Technische Universität München
TUM School of Management

Development of an HMI Concept to Communicate the Availability of a Conditional Automated Driving System

Simon Danner, M.Sc.

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung eines

Doktors der Philosophie (Dr. phil.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Alwine Mohnen

Prüfende der Dissertation:

1. Prof. Dr. Klaus Bengler
2. Prof. Dr. Jutta Roosen

Die Dissertation wurde am 24.06.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Management am 25.02.2025 angenommen.

Danksagung

Diese Dissertation entstand, während ich als wissenschaftlicher Mitarbeiter am Lehrstuhl für Ergonomie der Technischen Universität München angestellt war in Kooperation mit der Audi AG. Ich möchte mich herzlich bei allen Personen bedanken, die mir während dieser Zeit zur Seite standen.

Besonderer Dank gilt meinem Doktorvater, Prof. Dr. Klaus Bengler, der mir die Chance bot, an den Lehrstuhl zu kommen und mein Forschungsvorhaben zu verwirklichen. In vielen Gesprächen hat er mir neue Wege und Gedanken aufgezeigt und mir dabei stets die Freiheit gelassen, das Projekt nach meinen Vorstellungen zu gestalten. Auch bei Audi konnte ich mich auf die Hilfe vieler Kolleginnen und Kollegen verlassen, hervorheben möchte ich jedoch die Unterstützung, die ich von Linda Köhler und Matthias Pfromm erfahren habe, sowohl auf fachlicher wie auf persönlicher Ebene.

Die Zeit am Lehrstuhl war eine sehr besondere für mich. Ich habe viel gelernt und mich persönlich weiterentwickelt, das Forschen und tiefe Eintauchen in mein Themengebiet haben mir unglaublich viel Spaß gemacht. Es waren jedoch die Kolleginnen und Kollegen, die dafür sorgten, dass ich mit Freude zur Arbeit ging und es mir auch manchmal schwerfiel, wieder den Nachhauseweg anzutreten. Ewiger Dank gilt insbesondere Alexander Feierle, ohne dessen wissenschaftliche, technische aber vor allem emotionale Unterstützung ich die ein oder andere Phase nicht so gut überstanden hätte.

Außerdem möchte ich mich bei meiner Familie bedanken, die mich während meines Studiums immer unterstützt und es mir somit überhaupt erst ermöglicht hat, die Stelle am Lehrstuhl anzutreten. Zu guter Letzt bedanke ich mich bei Emma, die mir in all den Jahren eine geduldige Zuhörerin war. Du hast mich von Anfang an bestärkt diese Doktorarbeit zu beginnen und mich über die ganze Zeit hinweg motiviert, am Ball zu bleiben. Ohne dich hätte ich das nicht geschafft.

Abstract

Higher level automated driving, such as SAE Levels 3 and 4, is predicted to come with several benefits like the possibility of conducting non-driving related activities (NDRAs) without monitoring the automation system. As this is seen as one of the most important advantages of automated driving, users' perception of the usefulness of such systems may – at least partially – depend on how well they are able to perform NDRAs, which in turn might depend on how long the vehicle will be able to drive automatically without the driver taking over control from the vehicle again. Since the availability of the automated driving system (ADSs) requires specific conditions to be met that might not be in users' direct control, how long the system is available must mostly be accepted as a given. Nonetheless, the availability of the ADS can be indicated by means of the human-machine interface (HMI), enabling users to plan their activities accordingly and to adjust their expectations, avoiding disappointment. The research on HMIs regarding ADSs has long focused on intervention requests by the system and information needs while driving automatically. This work aims to change that by providing research on requirements for HMIs regarding the availability of ADSs independent of the system state and requests to intervene. First, information needs were collected using explorative research approaches. These information needs were then translated into HMI concepts and tested regarding their benefits for users in driving simulator studies. Information displays that were found to be useful for study participants were combined into one final HMI concept providing information on why the ADS is not available and when it will become available. Furthermore, before the ADS is activated, the HMI concept displays how long it will be available after activation, giving users the possibility to evaluate whether they want to activate the system and whether the availability duration fits their needs. These information elements were shown to increase acceptance towards the ADS and usability of the HMI display. Moreover, study results indicate that task performance in NDRAs and perceived system understanding can be increased by providing transparency regarding the availability of the system. Hence, this thesis represents an important contribution to the literature on human-machine interaction related to automated driving, as it fills a gap that was previously overlooked in research and helps developers to create HMIs that improve users' acceptance, which is associated with the actual usage of a system.

Zusammenfassung

Automatisiertes Fahren höherer Stufen, wie SAE Level 3 und 4, wird mit mehreren Vorteilen verbunden sein. Zum Beispiel mit der Möglichkeit für Nutzende, Aktivitäten auszuüben, die nicht mit dem Fahren zusammenhängen, ohne das Automatisierungssystem überwachen zu müssen. Da dies als einer der wichtigsten Vorteile des automatisierten Fahrens angesehen wird, könnte der wahrgenommene Nutzen solcher Systeme zumindest teilweise davon abhängen, wie gut Nutzende in der Lage sein werden, fahrfremde Tätigkeiten (NDRAs) auszuüben, was wiederum davon abhängen könnte, wie lange automatisiertes Fahren möglich sein wird, ohne wieder die Kontrolle über das Fahrzeug übernehmen zu müssen. Da die Verfügbarkeit des automatisierten Fahrsystems (ADS) an bestimmte Bedingungen geknüpft sein wird, die die Nutzer häufig nicht beeinflusst können, muss die Dauer der Verfügbarkeit dieses Systems meist als gegeben hingenommen werden. Dennoch kann durch die Mensch-Maschine-Schnittstelle (HMI) Transparenz bezüglich der Verfügbarkeit des ADS geschaffen werden, damit die Nutzer ihre Aktivitäten entsprechend planen und ihre Erwartungen anpassen können, um Enttäuschungen zu vermeiden. In der Forschung zu HMIs im Zusammenhang mit ADSs standen Übernahmeaufforderungen und Informationsbedarfe während des automatisierten Fahrens lange Zeit im Mittelpunkt. Die vorliegende Arbeit zielt darauf ab, dies zu ändern und erforscht die Anforderungen an ein HMI bezüglich der Verfügbarkeit von ADSs unabhängig vom Systemzustand und den Übernahmeaufforderungen. Zunächst wurden Informationsbedarfe mit Hilfe explorativer Forschungsmethoden gesammelt. Diese Bedarfe wurden dann in HMI-Konzepte übersetzt und in Fahrsimulatorstudien auf ihre Vorteile für Nutzende getestet. Informationen, die sich als nützlich erwiesen, wurden in einem endgültigen HMI-Konzept zusammengefasst, das Auskunft darüber gibt, warum das ADS nicht verfügbar ist und wann es verfügbar sein wird. Außerdem wird vor der Aktivierung des ADS angezeigt, wie lange es nach der Aktivierung zur Verfügung stehen wird, um den Nutzenden die Möglichkeit zu geben, selbst zu beurteilen, ob sie das System aktivieren wollen und ob die Dauer der Verfügbarkeit ihren Bedürfnissen entspricht. Es hat sich gezeigt, dass die Anzeige dieser Informationen die Akzeptanz bezüglich des ADS und die Benutzerfreundlichkeit der HMI-Anzeige erhöhen. Darüber hinaus deuten die Studienergebnisse darauf hin, dass die Leistung in der Ausführung von NDRAs und das wahrgenommene Systemverständnis durch die Bereitstellung von Transparenz bezüglich der Verfügbarkeit des Systems erhöht werden können. Diese Arbeit stellt somit einen wichtigen Beitrag zur Literatur über die Mensch-

Maschine-Interaktion im Zusammenhang mit dem automatisierten Fahren dar, da sie eine Lücke schließt, die bisher in der Forschung übersehen wurde und Entwicklern helfen kann, HMIs zu erstellen, die die Akzeptanz der Nutzenden verbessern, welche wiederum mit der tatsächlichen Nutzung eines Systems assoziiert ist.

