge vom 45. Kongress der Deutschen Gesellschaft für Psychologie* (pp. 275-288). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human Factors, 59(5)*, 734-764.

Eriksson, A., & Stanton, N. (2017). Takeover times in highly automated vehicles: Noncritical transitions to and from manual control. *Human Factors, 59(4)*, 689-705.

ERTRAC (2017). *Automated driving roadmap*. Retrieved from https://www.ertrac.org/uploads/documentsearch/id48/ERTRAC_Automated_Driving_2017.pdf

Feldhütter, A., Gold, C., Hüger, A., & Bengler, K. (2016). Trust in automation as a matter of media influence and experience of automated vehicles. *Proceedings of the Human Factors and Ergonomics Society 2016 Annual Meeting*, 2024-2028. doi:10.1177/1541931213601460

Feldhütter A., Gold C., Schneider S., & Bengler K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. In C. Schlick, S. Duckwitz, F. Flemisch, M. Frenz, S. Kuz, A. Mertens, & S. Mütze-Niewöhner (Eds.), *Advances in Ergonomic Design of Systems, Products and Processes* (pp. 309-318). Berlin: Springer.

Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., & Beller, J. (2012). Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. *Cognition, Technology & Work, 14*, 3-18.

Ford Motor Company (2019). *2019 Edge Owner's Manual* (3rd ed.). Retrieved from http://www.fordservicecontent.com/Ford_Content/Catalog/owner_information/2019-Ford-Edge-OwnersManual-version-3_om_EN-US_06_2019.pdf

Gasser, T. M. (2016). Fundamental and special legal questions for autonomous vehicles. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous driving. Technical, legal and social aspects* (pp. 523-552). Berlin: SpringerOpen.

Gasser, T. M., Arzt, C., Ayoubi, M., Bartels, A., Bürkle, L., Eier, J., . . . Vogt, W. (2012). Rechtsfolgen zunehmender Fahrzeugautomatisierung. Gemeinsamer Schlussbericht der Projektgruppe [Consolidated final report of the project group: Legal consequences of an increase in vehicle automation]. *Berichte der Bundesanstalt für Straßenwesen, F 83*.

Gold, C. (2016). *Modeling of take-over performance in highly automated vehicle guidance* (Doctoral dissertation, Technische Universität München, Germany). Retrieved from https://mediatum.ub.tum.de/doc/1296132/document.pdf

Gold, C., Damböck, D., Bengler, K., & Lorenz, L. (2013). Partially automated driving as a fallback level of high automation. *6. Tagung Fahrerassistenz*. Munich: Germany. Retrieved from: https://mediatum.ub.tum.de/1187198

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*, 1938-1942. doi: 10.1177/1541931213571433

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. doi:10.1016/j.promfg.2015.07.847

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.

Goncalves, J., Happee, R., & Bengler, K. (2016). Drowsiness in conditional automation: proneness, diagnosis and driving performance effects. *Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) 2016*. Rio de Janeiro: Brazil. doi: 10.1109/ITSC.2016.7795658

Green, M. (2000). “How long does it take to stop?” Methodological analysis of driver perception-brake times. *Transportation Human Factors, 2(3)*, 195-216.

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3(1)*, 42-56.

Happee, R., Gold, C., Radlmayer, J., Hergeth, S., & Bengler, K. (2017). Take-over performance in evasive manoeuvres. *Accident Analysis and Prevention, 106*, 211-222.

Hoedemaeker, M., & Brookhuis, K. A. (1998). Behavioural adaptation to driving with an adaptive cruise control (ACC). *Transportation Research Part F: Traffic Psychology and Behaviour, 1*, 95-106.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57(3)*, 407-434.

Hoffmann, J., & Gayko, J. (2009). Fahrerwarnelemente [Driver warning signals]. In H. Winner, S. Hakuli, & G. Wolf (Eds.), *Handbuch Fahrerassistenzsysteme. Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (pp. 343-354). Wiesbaden, Germany: Vieweg+Teubner.

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.). *A Cognitive Approach to Situation Awareness* (pp. 155-175). Aldershot, UK: Ashgate.

Jamson, A.H., Merat, N., Carsten, O. M. J., & Lai, F. C. H. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies, 3*0, 116-125.

Jensen, A. F. (2014). *Assessing the impact of direct experience on individual preferences and attitudes for electric vehicles* (Doctoral dissertation, Department of Transport, Technical University of Denmark). Retrieved from http://orbit.dtu.dk/files/120656763/Phd_project_Anders_Fjend-bo_Jensen_m_omslag.pdf

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4(1)*, 53–71.

Josten, J., Schmidt, T., Philipsen, R., Eckstein, L., & Ziefle, M. (2017). Privacy and initial information in automated driving – Evaluation of information demands and data sharing concerns. In *Proceedings of the 28th IEEE Intelligent Vehicles Symposium (IV 2017)*, 541-546. doi:10.1109/IV-S.2017.7995774

Josten, J., Schmidt, T., Philipsen, R., Eckstein, L., & Ziefle, M. (2018). What to expect of automated driving – Expectations and anticipation of system behavior. In N. A. Stanton (Ed.), *Advances in Human Aspects of Transportation*, 606-617. doi:10.1007/978-3-319-60441-1_59

Josten, J., Zlocki, A., & Eckstein, L. (2016). Untersuchung der Bewältigungsleistung des Fahrers von kurzfristig auftretenden Wiederübernahmesituationen nach teilautomatischem, freihändigem Fahren [Investigation of the driver's handling of short-term takeover situations after partially automated, hands-free driving]. *FAT-Schriftenreihe, 289.* Retrieved from https://www.vda.de/de/services/Publikationen/fat-schriftenreihe-289.html

Kaptein, N. A., Theeuwes, J., & van der Horst, R. (1996). Driving Simulator Validity: Some Considerations. *Transportation Research Record, 1550*, 30-36.

Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinitaet erfassen – der Fragebogen TA-EG [Measuring technical affinity – the questionnaire TA-EG]. *8. Berliner Werkstatt Mensch-Maschine-Systeme (ZMMS Spektrum), 22(29)*, 196–201.

Kerschbaum, P., Lorenz, L., & Bengler, K. (2014). Highly automated driving with a decoupled steering wheel. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting*, 1686-1690. doi: 10.1177/1541931214581352

Knappe, G., Keinath, A., & Meinecke, C. (2006). Empfehlungen für die Bestimmung der Spurhaltegüte im Kontext der Fahrsimulation [Recommendations for the assessment of lane keeping quality in the context of driving simulation]. *MMI-Interaktiv, 11*, 3-13.

Körber, M., & Bengler, K. (2014, September). *Potential individual differences regarding automation effects in automated driving*. Interacción 2014, Puerto de la Cruz, Tenerife, Spain.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing, 3*, 2403-2409. doi:10.1016/j.promfg.2015.07.499

Körber, M., Gold, C., Lechner, D., & Bengler, K. (2016). The influence of age on the take-over of vehicle control in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 39*, 19-32.

Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour, 32*, 127-140.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46(1)*, 50-80.

Lorenz, L., & Hergeth, S. (2015). Einfluss der Nebenaufgabe auf die Überwachungsleistung beim teilautomatisierten Fahren [Influence of the secondary task on monitoring performance during partial automated driving]. *VDI-Berichte, 2264*, 159-171.

Louw, T., Kountouriotis, G., Carsten, O., & Merat, N. (2015). Driver inattention during vehicle automation: How does driver engagement affect resumption of control? *4th International Driver Distraction and Inattention Conference*. Sydney: Australia.

Lu, Z., Coster, X., & de Winter, J. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied Ergonomics, 60*, 293-304.