Contents

Acronyms

ACC:	Adaptive Cruise Control
ADAS:	Advanced Driving Assistance Systems
ADS:	Automated Driving System
AV:	Automated Vehicle
CAD:	Conditionally Automated Driving
CID:	Central Information Display
HMI:	Human-Machine Interface
IC:	Instrument Cluster
ITU:	Intention to Use
NDRA:	Non-Driving Related Activity
ODD:	Operational Design Domain
PEoU:	Perceived Ease of Use
PU:	Perceived Usefulness
RtI:	Request to Intervene
SAE:	Society of Automotive Eningeers
TAM:	Technology Acceptance Model
TBP:	Theory of Planned Behaviour
UTAUT:	Unified Theory of Acceptance and Use of Technology
UX:	User Experience

1 Introduction

Automated driving at higher levels is associated with several benefits, including the possibility to engage in activities that have no connection with the actual task of driving (Pfleging et al., 2016) or increased road safety (Winkle, 2016). Before benefits like these can come to fruition, users will have to understand automated driving systems to be able to operate them safely and satisfactorily. Looking at existing advanced driving assistance systems (ADAS) like the adaptive cruise control (ACC), it seems like users often struggle to understand the systems' functionalities or do not know about their limitations (Beggiato & Krems, 2013). As it can be assumed that higher levels of automated driving systems (ADSs) are more complex than an ACC for example (Parasuraman & Riley, 1997), enabling the users to understand and use them correctly becomes even more important. The limitations of an ADS influence the users' possibilities to engage in non-driving related activities (NDRAs), and thus can impact the perceived usefulness of the system as well as the usage behaviour. As will be further explained in the next chapter, limited availability of the system must be communicated clearly to help users adjust their expectations and plan their NDRAs accordingly, ensuring that the benefits of an ADS will actually be perceived as such.

An overview of the state of research regarding expectations and attitudes towards automated driving along with methodological approaches to study the same is given in Chapter 2, which is the theoretical background chapter of this work. Chapter 3 gives an overview of the methodology and research questions of this thesis, while Chapter 4 presents an explorative investigation of which information users might need to develop an appropriate mental model of the availability and limitations of ADSs of higher levels. The development of different HMI concepts, which are based on the found information requirements, and the testing of the same in driving simulator studies is described in Chapter 5. The findings of the experiments result in a final HMI concept which is presented in Chapter 6. A general discussion of the outcome of this work is presented in Chapter 7 and recommendations on how to conceptualise an in-vehicle HMI display with a focus on conveying important information on the system's functionalities and limitations in different system states to increase transparency are given in Chapter 8. Moreover, methodological insights gained from investigating different information needs regarding ADSs are presented in the respective chapters of the different studies as well as in the general discussion.

2 Theoretical Background

2.1 Manual and Automated Driving

The term automated driving does not refer to a specific function or definition. Rather, it is used to distinguish certain driver-assistance functions from manual driving. Manual driving can be divided into three different tasks: the primary, the secondary, and the tertiary driving task. The primary driving task involves the mechanical activities that are directly involved in driving a vehicle, such as steering and accelerating. The secondary driving task includes the activities that are performed to communicate with the environment and aim at supporting the primary driving task, such as operating the windshield wipers when it rains. The tertiary driving task includes, for example, operating the air conditioning or the radio and is therefore not directly related to operating the vehicle (Bubb, 2003).

The primary driving task can be further subdivided into three levels, the navigation, the manoeuvre, and the stabilisation level (Donges, 1982). The navigation level comprises the selection and planning of a route as well as the calculation of the time required for the route. The manoeuvre level is where actions such as changing lanes, turning or overtaking take place. All interventions that serve to stabilise the vehicle concerning the current lane and speed are associated with the stabilisation level (Donges, 2015).

ADASs primarily serve to increase efficiency and safety when driving (Bengler et al., 2014). However, the development of such systems also aims at increasing comfort while driving (Bubb & Bengler, 2015; Wallentowitz & Reif, 2011). ADASs can support the driver in each of the three driving tasks, for example in navigation, lane keeping or even in the tertiary driving task, such as by providing information on service activities to be performed or by assisting in the operation of the infotainment system (Bubb & Bengler, 2015). These support systems can be divided into three categories (Gasser et al., 2015): category A describes information and warning systems that do not have a direct influence on vehicle guidance, but only have an effect through the driver. One example of such a system is the traffic sign recognition. Category B includes systems that continuously automate specific driving tasks. This means that these functions have a direct influence on vehicle guidance, although the driver is always able to override them, as they are mainly comfort functions. The ACC is an ADAS that falls in this category. Category C, on the other hand, describes systems that intervene in vehicle control in emergencies. This means that there is a direct influence on the control of the vehicle, as in category B, but this influence is not continuous and is limited to hazardous and emergency situations. One example of this is the emergency brake assistant.

The present study takes on the definition and categorisation of automated driving systems as proposed by the Society of Automotive Engineers (SAE) (SAE, 2021). The SAE taxonomy describes all levels of automation from Level 0 to Level 5. Level 0 is manual driving, where all driving tasks must be performed by the driver. On this level, no assistance from the system exists – not even on the stabilisation level (Donges, 1982). Level 1 refers to assisted driving, where one stabilisation task is performed by the vehicle. This means that one ADAS of category B is active. Hence, the system takes over either longitudinal or lateral guidance, but not at the same time. Level 2 automated driving is when the system has both lateral and longitudinal control and is therefore a combination of two ADASs of category B. At this level, the system can take over the stabilisation task independently, but the driver must still always observe and monitor the system and the functionality of the ADAS to be able to intervene if necessary. The system does not have to provide warnings in the event of errors or failure. In Level 3, the vehicle also takes over longitudinal and lateral guidance within a specific operational design domain (ODD), but the user of the function may engage in NDRAs during active automation. In the event of a dangerous situation or when a system limit is reached, the driver is informed by the system with a certain lead time so that control of the vehicle can be taken over again. This is called a request to intervene (RtI). Thus, the user of the system must always be able to take over while performing an NDRA. This level of automated driving is called conditionally automated driving (CAD), since specific, previously defined conditions must be met before the system can be enabled. In comparison, a Level 4 ADS can execute a minimal risk manoeuvre in case the user does not respond to the RtI in time, and thus the user of the system no longer serves as a necessary fallback level and is not expected to do so. Level 5 ADS or autonomous driving, on the other hand, can handle all traffic situations and is no longer limited to a specific ODD. In theory, a human operator is no longer needed here (SAE, 2021). Figure 2.1. gives an overview of all automated driving levels according to the SAE manual standard.

Level	Name	Narrative definition	DDT		DDT fallback	ODD
			Sustained lateral and longitudinal vehicle motion control	OEDR		
Driver* performs part or all of the *DDT						
0	No Driving Automation	The performance by the *driver* of the entire *DDT*, even when enhanced by *active safety systems*.	*Driver*	*Driver*	*Driver*	n/a
1	Driver Assistance	The *sustained* and *ODD*-specific execution by a *driving automation system* of either the *lateral* or the *longitudinal vehicle motion control* subtask of the DDT (but not both simultaneously) with the expectation that the *driver* performs the remainder of the *DDT*.	*Driver* and *System*	*Driver*	*Driver*	Limited
2	Partial Driving Automation	The *sustained* and *ODD*-specific execution by a *driving automation system* of both the *lateral* and *longitudinal vehicle motion control* subtasks of the *DDT* with the expectation that the *driver* completes the *OEDR* subtask and *supervises* the *driving automation system*.	***System***	*Driver*	*Driver*	Limited
***ADS* ("*System*") performs the entire *DDT* (while engaged)**						
3	Conditional Driving Automation	The *sustained* and *ODD*-specific performance by an *ADS* of the entire DDT with the expectation that the *DDT fallback-ready user* is *receptive* to *ADS*-issued *requests to intervene*, as well as to *DDT performance-relevant system failures* in other *vehicle* systems, and will respond appropriately.	*System*	***System***	*Fallback-ready user (becomes the driver during fallback)*	Limited
4	High Driving Automation	The *sustained* and *ODD*-specific performance by an *ADS* of the entire *DDT* and *DDT fallback* without any expectation that a *user* will respond to a *request to intervene*.	*System*	*System*	***System***	Limited
5	Full Driving Automation	The *sustained* and unconditional (i.e., not *ODD*-specific) performance by an *ADS* of the entire *DDT* and *DDT fallback* without any expectation that a *user* will respond to a *request to intervene*.	*System*	*System*	*System*	**Unlimited**

Figure 2.1: Overview of the automated driving levels (SAE, 2021) (DDT: Dynamic Driving Task; OEDR: Object and Event Detection and Response).

One of the first use cases of higher-level ADSs like CAD is driving automatically on highways. Such an automated system could be called a highway pilot (Wachenfeld et al., 2015). The ODD in this case would be the highway, but very likely not all sections of it. System boundaries may also be congestion ends, highway ends, acceleration and deceleration lanes, and distributor lanes (BaSt, 2012). This implies that there must necessarily be a period of non-availability of the CAD before it is activated on the highway. Moreover, it cannot be guaranteed that a highway trip can be driven automatically without any interruptions. The first Level 3 ADS was released by Mercedes Benz. It works only on highways with a maximum velocity of 60 km/h (Mercedes-Benz, 2022).