Manzey, D. (2008). *Systemgestaltung und Automatisierung* [System design and automation]. In P. Badke-Schaub, G. Hofinger, & K. Lauche (Eds.). *Human Factors. Psychologie sicheren Handelns in Risikobranchen* (pp. 307-324). doi:10.1007/978-3-540-72321-9.

Matthews, G., Saxby, D. J., Funke, G. J., Emo, A. K., & Desmond, P. A. (2011). Driving states of fatigue or stress. In D. L. Fisher, M. Rizzo, J. K. Caird, & J. D. Lee (Eds.). *Handbook of driving simulation for engineering, medicine, and psychology* (pp. 29-1–29-10). Boca Raton: CRC Press.

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. M. J. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors, 54(5)*, 762-771.

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behavior when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 274-282.

Metz, B., & Landau, A. (2015). Ist der Fahrer aufmerksam? Vorstellung eines Modells zur Beschreibung und Bewertung des Blickverhaltens des Fahrers [Is the driver attentive? Presentation of a model for the description and rating of drivers' gaze behaviour]. *VDI Berichte, 2264*, 187-197.

Mok, B. K.-J., Johns, M., Lee, K. J., Ive, H. P., Miller, D., & Ju, W. (2015). Timing of unstructured transitions of control in automated driving. *2015 IEEE Intelligent Vehicles Symposium (28.06.-01.07.)*, 1167-1172. Seoul, Korea. doi:10.1109/IVS.2015.7225841

Morando, A., Victor, T., & Dozza, M. (2019). A reference model for driver attention in automation: Glance behavior changes during lateral and longitudinal assistance. *IEEE Transactions on Intelligent Transportation Systems, 20, 2999-3009*. doi:10.1109/TITS.2018.2870909

Morgan, P. L., Alford, C., Williams, C., Parkhurst, G., & Pipe, T. (2018). Manual takeover and handover of a simulated fully autonomous vehicle within urban and extra-urban settings. In N. Stanton (Ed.), *Advances in Human Aspects of Transportation*, 760-771. doi:10.1007/978-3-319-60441-1_73

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics, 39(3)*, 429-460.

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? *Transportation Research Part F: Traffic Psychology and Behaviour, 35(11)*, 185-198.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors, 54(5)*, 734-746.

Neubauer, C., Matthews, G., & Saxby, D. (2012). The effect of cell phone use and automation on driver performance and subjective state in simulated driving. *Proceedings of the Human Factors and Ergonomics Society, 56th Annual Meeting*, 1987-1991. doi:10.1177/1071181312561415

Neubauer, C., Matthews, G., & Saxby, D. (2014). Fatigue in the automated vehicle: Do games and conversation distract or energize the driver? *Proceedings of the Human Factors and Ergonomics Society, 58th Annual Meeting*, 2053-2057. doi:10.1177/1541931214581432

Neukum, A., & Krüger, H.-P. (2003). Fahrerreaktionen bei Lenksystemstörungen - Untersuchungsmethoden und Bewertungskriterien [Driver reactions to steering system failures – Methodology and criteria for evaluation]. *VDI-Berichte, 1791*, 297-318.

Olson, P. L. (1989). Driver perception response time. *SAE Transactions, Section 6: Journal of Passenger Cars, 98*, 851-861.

Othersen, I. (2016). *Vom Fahrer zum Denker und Teilzeitlenker. Einflussfaktoren und Gestaltungsmerkmale nutzerorientierter Interaktionskonzepte für die Überwachungsaufgabe des Fahrers im teilautomatisierten Modus* [From driver to thinker to part-time controller. Influencing factors and design characteristics of human-centered interaction concepts for the monitoring task of the driver in the partial automated mode]. (Doctoral dissertation, Technische Universität Braunschweig, Germany). doi:10.1007/978-3-658-15087-7

Othersen, I., Petermann-Stock, I., & Vollrath, M. (2014). Bitte überwachen! – Eine Analyse des teilautomatisierten Fahrens [Please supervise! – An analysis of partial automated driving]. *VDI-Berichte, 2223*, 229–248.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, misuse, abuse. *Human Factors, 39(2)*, 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30(3)*, 286-297.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making, 2(2)*, 140-160.

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 252-263.

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving: Impact of trust and practice on manual control recovery. *Human Factors, 58(2)*, 229-241.

Petermann-Stock, I., Hackenberg, L., Muhr, T., Josten, J., & Eckstein, L. (2015). „Bitte übernehmen Sie das Fahren!“ Ein multimodaler Vergleich von Übernahmestrategien [‚Please take-over the driving task!‘ A multi-modal comparison of takeover strategies]. In Intelligente Transport- und Verkehrssysteme und -dienste Niedersachsen e.V. (Ed.), *AAET: Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel. Beiträge zum 16. Braunschweiger Symposium* (pp. 345-369). Braunschweig: ITS Niedersachen e.V.

Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. *Human Factors, 47(4)*, 840-852.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situation and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, 58(1)*, 2063–2067. doi:10.1177/1541931214581434

Reinisch, P. (2012). *Eine risikoadaptive Eingriffsstrategie für Gefahrenbremssysteme* [A risk adaptive intervention strategy for emergency braking systems]. (Doctoral dissertation, Fakultät für Ingenieurwissenschaften, Universität Duisburg-Essen, Germany). Retrieved from https://duepublico2.uni-due.de/receive/duepublico_mods_00029942

Rudin-Brown, C. M., Parker, H. A., & Malisia, A. R. (2003). Behavioral adaptation to adaptive cruise control. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, 1850-1854. doi:10.1177/154193120304701604

Saad, F. (2007). Dealing with behavioural adaptations to advanced driver support systems. In C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments. Critical issues in driver interactions with intelligent transport systems* (pp. 147-161). Heidelberg, Germany: Springer.

SAE (2018). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles* (Ground Vehicle Standard J3016_201806). Retrieved from https://saemobilus.sae.org/content/j3016_201806.

Schaller, T., Schiehlen, J., & Gradenegger, B. (2008). Stauassistenz - Unterstützung des Fahrers in der Quer- und Längsführung: Systementwicklung und Kundenakzeptanz [Traffic jam assistance – Driver assistance in lateral and longitudinal guidance: System development and customer acceptance]. *3. Tagung Aktive Sicherheit durch Fahrerassistenz*. München, Germany. Retrieved from https://mediatum.ub.tum.de/doc/1145106/1145106.pdf

Schlick, C., Bruder, R., & Luczak, H. (2010). *Arbeitswissenschaft* [Occupational Science] (3rd ed.). Heidelberg, Germany: Springer.

Schmidt, E. A., Kincses, W. E., Schrauf, M., Haufe, S., Schubert, R., & Curio, C. (2007). Assessing drivers' vigilance state during monotonous driving. *Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 138-145.

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human Factors, 50(2)*, 264–275.

Seppelt, B. D., Seaman, S., Lee, J., Angell, L. S., Mehler, B., & Reimer, B. (2017). Glass half-full: On-road glance metrics differentiate crashes from nearcrashes in the 100-car data. *Accident Analysis and Prevention, 107*, 48–62.

Seppelt, B. D., & Victor, T. W. (2016). Potential solutions to human factors challenges in road vehicle automation. In G. Meyer & S. Beiker (Eds.), *Road Vehicle Automation 3, Lecture Notes in Mobility* (pp. 131–148). doi:10.1007/978-3-319-40503-2_11.

Skottke, E.-M., Debus, G., Wang, L., & Huestegge, L. (2014). Carryover effects of highly automated convoy driving on subsequent manual driving performance. *Human Factors, 56(7)*, 1272-1283.