2.2 Acceptance and Intention to Use

ADSs of higher levels, i.e. SAE level 3 and above (SAE, 2021), that allow for automated driving at velocities above 60 km/h, are not yet commercially available. Still, this is a topic discussed by the broader public. Automated vehicles can be seen in movies and in car manufacturers' advertising campaigns, whereby these vehicles rarely correspond to the current state of technology and could lead to people's expectations being exaggerated.

When Level 2 automated driving functions that are already available today, like the Tesla Autopilot, are advertised ambiguously (Giesen, 2016), dangerous behaviour can occur when, for example, the monitoring role is no longer consistently performed by the driver (Morando et al., 2021). The fact that expectations regarding current Level 2 ADSs sometimes exceed their functionalities leads to the assumption that expectations regarding ADSs of higher levels will even be higher and might not be met by these systems after their market introduction.

This subchapter summarises research on what future users expect from ADSs and explores if these systems will be able to fulfil these expectations. It will be further discussed how possible discrepancies between expectations and real system capabilities might affect the attitudes of future users towards ADSs.

2.2.1 Expectations Regarding Automated Driving

The expectations associated with the introduction of ADSs comprise increased efficiency and road safety (Lemmer, 2016; Winkle, 2016), and other comfort-related benefits. In an online survey with more than 3000 participants from Denmark (Nielsen & Haustein, 2018), a number of specific expectations were identified, such as arriving relaxed at the destination (which was mentioned most often), not having to park manually and not having to pay attention to traffic. Some participants also expected reduced fuel consumption and thus a reduced negative impact on the environment.

The second most frequently mentioned expected benefit in this survey study was the possibility to devote time to other activities while driving (Nielsen & Haustein, 2018). Other studies have also shown that the possibility to conduct NDRAs is one of the most important advantages of ADSs for future users (Pettersson & Karlsson, 2015; Pfleging et al., 2016). However, the literature also indicates that there can be inter-individual differences in the perception of these systems. For example, only one in four respondents in the above-mentioned Danish survey study believed that ADSs can increase safety. Overall, this study differentiated between enthusiasts, sceptics, and those with a neutral attitude. There were differences between these groups regarding the frequency of mentioning certain expectations. The enthusiasts were those who mentioned positive expectations more frequently. This group consisted to a greater extent of younger, male participants who lived in urban areas, while sceptics tended to belong to the older population in rural areas (Nielsen & Haustein, 2018). In addition, married people had fewer concerns about NDRAs and costs and more concerns about the safety of ADSs. People with higher incomes were more concerned about liability and less concerned about relinquishing control, while people with lower incomes were also more concerned about safety. Car enthusiasts were more concerned about giving up control (Howard & Dai, 2013).

Regardless of inter-individual differences, it appears that there are some specific expectations among the population about the potential benefits of ADSs which must be considered when developing these systems. If expectations are neither met nor corrected, intention to use the system and acceptance might decrease and expected advantages cannot come to fruition (Naujoks, Forster, et al., 2017b).

2.2.2 Technology Acceptance and ADSs

Acceptance in the driving context is a construct that the literature has defined in several different ways. In a comprehensive review, Adell, Várhelyi, and Nilsson (2018a) postulated five different categories into which the diverse definitions of acceptance can be divided. The first refers to acceptance as the willingness to accept a system, a measure or even, for example, a law. In the second category, acceptance is equated with the degree to which users' needs and requirements are met and can thus be described as the usefulness of a system. The third category is similar to the second one, but here rational usefulness is supplemented by emotional components. In the fourth category, acceptance is outlined as the willingness to use – in other words, the intention to adopt a system. The fifth category refers to the actual use of a system and not just to the willingness to use it.

Acceptance can also be differentiated between attitudinal and behavioural acceptance (Saad, 2004). Attitudinal acceptance is based on emotions and experiences and is therefore considered the basis for accepting a system. Behavioural acceptance, on the other hand, refers to observable behaviour, i.e., actual use or the intention to use. A distinction can also be made between contextual and conditional acceptance. Contextual acceptance refers to the willingness to use a system in a certain setting (Saad, 2004), while conditional acceptance depends on certain conditions, e.g., the willingness to only use a system provided it can be switched off at any given time (Adell, Várhelyi, & Nilsson, 2018a). Adell (2009) researched these different definitions and concluded that acceptance in the driving context is the willingness of a person to include a system in their driving or, if the system is not yet available, the intention to use it.

The most frequently used models of acceptance, defining what factors can be used to predict it (Rahman, 2016), are presented in the following.

The *Technology Acceptance Model* (TAM) (Davis, 1989) – which was not developed for assessing technology acceptance of ADSs or manual driving but in the context of information technology – is one of the most widely used models explaining the adoption of new technological systems (Park & Kim, 2014). It consists in its original form of two main factors influencing the usage of systems: *perceived usefulness* (PU) and *perceived ease of use* (PEoU). According to the authors, PU and PEoU influence the attitude towards showing a certain behaviour, which in turn predicts the *intention to use* (ITU) a system. The actual usage is then believed to be influenced by this intention. PU is defined as the degree to which a person believes that a certain system would help in achieving their goal and to which extent using this system would enhance their performance. PEoU is described as the extent to which a person believes that using this system is free of effort. Venkatesh and Davis (2000) extended the original TAM and added several more predicting factors to it: *experience*, *voluntariness*, *subjective norms*, *image*, *job relevance*, *output quality*, and *result demonstrability*. Figure 2.2 depicts the extended TAM version containing the original TAM framed by a rectangle. In this depiction, the attitude towards usage is not included since this factor was removed from the extended version. Thus, this model incorporates the idea that more than the attitude towards usage can influence the actual usage behaviour (Larue et al., 2015).

Although the original TAM was not developed for the vehicle context, it has been used in research to assess driver acceptance (Larue et al., 2015; Park & Kim, 2014; Rödel et al., 2014). The suitability of the model to investigate technology acceptance in the field of ADAS was tested by Rahman et al. (2017). The authors suggest that it can be used in this context and leads to meaningful results. Yuen et al. (2021) found in their study that PU as well as PEoU positively influence users' ITU and are influenced by perceived characteristics of innovation such as result demonstrability and relative advantage. Panagiotopoulos and Dimitrakopoulos (2018) state that

PU is the strongest predictor of acceptance towards automated vehicles (AVs) while Choi and Ji (2015) state that PU might be of more importance to users of ADSs than PEoU.

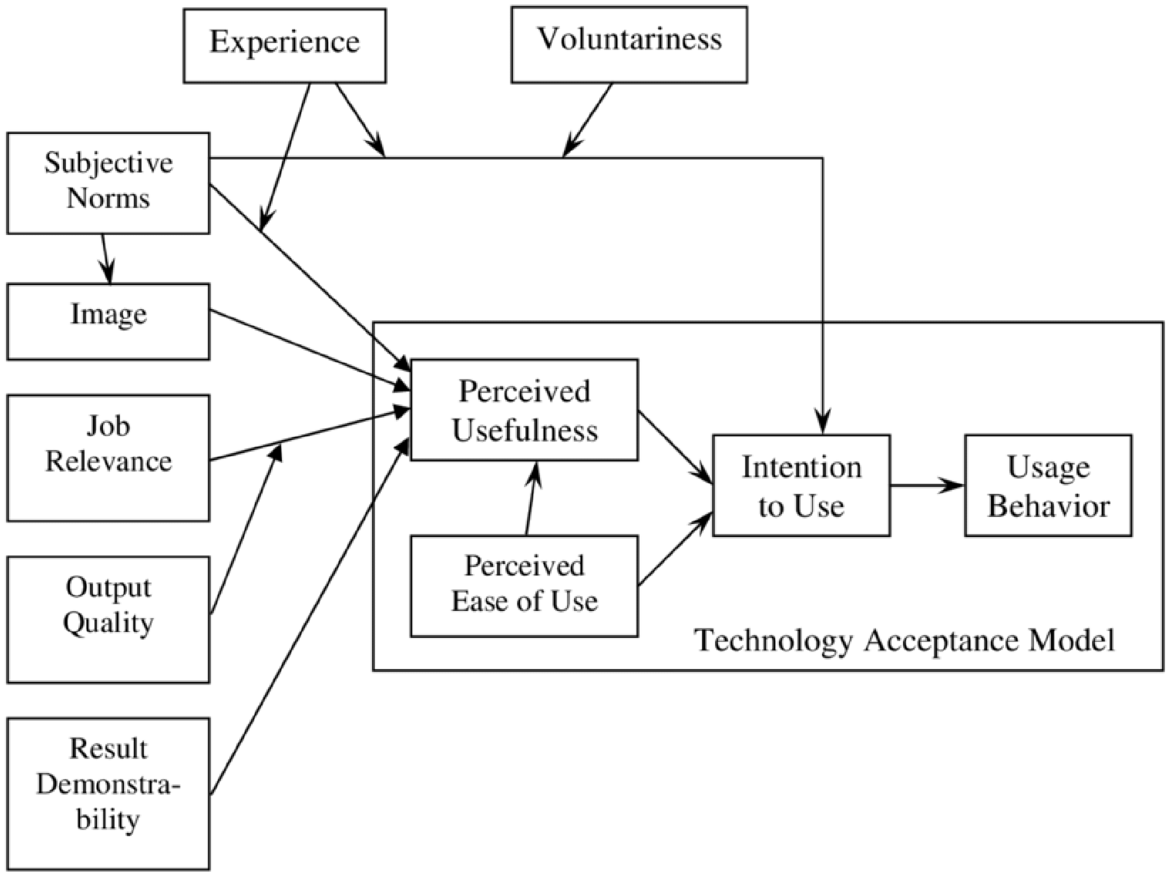


Figure 2.2: Extended TAM (original TAM within the rectangle) (Venkatesh & Davis, 2000).