Stanton, N. A., & Marsden, P. (1996). From fly-by-wire to drive-by-wire: Safety implications of automation in vehicles. *Safety Science, 24(1)*, 35-49.

Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driver performance. *Ergonomics, 41(7)*, 1014-1028.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics, 48(10)*, 1294-1313.

Strand, N., Nilsson, J., Karlsson, I. C. M., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 218-228.

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention, 35*, 381-391.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*, 629-646.

UNECE (2017a). *Consolidated Resolution on the Construction of Vehicles (R.E.3) – Revision 6*. Retrieved from www.unece.org/trans/main/wp29/wp29wgs/wp-29gen/wp29resolutions.html

UNECE (2017b). *UN Regulation No. 79 (Addendum 78) - Revision 3*. *Uniform provisions concerning the approval of vehicles with regard to steering equipment.* Retrieved from https://www.unece.org/fileadmin/DAM/tra-ns/main/wp29/wp29regs/2017/R079r3e.pdf

Van der Laan, J. D., Heino, A., & de Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies, 5(1)*, 1-10.

Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and user behavior. *MIS Quarterly, 24(1)*, 115-139.

Verster, J. C., & Roth, T. (2011). Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP). *International Journal of General Medicine, 4*, 359-371.

Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005). Sensitivity to eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*, 167-190.

Victor, T., Tivesten, E., Gustavsson, P., Johannson, J., Sangberg, F., & Aust, M. L. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human Factors, 60(8)*, 1095-1116.

Vollrath, M., & Krems, J. F. (2011). *Verkehrspsychologie: ein Lehrbuch für Psychologen, Ingenieure und Informatiker* [Traffic psychology: A textbook for psychologists, engineers and computer scientists]. Stuttgart, Germany: Kohlhammer.

Vollrath, M., Schleicher, S., & Gelau, C. (2011). The influence of Cruise Control and Adaptive Cruise Control on driving behaviour - A driving simulator study. *Accident Analysis and Prevention, 43*, 1134-1139.

Volvo Car Corporation (2019). *XC60 Betriebsanleitung* [XC60 owners' manual]. Retrieved from https://az685612.vo.msecnd.net/pdfs/24e5fcf7173ec8e-50751d5d40148269b1e82a077/XC60_OwnersManual_MY19_de-DE_TP-27492.pdf

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour, 26*, 227-237.

Ward, N. J. (2000). Automation of task processes: An example of intelligent transportation systems. *Human Factors and Ergonomics in Manufacturing, 10(4)*, 395-408.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, 50(3), 433-441.

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. *Journal of Safety Research, 33*, 117-128.

Weinberger, M., Winner, H., & Bubb, H. (2001). Adaptive cruise control field operational test – the learning phase. *JSAE Review, 22*, 487-494.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3(2)*, 159-177.

Wickens, C. D. (2008a). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors, 50(3)*, 397-403.

Wickens, C. D. (2008b). Multiple resources and mental workload. *Human Factors, 50(3)*, 449-455.

Wright, T. J., Samuel, S., Borowsky, A., Zilberstein, S., & Fisher, D. L. (2016). Experienced drivers are quicker to achieve situation awareness than inexperienced drivers in situations of transfer of control within a Level 3 autonomous environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*, 270-273. doi:10.1177/15419312 13601062

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. *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems (ITSC 2013), 2012-2019*. doi:10.1109/ITSC.2013.6728525

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 and Prevention, 78*, 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, 92*, 230-239.

12 Appendix

12.1 System Instruction for Hands-Off Group in Survey (German)

Versetzen Sie sich für die Beantwortung der folgenden Fragen bitte in die unten beschriebene Situation. Lesen Sie die folgenden Informationen, vor allem auch die Details im hervorgehobenen Abschnitt, bitte sorgfältig durch.

Sie haben einen Neuwagen gemietet, der mit einem System zur automatisierten Autobahnfahrt ausgestattet ist. Da die Funktion noch nicht sehr weit verbreitet ist, hat die Fahrzeugvermietung folgende Informationen zur Funktion aus dem Fahrzeughandbuch zusammengestellt:

- Das System übernimmt nach Aktivierung durch den Fahrer die Lenkung zur Spurhaltung, die Geschwindigkeitsregelung und die Regelung des Abstands zum Vorderfahrzeug.

- Der Fahrer muss auch bei aktiviertem System jederzeit die Umgebung beobachten und sofort eingreifen können, sollte das System seine technischen Grenzen erreichen.

- Wenn das System aktiv ist, dürfen beide Hände vom Lenkrad genommen werden.

Im Folgenden werden Ihnen einige Aussagen zum gerade beschriebenen System präsentiert.

Wählen Sie bitte die Option aus, die Ihre Meinung am besten widerspiegelt.

12.2 System Instruction for Hands-On Group in Survey (German)

Versetzen Sie sich für die Beantwortung der folgenden Fragen bitte in die unten beschriebene Situation. Lesen Sie die folgenden Informationen, vor allem auch die Details im hervorgehobenen Abschnitt, bitte sorgfältig durch.

Sie haben einen Neuwagen gemietet, der mit einem System zur automatisierten Autobahnfahrt ausgestattet ist. Da die Funktion noch nicht sehr weit verbreitet ist, hat die Fahrzeugvermietung folgende Informationen zur Funktion aus dem Fahrzeughandbuch zusammengestellt:

- Das System übernimmt nach Aktivierung durch den Fahrer die Lenkung zur Spurhaltung, die Geschwindigkeitsregelung und die Regelung des Abstands zum Vorderfahrzeug.

- Der Fahrer muss auch bei aktiviertem System jederzeit die Umgebung beobachten und sofort eingreifen können, sollte das System seine technischen Grenzen erreichen.

- Auch wenn das System aktiv ist, sind Sie dazu verpflichtet immer mindestens eine Hand am Lenkrad zu behalten.

Im Folgenden werden Ihnen einige Aussagen zum gerade beschriebenen System präsentiert.

Wählen Sie bitte die Option aus, die Ihre Meinung am besten widerspiegelt.

12.3 System Instruction for Short-Termed Takeover Situations (German)

The following instruction for short-termed takeover situations has prior been published in Josten et al. (2016, p. 106). The instruction was adapted regarding the system limits and speed range for Study 3 (see Josten et al., 2016, p. 107).

„Ich werde Ihnen nun einen beispielhaften Stauassistenten zeigen, der sich so bedienen lässt, wie auch die beiden in den folgenden Versuchsfahrten zu testenden Stauassistenten. Alle Informationen zur Bedienung, die für eines der beiden anderen Systeme spezifisch sind, werde ich Ihnen jeweils direkt vor der entsprechenden Versuchsfahrt erläutern. Wir werden dann gleich während einer kurzen Staufahrt die Bedienung des Assistenten, also das Einschalten und auch Übersteuern, z.B. zum Abfahren von der Autobahn, üben.
Die heute zu testenden Stauassistenten übernehmen in Stausituationen die gesamte Fahraufgabe für Sie, das heißt, der Assistent regelt sowohl die Geschwindigkeit, als auch den Abstand zum Vorderfahrzeug und die Spurhaltung. Während der Stauassistent aktiv ist, dreht sich das Lenkrad bei Lenkbewegungen mit. Genau wie bei der Nutzung von anderen Assistenzsystemen, z.B. Tempomat oder Adaptive Cruise Control, sollten Sie als verantwortlicher Fahrer Systemstatus, Fahrzeugverhalten und Umgebungsverkehr bei aktiviertem System durchgehend beobachten.
Die Systeme funktionieren jeweils in einem Geschwindigkeitsbereich zwischen 0 und 50 km/h. Der aktuelle Systemstatus lässt sich jederzeit im Display hinter dem Lenkrad ablesen.