Another acceptance model aiming at explaining behaviour – not only with regard to technology – is the *Theory of Planned Behaviour* (TPB) (Ajzen, 1991). According to this model, three factors of behavioural intention exist that can predict technology usage behaviour: *attitude*, *subjective norms*, and *perceived behavioural control*. Besides directly influencing the behavioural intention, the *perceived control* also indirectly affects actual usage and is therefore not fully mediated (Rahman et al., 2017). This model was also used in various studies to investigate system usage behaviour in the driving context (Larue et al., 2015; Rödel et al., 2014) and was found to be suitable for measuring technology acceptance (Rahman et al., 2017).

The *Unified Theory of Acceptance and Use of Technology* (UTAUT) (Venkatesh et al., 2003) is a further acceptance model that postulates two direct determinants of usage (see Figure 2.3): *behavioural intention* and *facilitating conditions*. The former is in turn influenced by *social influence*, *performance expectancy*, and *effort expectancy*. The model is based on the TAM and the TPB and other acceptance models and thus shows overlap with those models (Adell, Várhelyi, & Nilsson, 2018b). *Gender*, *experience*, *age*, and *voluntariness of use* serve as moderating factors in this model. Adell (2009) was the first to propose the UTAUT for measuring technology acceptance in the vehicle context and conducted a pilot study. The original items of the model were adapted for this purpose, to be used for the field study. The results showed that *performance expectancy* was the most significant factor in predicting *behavioural intention*, followed by *social influence* (Adell et al., 2011).

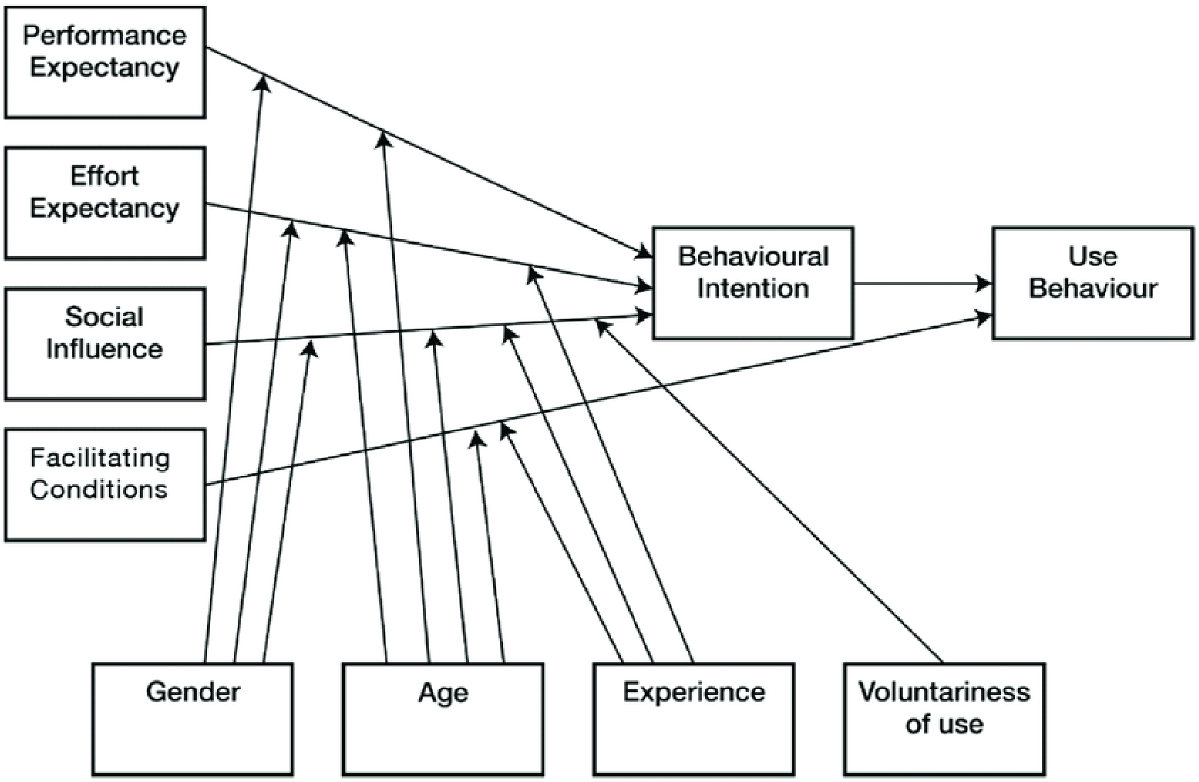


Figure 2.3: The UTAUT model (Venkatesh et al., 2003).

The extended version of the TAM consists of more influential factors than the original one but only one of the additional factors – *subjective norms* – has a direct influence on ITU. The other factors either influence PU directly or serve as moderator variables. As *subjective norms* depend on social norms regarding a system and, since the system of interest in this work is not yet available, this factor cannot be measured and would be of no use for investigating the effects of an HMI concept on acceptance.

Some main differences between the TBP and the original TAM are that the TBP was not created for measuring acceptance regarding technology as well as the dimension of *subjective norms* that exists in the TBP but not in the original TAM and cannot be used in this work as explained above. The UTAUT contains similar factors as the TAM, since *performance expectancy* can be interpreted analogously to PU and *effort expectancy* to PEoU. Besides these two factors, only *social influence* directly influences the *behavioural intention* which is operationalised like ITU of the TAM. This additional factor of *social influence* would not add valuable information in the context of this work. Furthermore, the remaining moderator variables of the UTAUT like *gender, age* or *voluntariness of use* can be controlled in experimental settings.

Taking the considerations above into account, as well as the fact that it is assumed to be one of the most used models when it comes to describing the adoption of new technology in driving context in general (Park & Kim, 2014; Zhang et al., 2021) and also automated driving (Jing et al., 2020; Zhang et al., 2021), the original TAM was chosen to serve as a basis for defining acceptance in the present paper. Its simplicity was assumed to aid in describing the results of the studies carried out for this thesis more easily, especially as no path analyses were conducted to test the single predictors. Moreover, with PU as a factor, the TAM includes the most important predictor of ITU and therefore acceptance. For an overview of different determinants of acceptance towards AVs, the review article from Jing et al. (2020) is recommended.

Since acceptance towards ADSs is deeply linked to people's expectations regarding the performance of a system and their own effort during usage (Adell, Várhelyi, & Nilsson, 2018a; Nordhoff et al., 2016), it is important to be aware of future users' expectations when designing a system that is subject to public discussion before it is available on the market. For example, future users will have expectations regarding which NDRAs will be feasible while driving automatically. The PU – and therefore the acceptance – might be dependent on how the users' task performance conducting NDRAs will be improved by using the system (Davis, 1989; Naujoks, Forster, et al., 2017b; Venkatesh et al., 2003). Hence, it must be taken into account during development in which tasks ADSs are intended to provide support to users. Users' knowledge of a system and its benefits seems to be able to influence attitudes towards it (König & Neumayr, 2017). Both the expectations regarding the use of a technical system and the knowledge about its functioning are stored in the user's mental model (Stanton & Young, 2000), which will be discussed in the next subchapter.

2.2.3 Mental Models and ADS

A mental model is a structure that reflects the user's knowledge and assumptions about a system, which are factors that can influence acceptance and hence usage behaviour (Itoh, 2012; Kazi et al., 2007; Vorm & Combs, 2022). It includes the user's understanding of what a system contains, how this content is linked and how the system works (Carroll & Olson, 1987). These assumptions are subject to falsification tests during usage: if they are not falsified, they are considered valid. All this takes place without conscious processes (Stanton & Young, 2000).