Ich werde Ihnen die verschiedenen Zustände nun einzeln vorführen. Das aktuelle Symbol (grau, gestrichelte Linien) bedeutet, dass das System aus ist und zurzeit auch nicht eingeschaltet werden kann, weil die Voraussetzungen zur Nutzung nicht erfüllt sind. Erkennt das System eine Stausituation, die sich im zugelassenen Geschwindigkeitsbereich bewegt, wird Ihnen die Aktivierung des Systems angeboten (Symbol „System bereit“ plus Hinweiston). Sollte diese Situation im Versuch eintreten, bitte ich Sie, das System auf den Hinweis hin einzuschalten, denn nur durch eine entsprechend durchgehende Nutzung des Systems können Sie genug Erfahrung sammeln, um uns anschließend eine fundierte Bewertung des Systems zu geben. Sie können das System über die entsprechende Lenkradtaste (unten links) einschalten. Das System-Symbol sollte Ihnen anschließend anzeigen, dass das System aktiv ist. Wir werden das Einschalten in der kommenden Fahrt üben.
Wollen Sie die Autobahn verlassen (*im Versuch nur auf meine Anweisung hin*), können Sie das System durch Gaspedal, Bremse und Lenkbewegungen übersteuern. Achten Sie daher bitte darauf, den Lenkbewegungen des Fahrzeugs nicht entgegen zu arbeiten, da Sie das System sonst unbeabsichtigt ausschalten

könnten. In der folgenden Fahrt lassen Sie bitte die Hände am Lenkrad, solange das System aktiviert ist.
Sollte das System zu irgendeinem Zeitpunkt an seine Systemgrenzen gelangen, z.B., weil die Geschwindigkeit des Umgebungsverkehrs zu hoch wird, schaltet es sich mit akustischem Warnton und entsprechendem Symbol ab. In diesem Fall müssen Sie als Fahrer die Lenkung und Geschwindigkeitsregelung übernehmen."

12.4 System Instruction for Planned Takeover Situations (German)

Der Autobahnassistent ist ein **für Autobahnen und autobahnähnliche Straßen ausgelegtes Assistenzsystem**, das nach Aktivierung die gesamte Fahraufgabe übernimmt. Der Assistent regelt im aktiven Zustand sowohl die **Geschwindigkeit**, als auch den **Abstand** zu einem vorhandenen Vorderfahrzeug und die **Spurhaltung**. Während der Assistent aktiv ist, dreht sich das Lenkrad bei Lenkbewegungen mit.

Die Zielgeschwindigkeit des Autobahnassistenten beträgt 130 km/h. Solange keine Verkehrsschilder oder Vorderfahrzeuge die Geschwindigkeit beschränken, wird der Assistent auf 130 km/h regeln. In Bereichen mit geltender Geschwindigkeits-begrenzung unter 130 km/h regelt der Assistent automatisch auf die auf den Verkehrszeichen dargestellte Höchstgeschwindigkeit. Befindet sich vor Ihnen ein langsameres Fahrzeug auf der Straße, wird der Assistent sich an die Geschwindigkeit dieses Fahrzeuges anpassen und diesem mit einem geschwindigkeitsangepassten, voreingestellten Abstand folgen.

Der Assistent orientiert sich mit Hilfe mehrerer Frontkameras auf der Straße. Das System nutzt zur Orientierung die Fahrbahnmarkierungen und richtet das Fahrzeug spurmittig aus. Über die Kameras werden vorausfahrende Fahrzeuge im eigenen Fahrstreifen zuverlässig erkannt.

WICHTIG:

Genau wie bei der Nutzung von anderen Assistenzsystemen, z.B. Tempomat (CC) oder Adaptive Cruise Control (ACC), müssen Sie als verantwortlicher Fahrzeugführer **Systemstatus, Fahrzeugverhalten und Umgebungsverkehr bei aktiviertem System durchgehend beobachten. Auch bei aktiviertem Assistenten sind Sie als Fahrer für die Fahrzeugführung verantwortlich.**

Beachten Sie bitte folgende wichtige Informationen zur Funktionsweise des Assistenten:

- **Der Assistent erkennt Fahrzeuge vor sich zuverlässig, jedoch keine Fahrzeuge auf der Nebenspur.** Fahren Fahrzeuge auf dem linken Fahrstreifen langsamer als der Assistent, müssen Sie die Fahrzeugführung rechtzeitig übernehmen, um das andere Fahrzeug nicht von rechts zu überholen.
- **Das System beherrscht keine Fahrstreifenwechsel.** Wollen Sie den Fahrstreifen wechseln, müssen Sie das System übersteuern oder abschalten und können es nach Rückkehr auf den rechten Fahrstreifen wieder aktivieren.
- Detektiert der Assistent einen Gegenstand oder ein **Hindernis im eigenen Fahrstreifen**, auf den er nicht in erforderlicher Weise (z.B. Ausweichen) reagieren kann, wird er sich abschalten und Sie gleichzeitig durch einen entsprechenden Hinweis im Display hinter dem Lenkrad (siehe Abbildung 2) benachrichtigen.
- **Verschmutzte, fehlende oder temporäre (z.B. gelbe) Fahrbahnmarkierungen** werden vom Assistent nicht immer korrekt erkannt. Eventuell wird der Assistent die Kontrolle an Sie zurück übergeben, wenn er eine korrekte Ausrichtung des Fahrzeugs auf dem Fahrstreifen nicht sicherstellen kann.

AKTIVIERUNG DES ASSISTENTEN

Um den Assistenten aktivieren zu können, muss die Fahrzeuggeschwindigkeit **über 60 km/h** liegen. Der Assistent kann nur auf dem **rechten Fahrstreifen** aktiviert werden. Der Assistent prüft diese Voraussetzungen und gibt ein Angebot zum Aktivieren aus, wenn diese erfüllt sind. Er kann anschließend über die **Systemtaste 1** aktiviert werden.

Um den Assistenten aktivieren zu können, dürfen Sie zeitgleich mit dem Tastendruck keine eigenen Eingaben mehr vornehmen, das heißt, weder lenken, noch Gas geben oder bremsen. Nur dann lässt sich der Assistent über die Systemtaste 1 aktivieren.

Abbildung 1: Position der Systemtaste 1 (Einschalten des Assistenten) in der Mittelkonsole.

Der aktuelle Status des Assistenten wird durchgehend im Display hinter dem Lenkrad angezeigt. Alle vom Assistenten verwendeten Symbole und ihre Bedeutung finden Sie in Abbildung 2. Das Einschaltangebot und ein Abschalten werden zusätzlich durch Tonsignale kommuniziert.