The knowledge that system developers have of a system differs from that of its users, which leads to a difference in their mental models. Therefore, it is important to involve users in the development process to understand these discrepancies and to be able to address them accordingly. In order to make a system user friendly, the interaction problem that exists in the communication between humans and computers should be minimised so that users can solve their actual problems while using the technical system as a help (Bullinger, 1985). In such an interaction, the system is called target system. The complete and consistent representation of the target system is called conceptual model and is created by the developers. The attempt of scientists to represent the users' mental model of a system is called the scientists' conceptualisation (Norman, 1987).

The concept of mental models has been considered in the context of automated driving, as it seems to be fundamental that users know about the functionalities of the ADS and thus what actions are expected from them (Carsten & Martens, 2018). Incorrect mental models could, for example, lead to critical situations before and after transitions (Borojeni et al., 2018), and may lead to a misplaced level of trust in automation and thus to negative behavioural adaptations (Ebnali et al., 2019). One study found that users of ACC – especially at the beginning of usage – have difficulties understanding system limitations (Larsson, 2012). However, a longitudinal study has shown that incorrect mental models of one experimental group, caused by incorrect prior information regarding the system, converged over time through usage with the correct mental models of the other experimental group, caused by correct prior information. The same study also showed that information that was learnt regarding system limitations can be erased from the mental model when the respective limitations are not experienced during repeated usage (Beggiato & Krems, 2013). Hence, feedback about the ADS, its capabilities, and the limits of its capabilities is important to allow users to form a coherent mental model (Beggiato, Hartwich, et al., 2015). Moreover, any adjustment of an incorrect mental model by corrective information from the environment leads to increased cognitive workload (Beggiato

& Krems, 2013). This would also mean, for example, that when a user expects an ADS to be available for a certain amount of time and this expectation is not fulfilled, workload will be increased – not only by the process of taking over itself but also by adjusting the mental model.

Therefore, mental models store knowledge that is gathered about a system during use, but which may also already exist before the first use, and which can be shaped by different information channels. In the case of ADSs, it can be assumed that many laypersons already have a mental model of these systems – even of higher levels of automation – even though such systems are not yet available for purchase. These mental models include expectations, which may be influenced by advertising, movies, and other sources. In this case, the mental model would be formed top-down based on existing knowledge structures (Beggiato & Krems, 2013). Up to and including SAE Level 4, every ADS has limitations (SAE, 2021) and thus, these systems will never be able to meet the expectations of future users who have an almost perfectly functioning automation represented in their mental models (Dzindolet et al., 2003; Kraus et al., 2020).

Mental models cannot only be formed top-down (based on existing knowledge structures) but also bottom-up through situational information from the environment (Beggiato & Krems, 2013), which explains why they are not static – they may change over time (Bach, 2000). Multiple interactions between users and machines contribute to the creation of mental models (Forster, Hergeth, Naujoks, Beggiato, et al., 2019) and incoming information is able to update them. Thus, it can be assumed that by presenting information on system functionalities, incorrect mental models based on false expectations can be corrected. The user interface, or system image (Bullinger, 1985), can be used to influence the user's mental model, since a discrepancy between the system image and the user's model leads to an adjustment of the latter. However, such discrepancies between the mental model and the environment lead to an increase in cognitive workload (Beggiato & Krems, 2013), and therefore presentation of information should be carefully considered. The next subchapter focuses on user interfaces in cars for conveying information, and their potential effects on the users' attitudes and explores what needs to be factored in during the design of such an interface.

2.3 The Human-Machine Interface

The HMI is used for communication between the driver and the vehicle. All input and output systems fall into the HMI category. In the early days of driving, these HMIs were mainly mechanical and conveyed relevant information to the driver such as speed, engine speed or fuel level. Over time, these HMIs no longer proved to be sufficient, as drivers wanted entertainment while driving, for example, beyond the mere provision of information. With the introduction of the radio and navigation systems, the in-vehicle HMI evolved into a complex infotainment system that provides information relevant to driving as well as comfort and entertainment. Today, the scope of the infotainment system includes telephony, navigation, TV, vehicle setting, media, and driving and automation-related information (Meixner et al., 2017). Bengler et al. (2020) created a framework for different HMI components, dividing the HMI in different types, namely the dynamic HMI, the vehicle HMI, the infotainment HMI, the external HMI, and the automation HMI, which contain all information a user needs to interact with an ADS. The following section provides a short overview of the general HMI in the vehicle context.

Inputs are made in the vehicle via control elements. Based on the taxonomy of Rühmann (1993) and Eckstein (2001), Bruder and Didier (2015) have made a human-centred classification of input modalities. According to this classification, hand/arm inputs are required for controlling the vehicle using the steering wheel as well as for operating the secondary and tertiary driving tasks, such as setting the turn signal and operating a touchscreen. Leg and foot movements, on the other hand, are mainly required for pedal inputs. Other input modalities mentioned in this classification are weight shifting, voice input, eye movement, as well as gestures and facial expressions (Bruder & Didier, 2015).

Input devices in the vehicle are mostly haptic. Thus, in addition to the elements for the primary driving task, there are hard and soft keys that are often located in the centre console and on the steering wheel. In addition, rotary controllers as central input devices and levers behind the steering wheel are common in current vehicles. However, there is a present trend towards touchscreens, allowing for a large variety of inputs to be made in one single place in the vehicle (Meixner et al., 2017). The operating tasks include switching a function on and off and setting values and permanent inputs that are necessary for driving the vehicle (Bengler, Pfromm, & Bruder, 2015).

Concerning the output of human-machine interactions, signals can be divided into three different types: acoustic, optical, and haptic (Schmidtke, 1993; Timpe & Baggen, 2000). The visual sense processing optical signals is still the most important when driving (Meixner & Müller, 2017). While the traffic environment is perceived visually, driving task and comfort-related information is also mainly presented using this channel, but often combined with an acoustic stimulus. Due to the high demand put on the optical sense, presenting further complex information visually should be avoided (Bruder & Didier, 2015).

Visual cues in the vehicle can be further categorised into digital, analogue, pictorial, and contact analogue (Bubb, Bengler, et al., 2015), primarily used in infotainment displays. These include the instrument cluster (IC) behind the steering wheel, the central information display (CID) in the centre console, and the head-up display in the windshield (Knoll, 2015). Acoustic signals, on the other hand, can be divided into those that warn and indicate something and those that convey detailed information via speech. In the case of haptically conveyed information, a distinction must be made between surface haptics, which refers to tactile stimuli, and actuation haptics (Bubb, Bengler, et al., 2015).

2.3.1 Challenges for In-Vehicle HMIs

The focus of the present work is on the automation HMI and information needs concerning the ADS. Since some ADS-related information might have to be presented while driving manually, the challenges for in-vehicle HMIs are described in this subchapter also with reference to manual driving.

For the driver to be able to operate all the different functions, and to process the large volume of information conveyed through different channels, a careful design of the HMI and all related components is necessary. In addition to conventions and guidelines for the anthropometric design of the vehicle interior, which also concerns the positioning of the screens and controls (Bengler, Pfromm, & Bruder, 2015; Bubb, Grünen, & Remlinger, 2015; DIN, 1953; ISO, 2009), there are norms that address system ergonomics. These include requirements for the presentation of information in general (DIN ISO, 2017), requirements for the design of information displays in the vehicle (ISO, 2003), as well as guidelines for the development of in-vehicle information systems with particular reference to safety (Stevens et al., 2002) and system ergonomic design maxims (Bubb, Bengler, et al., 2015), to name only a few.

Since the complexity of the primary and secondary driving tasks is increasing due to the growing number of different functions and potential distractors, and at the same time the tertiary activities themselves are becoming more and more complex (Bengler, 2017; Bubb, Bengler, et al., 2015), it is important that constructs exist that can be used to assess the system-ergonomic quality of HMIs. One of these constructs is usability, published in the DIN ISO norm 9241, which describes the extent to which a product can be employed by users to perform tasks satisfactorily, efficiently, and effectively (DIN ISO, 2017). Satisfaction is defined as the extent to which the user's emotional and cognitive responses, which result from interaction with the product, are consistent with expectations and requirements. Efficiency describes the relationship between the resources used, including time, effort, materials, or money, and the results achieved. Effectiveness is the completeness and accuracy with which tasks are fulfilled or goals are achieved.