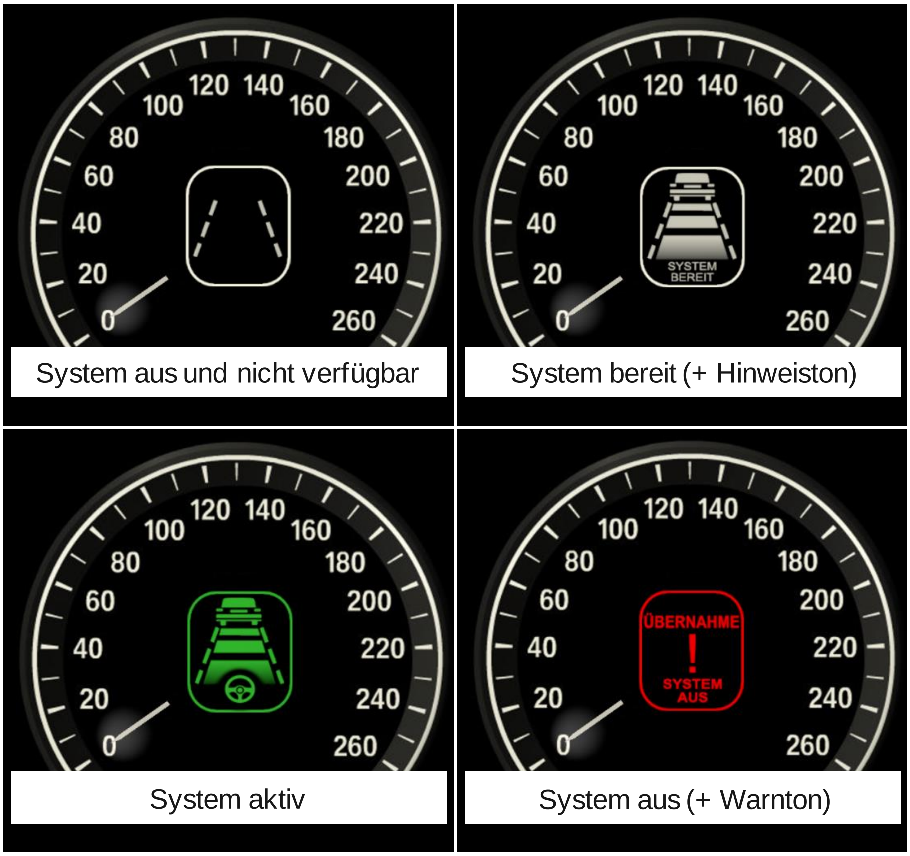

Abbildung 2: Vom Assistenten verwendete Symbole und ihre Bedeutung

WICHTIG:

Erkennt das Sensorset des Fahrzeugs eine Situation, in der der Assistent eine sichere Fahrzeugführung nicht gewährleisten kann, gibt der Assistent ein Übergabesignal aus (rotes Symbol mit Warnton, siehe Abbildung 2). **Mit Ausgabe des Übergabesignals ist der Assistent inaktiv, das heißt, Sie müssen spätestens mit Ausgabe des Übergabesignals die Fahrzeugführung wieder übernehmen.**

Sie können bei aktivem Assistenten jederzeit die **Kontrolle über die Fahrzeugführung übernehmen**, indem Sie lenken, bremsen oder auch beschleunigen. Der Assistent wird durch Ihren Eingriff in die Fahrzeugführung **abgeschaltet**. Zusätzlich lässt sich der Assistent durch erneuten Druck auf die Systemtaste 1 abschalten.

Sind die Anforderungen an die Nutzung des Assistenten nach dem Abschalten erneut erfüllt, gibt der Assistent eine **Einschaltaufforderung** („System bereit“) aus und kann durch Drücken der **Systemtaste 1** erneut aktiviert werden.

12.4.1 Instruction for Hands-On Group

Hinweis zur Bedienung:

Auch bei aktiviertem System muss sich zu jeder Zeit mindestens eine Hand am Lenkrad befinden. Achten Sie bitte darauf, den Lenkbewegungen des aktivierten Assistenten nicht entgegen zu arbeiten, da Sie das System sonst unbeabsichtigt ausschalten könnten. Der Assistent und die Verkehrsumgebung müssen durchgehend überwacht werden.

12.4.2 Instruction for Hands-Off and Manual Group

Hinweis zur Bedienung:

Während der Assistent aktiv ist, können beide Hände vom Lenkrad genommen werden. Achten Sie bitte darauf, den Lenkbewegungen des aktivierten Assistenten nicht entgegen zu arbeiten, da Sie das System sonst unbeabsichtigt ausschalten könnten. Der Assistent und die Verkehrsumgebung müssen durchgehend überwacht werden.

12.5 ADAS Experience Scale

ADAS experience scores were established based on the summed experience with the following systems: Cruise Control (CC), Adaptive Cruise Control (ACC), Traffic Jam Assist (TJA) and Lane Keeping Assist (LKA).

A detailed description of each ADAS was provided to participants of the driving simulator studies and the test track study before rating prior ADAS experience. Experience points per item ranged from 0 (‚Das System ist mir nicht bekannt‘; English: ‘*I am not familiar with this system*’) to 4 (‚Ich nutze das System regelmäßig‘; English: ‘*I use this system regularly*’).

Bitte geben Sie an, welche der folgenden Assistenzsysteme Ihnen bekannt sind, und welche Sie bereits genutzt haben bzw. regelmäßig nutzen. Wählen Sie bitte eine Option pro System aus.

Geschwindigkeitsregelanlage (auch: Tempomat)				
Das System ist mir **nicht** bekannt.	Das System ist mir vom Hören-Sagen oder theoretisch bekannt.	Ich habe das System bereits einmal genutzt.	Ich habe das System bereits mehrfach genutzt.	Ich nutze das System regelmäßig.
☐	☐	☐	☐	☐

12.6 Additional Figures

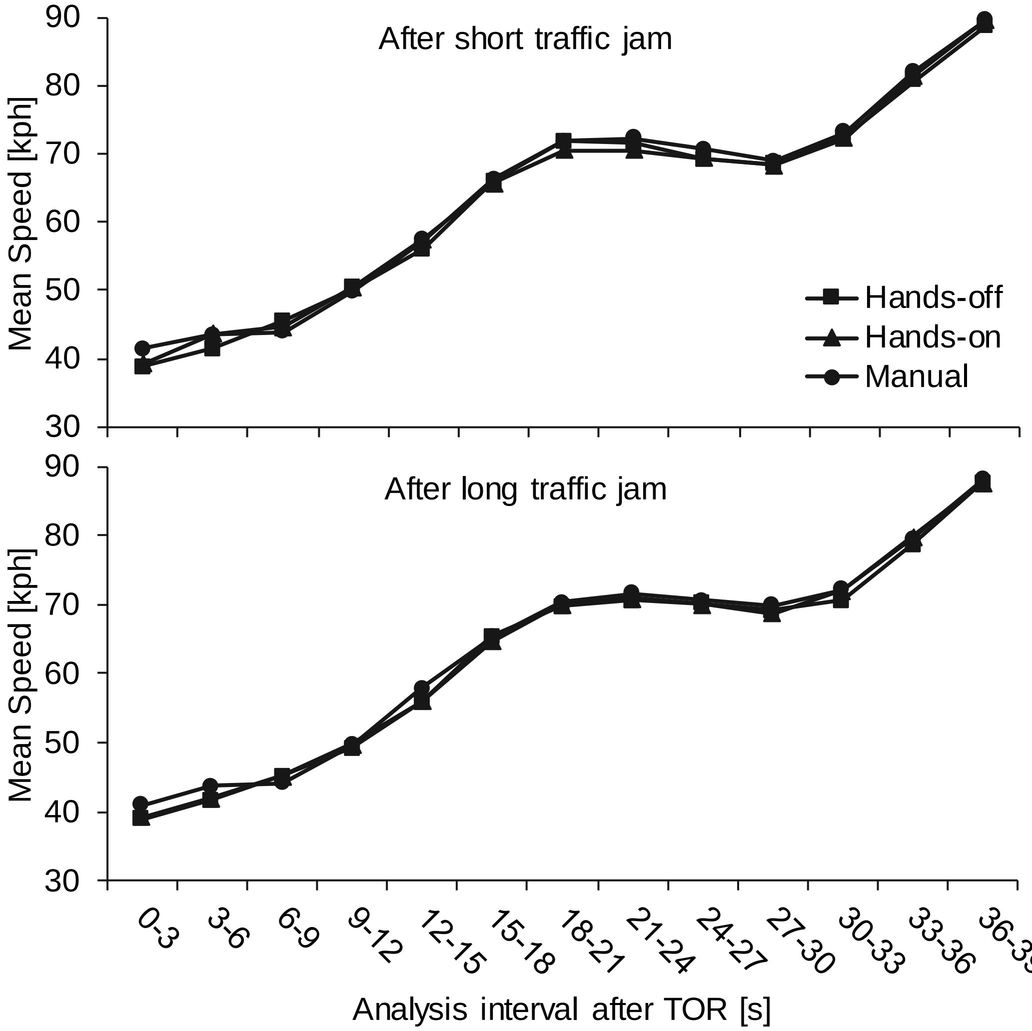

Figure 12-1: Mean speed for the sample of Study 1 (n = 21 for curve scenario) in the 13 intervals used for analysis of driver performance after short (upper graph) and long (lower graph) traffic jam driving with means for each of the three driving conditions.

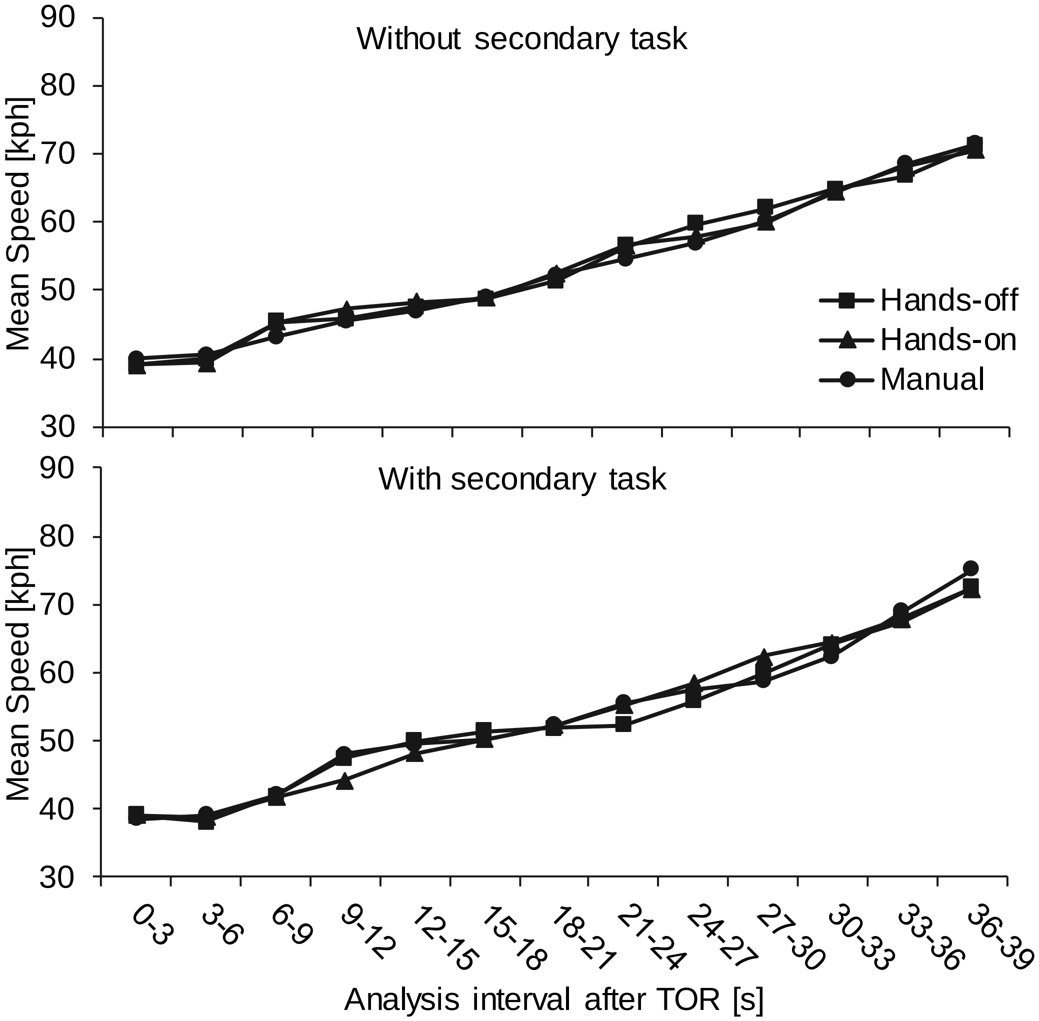

Figure 12-2: Mean speed for the sample of Study 2 (N = 12) in the 13 intervals used for analysis of driver performance without (upper graph) and with (lower graph) secondary task with means for each of the three driving conditions.

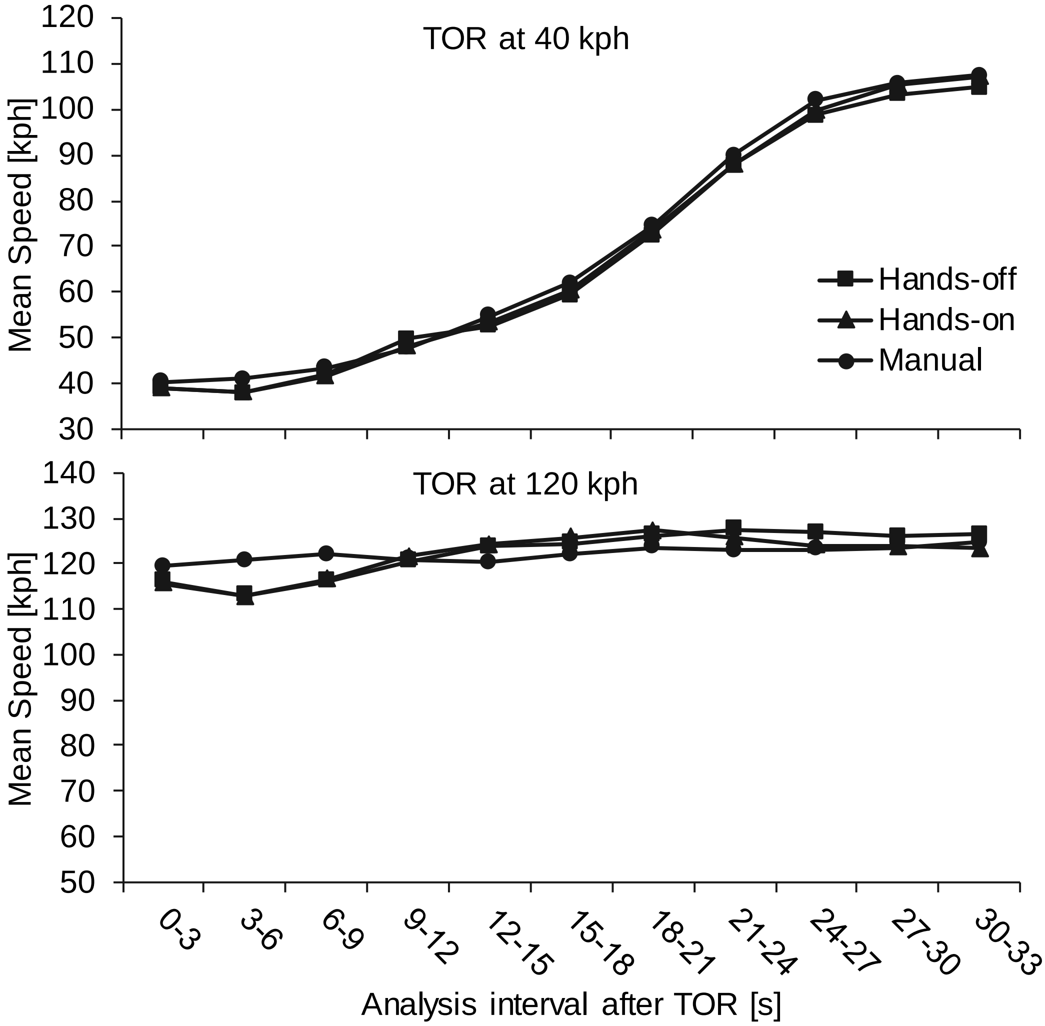

Figure 12-3: Mean speed for the sample of Study 3 in the 11 intervals used for analysis of driver performance at low speed (upper graph; n = 15) and at high speed (lower graph; n = 10) with means for each of the three driving conditions.