An exploration of usability should not leave out user experience (UX), as the definitions of both constructs often overlap (Bengler, 2017). In the ISO standard, UX is defined as the user's perception and reaction resulting from interaction with the system. It should be noted that UX refers to the reactions before, during and after usage (DIN ISO, 2017). UX includes other aspects that may be associated with a product and can have an influence on the emotional or cognitive reaction, such as brand image and the user's personality, state, and experiences with the product (Mirnig et al., 2015), and is therefore a broader construct than usability. Usability and UX are seen as selling points for vehicles (Bengler, 2017), since good usability positively influences the usage of a system in the vehicle as the probability of conducting tasks satisfactorily increases with increasing usability. Nonetheless, a high level of usability does not automatically imply an absence of distraction and therefore care must be taken to ensure that the level of workload is not too high (Bengler, Bubb, et al., 2015).

Workload is characterised by the situation surrounding the operator, including the work task, the environment and the systems available, and by the load that follows in response to this situation (Bubb, Vollrath, et al., 2015). Too high a workload is associated with a decrease in driving performance (Hancock & Verwey, 1997), making it imperative that critical driver distractions be considered in the development of driver assistance systems and HMIs (Bengler, 2017). According to the multiple resource theory (Wickens, 2008), different activities can be performed simultaneously without adverse effects as long as different and non-interfering perceptual channels are addressed. On the other hand, the simultaneous use of, for example, the visual channel by two different tasks can increase the workload and thus impair task performance, for instance when driving (Lansdown et al., 2004).

2.3.2 HMIs and Information Needs for Automated Driving

Research on HMIs regarding automated driving can be divided into three areas: methodological research, research on HMI modalities, and content-related issues such as information needs. Research on methodological issues deals, for example, with topics such as the creation of survey instruments to be able to test HMIs (Siebert et al., 2013) or the suitability of existing subjective measures for investigating HMIs in automated driving contexts (Forster et al., 2020). Other methodological issues relate to how studies should be designed to test HMIs, from creating suitable test scenarios for various HMI-related research questions to recommendations on how to check the safety of HMI displays (Hock et al., 2018; Naujoks et al., 2019; Naujoks et al., 2018; Schömig et al., 2020; Warg et al., 2020). In addition, there are research efforts that address the question of how to best test usability of HMI displays in automated driving (Albers et al., 2020; Forster et al., 2018). Albers et al. (2020) have formulated best practice recommendations that answer questions such as which test environment is most suitable or how a study sample should be selected when testing the usability of an HMI-system concerning automated driving. Their recommendations suggest that a study should take place in a driving simulator, that the study population should be representative of future users and that transitions between different levels of automation should be tested in the presence of several automation levels and in non-critical situations.

When considering the research that deals with the presentation of certain information as well as with different output modalities, it is helpful to divide them according to the respective system states in which the corresponding information is to be offered. This thesis follows the classification of the system states of CAD, hence a Level 3 ADS, into the following phases: manual driving with non-availability of the ADS; manual driving with availability of the ADS; automated driving including the take-over process when reaching a system limit. The following is a summary of the different information displays that were explicitly mentioned to be part of the HMI in studies investigating automated driving.

Only very few research papers were found that refer to displays of ADSs during manual driving while the ADS was unavailable. Only Wandtner et al. (2018) report having given information regarding the ADS in this state in a study. The participants in the experimental group experienced an overview of all availability sections on their route in each system state, i.e., even when the ADS was not available, visually in the centre console. However, there were no research questions on how this display during manual driving affected the drivers and their attitudes regarding the ADS.

For the manual driving phase with an available ADS, a number of research papers indicate how the availability of the ADS was indicated. In most works, a visual signal in the IC, in the HUD or via LED was coupled with an acoustic signal (Feldhütter et al., 2018; Kraus et al., 2020; Wandtner et al., 2018). In addition to this acoustic signal, Muthumani et al. (2020) describe having selected a voice output with the content “hands off” when availability was indicated. None of the cited papers reports having explicitly investigated the displays for ADS during manual driving.

More papers give insights into in-vehicle information content displayed in studies during automated driving, such as current speed (Feierle, Bücherl, et al., 2020; Feierle, Danner, et al., 2020; Hartwich et al., 2020; Hecht, Kratzert, & Bengler, 2020), speed limit (Feierle, Danner, et al., 2020; Naujoks, Forster, et al., 2017a), surroundings (Naujoks, Wiedemann, & Schömig, 2017), upcoming traffic events (Forster et al., 2016; Y. Yang et al., 2018), as well as the distance to upcoming traffic events (Naujoks, Forster, et al., 2017a). Moreover, navigation information (Feierle, Danner, et al., 2020; Hartwich et al., 2021; Oliveira et al., 2020) and details about the current traffic situation (Hartwich et al., 2020) are mentioned to be shown to participants while driving automatically in driving simulators. This information often aims at increasing transparency of what the vehicle perceives from the environment. Nonetheless, it can also be provided during manual driving and is not necessarily ADS-related and therefore does not fall into the category of an automation HMI (Bengler et al., 2020). Information content displayed in studies that is more directly related to ADSs is the current and the available driving mode (Feldhütter et al., 2018; Forster, Hergeth, Naujoks, Krems, & Keinath, 2019a; Lau et al., 2020), indicating whether the ADS is active or not. In addition, the distance to and reason for future manoeuvres can be presented to users along with the vehicle’s intentions (Hartwich et al., 2021; Naujoks, Forster, et al., 2017b; Wintersberger et al., 2018). Another piece of information mentioned in the literature is an indication of how well the ADS can handle the current situation (Kunze et al., 2019). In most studies, RtI-related information is presented to participants, e.g., in Feierle, Holderied, and Bengler (2020) and Gold (2016). Furthermore, the durations of current and upcoming sections of (non-)availability of the ADS are information needs mentioned in several papers (Beggiato, Hartwich, et al., 2015; Richardson et al., 2018; Wandtner et al., 2018). However, no influence of presenting the duration one can drive in automated driving mode on subjective evaluations or objective measures has been observed so far. There are no concrete research results on many of the information contents mentioned here, as these were presented in studies but were not subject of any specific research questions. Nevertheless, there is research that suggests that the need for information may depend on interpersonal and situational factors. For instance, users with less system experience (Beggiato, Hartwich, et al., 2015) or with lower initial trust in the system (Hartwich et al., 2020) may need more information on the ADS than comparison groups. According to one study, gender also plays a role in this regard, as does whether or not study participants engage with NDRAs during automated driving (Lau et al., 2020). An online survey concluded that less information about the system is needed during active engagement in NDRAs than without engagement (Hecht, Darlagiannis, & Bengler, 2020).

In some studies, visual displays such as ambient lights (Muthumani et al., 2020; Y. Yang et al., 2018) or augmented reality head-up displays (Schömig et al., 2018) were chosen to convey ADS-related information. Sometimes, information was also paired with an acoustic cue, for example regarding the RtI (Feierle, Holderied, & Bengler, 2020; Feldhütter et al., 2017; Gold et al., 2013). Nonetheless, information on automated driving is most frequently displayed visually in the IC in the studies mentioned above.

3 Research Questions, Methodology and Structure

This chapter briefly summarises the theoretical background and leads to the research questions this thesis aims to answer.

3.1 Research Questions

Higher level automated driving allows users to engage in NDRAs while the vehicle is in control of the driving task. In SAE Level 3, users serve as a fallback level in case a system limit is reached, while in SAE Level 4, the vehicle is able to handle all situations within the predefined ODD by itself. In both levels, however, users must perform the driving task themselves before reaching the ODD and again after leaving it. This means that an automated journey will always be of limited duration and that NDRAs that take longer than the automated driving section lasts must be interrupted. This thesis focuses on developing an HMI concept for such SAE Level 3 and 4 systems.

Research shows that the possibility to perform NDRAs is one of the most important reasons to purchase a vehicle with an integrated ADS and to activate the same. In addition, there is a picture of very high expectations among laypersons regarding the capabilities of future automated driving functions, which may not be fulfilled. A mismatch between the users' mental models that were built top-down based on expectations resulting from previous knowledge and the actual capabilities of the system can negatively influence the human-machine interaction. Moreover, unmet expectations – e.g. when the system's capabilities are overestimated – are likely to affect users' attitudes towards the ADS leading to decreased ITU and acceptance (Lee & See, 2004). Since acceptance is associated with actual usage and the benefits of an automation system can only be realised if it is used, it will be one construct this thesis focuses on. The usefulness – which is part of acceptance in the TAM – of an ADS can only be received by users if they use the system purposefully, meaning that the system actually helps them to achieve the goal they pursue with the activation of the ADS.