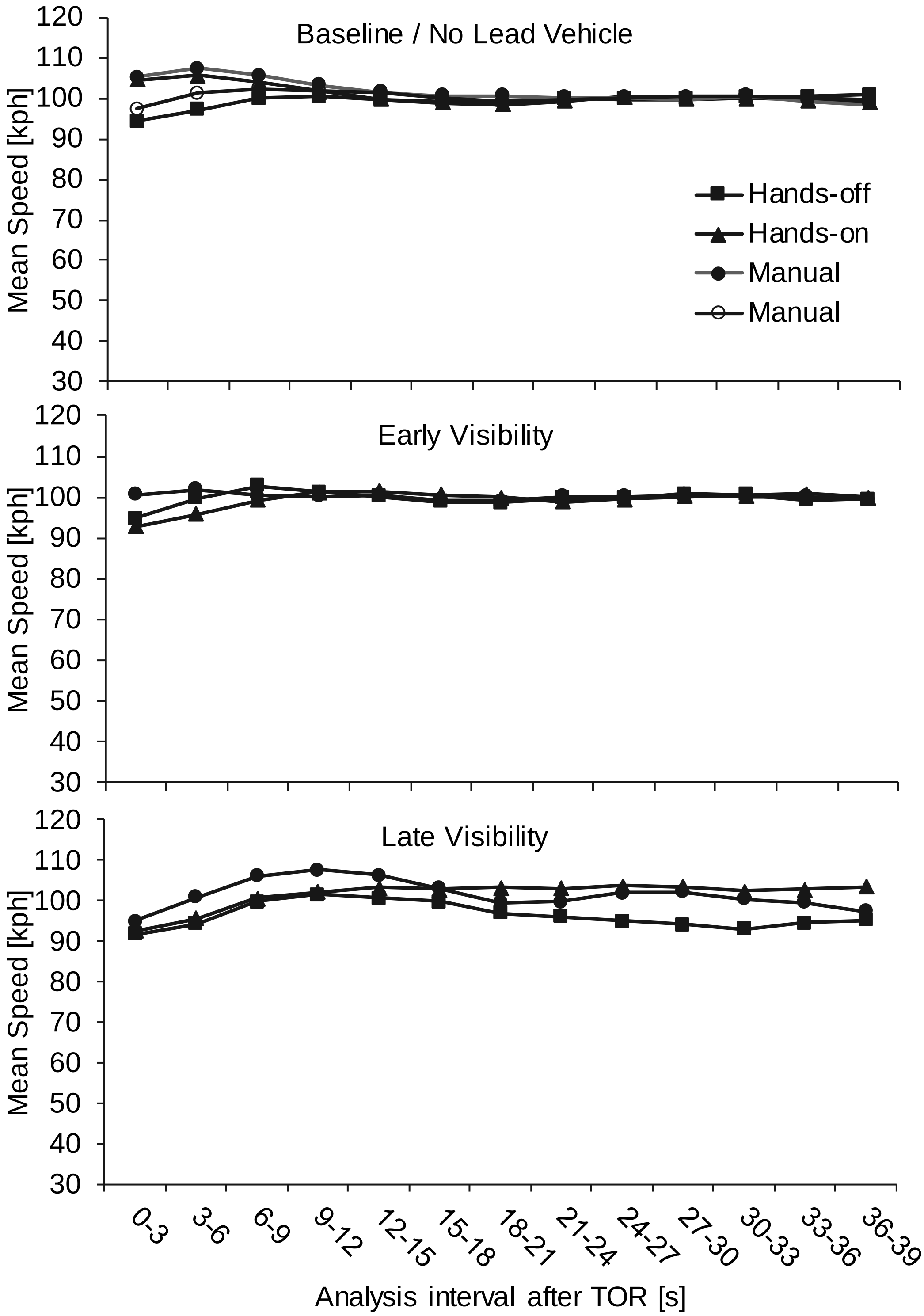

Figure 12-4: Mean speed for the sample of Study 4 in the 13 intervals used for analysis of driver performance without lead vehicle (upper graph; baseline performance; white dots represent the hands-off drive of the manual group), after early visibility (middle graph) and late visibility (lower graph). The first interval starts with the COG entering the new lane. Only valid data sets are considered for the single graphs.

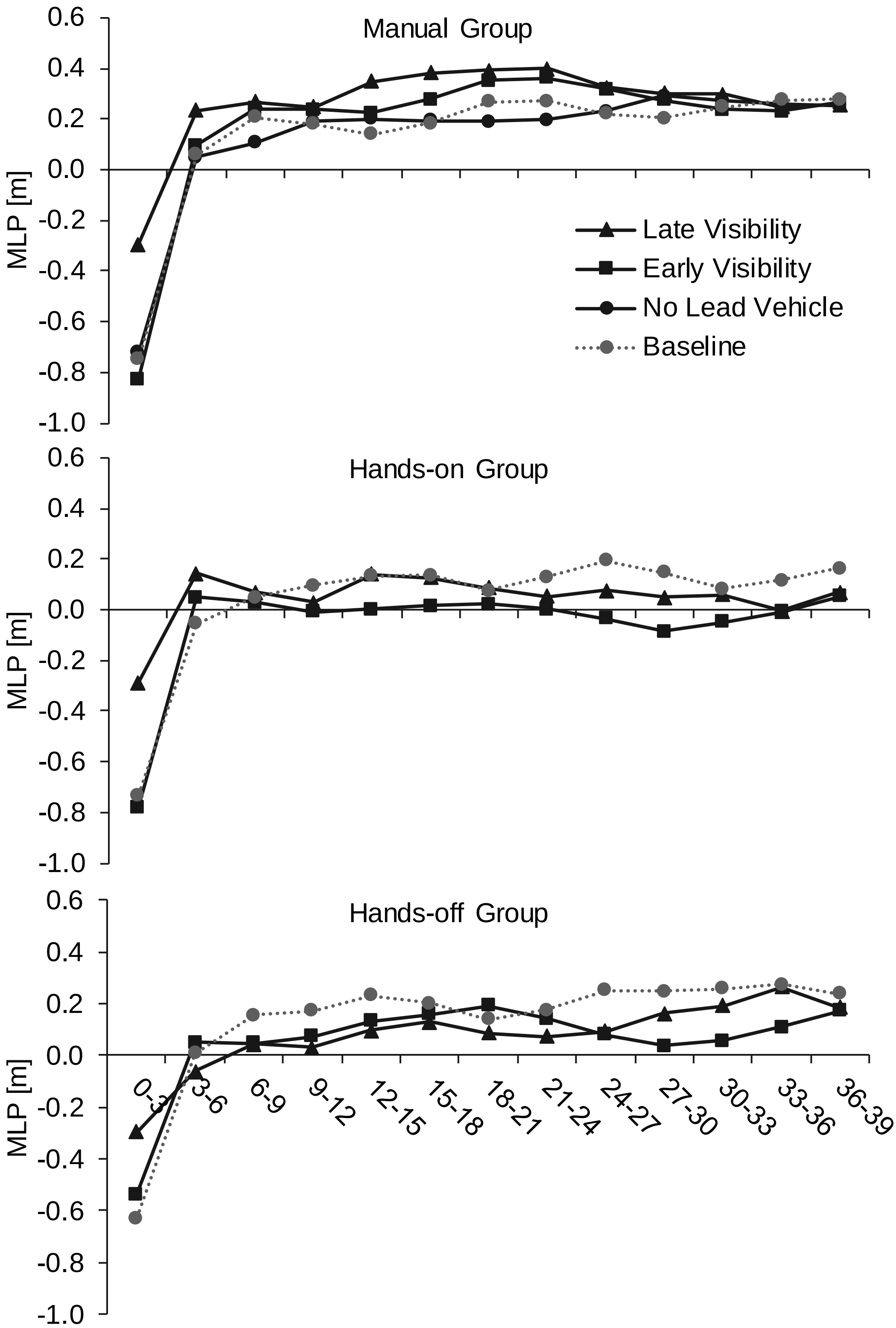

Figure 12-5: MLP in Study 4, aggregated for 3-s-intervals beginning with the COG entering the new lane. Only valid data sets are included for each driver group and condition. Grey dotted lines represent the according manual baseline drive without lead vehicle.

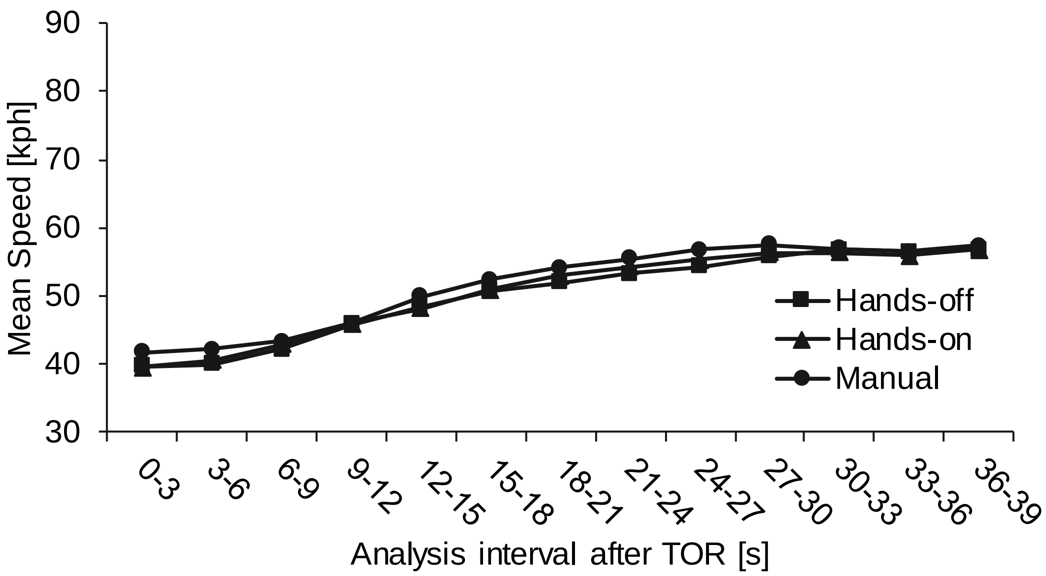

Figure 12-6: Mean speed for the sample of test track study (*N* = 12; Study 5) in the 13 intervals used for analysis of driver performance with means for each of the three driving conditions.

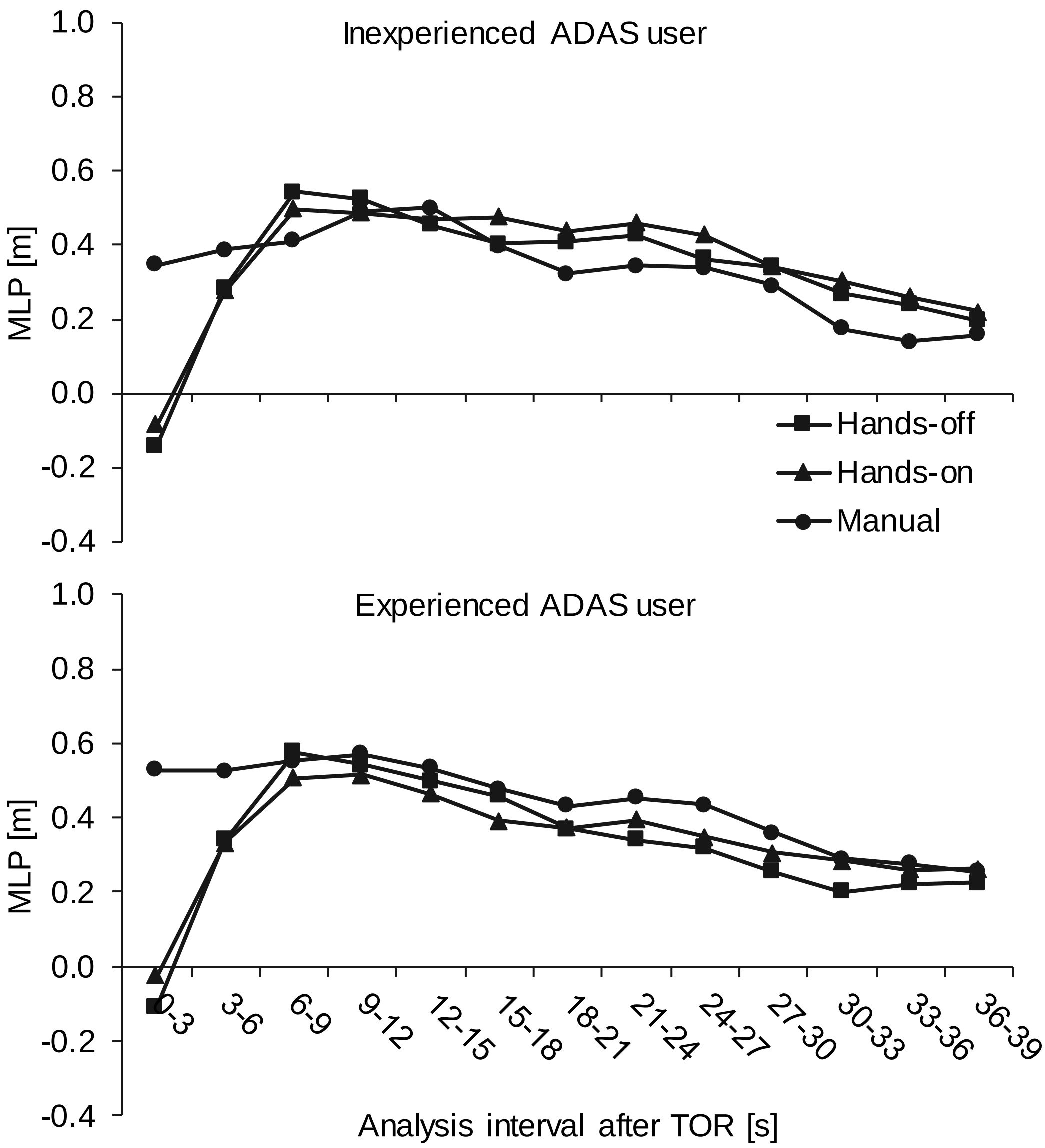

Figure 12-7: MLP for inexperienced (upper graph; sample means for *n* = 22) and experienced (lower graph; sample means for *n* = 26) ADAS user in the reference scenario of the simulator studies.