The high expectations of future users and the limited availability and thus limited feasibility of NDRAs could have a negative impact on the usage behaviour. As described above, the mental models of users can be adjusted to the reality of technical possibilities by providing information on the system. Hence, it is necessary to collect the information needs future users have regarding the availability of the ADS. The following research question was defined:

1. *Which information do users need regarding the ADS and its availability?*

Since the journey in a vehicle with an ADS can be divided into sections of manual driving with the automation being unavailable, manual driving with available automation, and automated driving, not including transitions, it must also be investigated in which system state the information addressed in the first research question should be presented:

2. *In which system state do the users need information regarding the ADS and its availability?*

The way HMIs are designed influences the usability of the system image and can also have effects on acceptance and mental workload. Therefore, the way information is presented in the vehicle can also affect driving safety and thus, the manner of interaction, especially when complex information regarding availability is displayed, will also be given consideration in this work:

3. *How should information regarding the ADS and its availability be conveyed?*

3.2 General approach

To answer the research questions listed above, this thesis followed the guidelines of user-centred design according to DIN EN ISO 9241-210 (International Organization for Standardization, 2019), which describes a development process for interactional systems that should be both usable and useful. An important part of this process is the understanding and involvement of the users or – if they do not yet exist – future users. This includes a comprehensive understanding of the tasks the system is intended to support and the environment and context in which the system will be used. In addition, the process provides for iterations of the single steps, as requirements may be added or changed during the development.

Specifically, requirements for a system are defined based on what is known about the task and the context of use. These requirements are then translated into design solutions, which in turn are evaluated. From each step of the process, it is possible to jump back to a previous step to make adjustments. For example, after evaluating a design solution, it is possible to go back to defining requirements or simply make a change in the design (Mirnig et al., 2015).

Applying these design principles to the research questions of this thesis, several steps were taken on the way to finding a solution. First, the users had to be well understood regarding their task of using the ADS in a meaningful way. For this purpose, a focus group interview was conducted with experts in the field of automated driving and/or in the field of human-machine interaction in the vehicle context. Based on the outcome of the focus group interview, an exploratory driving simulator study was conducted with laypersons. These two explorative investigations are reported in Chapter 4. Using the findings of these two different approaches, information requirements, for an HMI concept were defined. The requirements were then operationalised, and the fulfilment of these requirements was analysed regarding their benefits for future users. Based on literature research on each requirement, hypotheses were formulated, and tested through confirmatory research approaches using driving simulators. For this purpose, simple design solutions were used to enable the graphical presentation of specific information. The testing of the single information components is reported in Chapter 5. After evaluating the single information needs, those that were found to be relevant were combined into an overall concept and again evaluated in terms of usability and other measures considered relevant as outlined above, which is described in Chapter 6.

An alternative approach would have been to gather all information needs that are collected in the explorative study to one HMI concept instead of testing the single elements separately. This way it would have been possible to take more iterations of improving the complete HMI concept. Nonetheless, by using this approach it would have not been possible to trace on dependent variables back to one HMI element which is considered as a prerequisite to remove a concept component that serves no purpose.

3.3 Methodology and Questionnaires

For the confirmatory driving simulator studies open and semi-open questions to study participants were combined with quantitative surveys in the form of questionnaires, to ensure comparability within this work. Three questionnaires were used in each of the studies, except for the first exploratory one, all of which are described in the following. Reference is made to this in the respective chapters on individual studies.

The first of the three questionnaires was the Acceptance Scale (van der Laan et al., 1997), which measures acceptance of technical systems in a vehicle context on a semantic differential. The questionnaire measures two dimensions: *usefulness* and *satisfaction*. For this purpose, two opposing pairs of terms are juxtaposed, and study participants indicate which term applies more to them with a neutral option. An example is the comparison between the terms “useful” and “useless”. The TAM questionnaire was not used as an instrument to measure acceptance in this work as it was not developed for investigating automated driving acceptance, and therefore the items would have had to be adjusted to the driving context, which could have negatively influenced the validity of the test. The Car TAM questionnaire developed by Osswald et al. (2012) was also not chosen in this work since it is based on the UTAUT, and therefore includes dimensions and predictors that are not part of the more simple acceptance definition of the original TAM. The Acceptance Scale (van der Laan et al., 1997) on the other hand may not be based on the TAM but consists of two dimensions that are similar to the two predicting factors of this model. *Usefulness* can be assumed to be related to PU and *satisfaction* to be related to PEoU. Furthermore, it is frequently used when researching ADSs (Forster et al., 2020; Hecht, Kratzert, & Bengler, 2020; Körber et al., 2018; Zoellick et al., 2019). In addition, when the factor loadings of the corresponding dimensions are insufficient, an overall acceptance score comprising all items can be created (Beggiato, Pereira, et al., 2015). The same approach was chosen in this work.

As outlined above, usability is another important factor that should be considered when investigating in-vehicle HMIs. To measure this, the System Usability Scale (Brooke, 1996) was used in this work. This questionnaire consists of 10 items that are answered on a 5-point Likert scale, an example item being *I think I would like to use this system frequently*. The answer options cover a range from 0 to 4, from strongly disagree to strongly agree. The total sum of the answers is multiplied by 2.5 to obtain a total range from 0 to 100. It should be noted, however, that these are not percentages.

Workload can be increased by a new HMI concept or additionally displayed information and was surveyed in this work using the NASA Raw Task Load Index (NASA-rTLX) (Hart, 2006). It consists of six items with a response scale from 0 to 20. Each item describes its dimension discretely and can therefore be evaluated separately. However, it is also possible to obtain an overall value for workload by summing up the individual answers and scaling the sum to 100. The six dimensions are mental demand, physical demand, temporal demand, overall performance, effort, and frustration.

In the studies described in Chapter 5 and Chapter 6, these subjective measurements were used to compare HMI concepts or to test them against pre-defined threshold values. All HMIs and questionnaires used in this work were presented to participants in German. In the following chapters, the content of the HMIs has been translated to English when mentioned in the text to retain language consistency. Nonetheless, figures showing pictures of the HMIs are the original ones and therefore in German.

Hypotheses were formulated for each study before conducting statistical tests. In the respective chapters, all alternative hypotheses can be found but without explicit formulation of the null hypotheses. A null hypothesis is the logical inversion of the alternative hypothesis and can therefore be implicitly derived from latter (Field et al., 2013). For example, when the alternative hypothesis states that a difference in the outcome variable exists between difference groups, the null hypothesis would state that the outcome variable would be the same for all groups.

4 Information Needs Regarding the ADS Before Activation – An Exploratory Study

Transitions between manual and automated driving are largely studied in the context of RtIs, i.e., the takeover of control by the human driver after a period of automated driving, which can be explained by the criticality of this interaction (Lu et al., 2016). As mentioned above, the activation of ADSs has not received much attention from researchers so far. Lu et al. (2016) claim, for example, that this issue is trivial, similar to the activation of an ACC. However, many users do not even know all the functionalities and limitations of this comparatively simple ACC system (Beggiato & Krems, 2013), and therefore, the operation of an ADS – including its activation – will most likely be complex as well and thus requires investigation.

Specific conditions have to be met for users to be able to activate an ADS of Level 3 or 4 (SAE, 2021). Furthermore, users need an adequate mental model of the system and its capabilities and limitations to activate the system safely (Forster, Hergeth, Naujoks, Krems, & Keinath, 2019b). As these conditions can depend on external circumstances such as weather conditions and structural aspects of the road and can also change over time on familiar routes, it is necessary to provide the information required by the user while driving. The purpose of this study is to explore the information needs of users regarding ADS and its availability. The study described in this chapter was published in Danner, Pfromm, Limbacher, and Bengler (2020).

4.1 Research Questions

To be able to speak of a purposeful activation of the ADS, drivers must know what purpose they are pursuing with the activation, and whether this goal was fulfilled must be evaluated at the end of the automated drive. Thus, based on HMI concepts investigated in the literature so far (see Chapter 2.3.2), a user can only evaluate retrospectively if their goal has been achieved. However, to enable users to decide before an activation whether activating the ADS serves their purpose, this study exploratively collects information regarding the expectations of future users.

Future users are explicitly referred to as novices in this study, as they have no prior technical knowledge, which makes it a beneficial user group for gathering basic information needs. As described above, users with less experience tend to having a greater need for information

(Beggiato, Hartwich, et al., 2015). The first question refers to the expectations of these novices regarding the availability of ADS:

1. What do participants expect regarding the availability duration and frequency of L3 automated driving systems?

The second question refers to the information needs of novices regarding the availability of the ADS, especially before it has been activated:

2. What kind of information do participants need before activating the ADS?

Presuming that reasons for non-availability of the ADS are potentially not always intuitive, three more questions are explored regarding the assumption of novices and their information needs, along with one question regarding their intentions to use the ADS:

3. What reasons for non-availability do participants assume when the ADS is not available?
4. What kind of information do participants need regarding the ADS when it is not available?
5. Why would participants want to activate the ADS and why would they choose to drive manually when ADS is available?

4.2 Method

Two different qualitative methods were used to answer the research questions outlined above. Firstly, a focus group interview was carried out with experts in the field and secondly, a driving simulator study was conducted with laypersons.

4.2.1 Focus Group Discussion

The focus group interview was conducted with $N = 5$ experts on the topic of automated driving. All participants were employees of AUDI AG. In the context of this study, they were considered experts because they were either involved in the development of ADS as engineers or were engaged in the development of user interfaces for the communication between drivers and AVs as HMI experts. In addition, each of the participating experts had gained experience with novices in a prototypical AV and was therefore considered able to both report from their experience behaviour and statements of novices and give their professional assessment. The interview was scheduled for one hour and was recorded with the help of audio equipment after all participants had agreed. The conversation was moderated with the help of a previously defined interview guideline which consisted of four content-related topic blocks that served to trigger a professional discussion among the participants and originated from the above-mentioned research questions. These topics included:

- General reporting of experts' experience with novices in AVs
- The perceived expectations of novices regarding an ADS and its availability
- An assessment by the experts of how realistic these expectations are
- Elaboration of information, the presentation of which could reduce the potential discrepancy between novices' expectations and technical feasibility

Subsequently, the group discussion was transcribed and analysed through a qualitative content analysis according to Mayring (2015) with a focus on deductive allocation into categories. These categories corresponded to the four thematic blocks from above.

4.2.2 Driving Simulator Study

After analysing the results from the focus group interview, a driving simulator study was conducted to compare the findings from the expert interview with the behaviour and statements of study participants who were novices.

4.2.2.1 Participants

Overall, $N = 15$ participants took part in the study; 11 of them were men and 4 of them were women. The average age was $M = 27.5$ years ($SD = 3.1$). The average annual mileage among the participants was $M = 12.21$ km ($SD = 13.01$). Regarding driver assistance systems, 20% reported having experience with ACC systems, 13% with a lane keeping assistant and 13% with a parking assistant. None of the participants had had any prior experience with ADSs.

4.2.2.2 Simulator, Route, and Procedure

The study was conducted in a fix-base driving simulator at AUDI AG. The software Virtual Test Drive was used for track programming. The simulator setup is shown in Figure 4.1.

After the participants were informed about the trial and had given their written consent, they experienced a five-minute test drive to familiarise themselves with the driving simulator and the operation of the ADS, which was activated and deactivated by pressing a button on the steering wheel. During this drive, they were exposed to both manual and automated driving sections. Thus, they experienced the interactions for availability of the ADS in the HMI as well as RtIs.

Figure 4.1: Driving simulator setup.

Subsequently, four experimental drives followed. These took place on the same route in the simulator, differing from each other only in the availability of the ADS, the traffic density, or the duration of the journey. The order of the scenarios was randomised. Three of the drives lasted 5 minutes and one drive about 8. The ADS was available on three of the drives and on one it was not. When participants tried to activate the ADS when it was not available, a pop-up was displayed in the IC in German with the text "automation not available: route section not appropriate". Whenever the ADS was not available, the simulated road was missing the emergency lane. After each of the four journeys, participants answered semi-standardised interview questions. The answers were noted down by the conductor of the experiment. After completion of all drives, the participants rated the information needs that were raised in the focus group interview.

Before driving the first scenario, the participants were instructed to imagine that the upcoming trip was their daily commute home after work. They should imagine they wanted to use this drive to watch a short video on the tablet that was mounted in the cockpit below the CID. The participants were also instructed to leave the motorway at the next exit and stop at the motorway lay-by. They were asked to watch the video only while driving automatically and never while driving manually. If they had to take control of the vehicle, they would have to stop the video. In the first scenario, the ADS became available after about 20 seconds on the motorway until about 15 seconds before the highway exit. The video lasted 4 minutes and therefore the participants had the opportunity to finish watching it. Thus, this scenario is meant to represent the ideal case where users want to perform an NDRA and can complete it without interruption during an automated journey. Figure 4.2 shows the different scenarios.

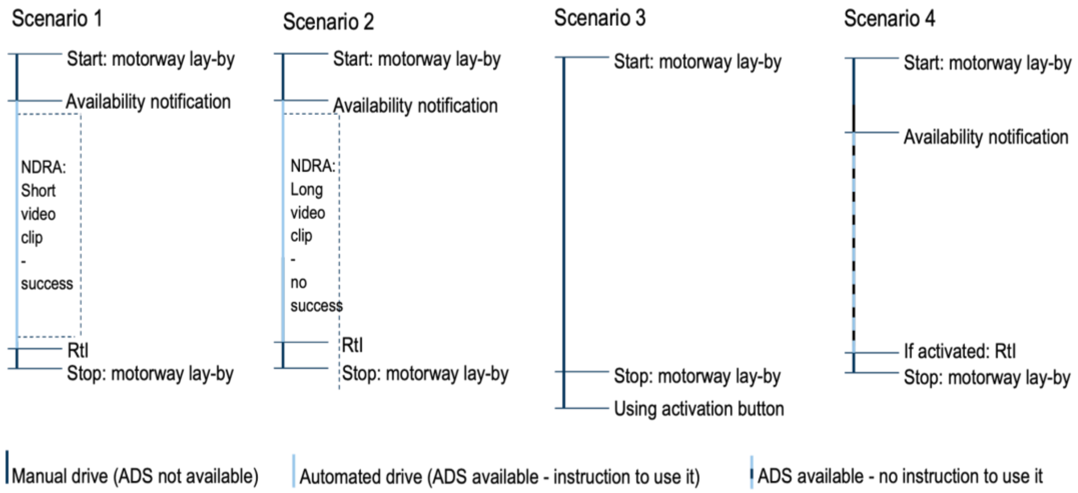


Figure 4.2: Scenarios experienced by study participants.

The instructions for the second scenario were similar. The only difference was that the participants were instructed not to take the next motorway exit but the one after that. The video to be watched during this drive was 8 minutes long. Thus, the video could not be finished until the RtI occurred again at the first motorway exit, i.e., at the same point as in the first scenario. For the remaining three minutes until the second exit and the completion of this experimental drive, the ADS was not available anymore. Therefore, the scenario simulated a situation in which users cannot complete their NDRA without interruption.

Before the third scenario, the participants were instructed in the same way as in the first scenario with the only difference that this time they were not asked to watch a video. They were instructed to drive as they pleased, manually or automatically, depending on the availability of the ADS of course. During this drive, the ADS did not become available. Hence, a scenario was simulated in which the unavailability of the ADS prevents the user from engaging in NDRAs, although they presumably expect availability. The sole deviation in the simulated route compared to the other scenarios was the absence of an emergency lane. If participants did not try to activate the ADS during this scenario, they were asked to use the activation button after stopping the car. This way it was ensured that every participant saw the notification stating non-availability.

The instruction for the fourth scenario was the same as for the third. However, on this trip the ADS did become available. In comparison to the other drives, the traffic density was also higher. Thus, an unpleasant highway drive was simulated.

4.2.2.3 Measures

After each drive, open questions were asked regarding information displays the participants would have wished for. Furthermore, the information needs raised in the focus group interview were rated on a 5-point Likert scale. Participants were asked how important and useful they would consider a presentation of the respective information.

4.3 Results

4.3.1 Focus Group Results and Discussion

According to the experts, novices' expectations of the ADS tended to be too high, especially if no prior experience with an ADS had been gained. Laypersons seemed to feel disturbed by RtIs and sometimes had difficulties understanding or even accepting system limits.

Novices seemed to expect infinite availability durations even if they knew about the possibility of RtIs. Although they were informed that the ADS they were experiencing would only work on the highway and that the availability of this system was subject to specific conditions, they still seemed surprised when they had to take control of the vehicle because the ADS did not work in certain areas. They also seemed to expect to be allowed to sleep while the ADS was activated, even if they knew they had to act as a fallback.

The experts considered these expectations not realistic and stated they cannot be fulfilled within the next few years with the introduction of Level 3 ADSs. They considered periods of 30 to 40 minutes of automated driving time to be realistic but assumed that interruptions would occur.

The discussants concluded that non-transparent system limits might lead to a discrepancy between the expectations of novices and actual technical functionalities in Level 3 ADSs. They felt that an indication of how long the ADS was available – presented before and after the activation – could help users to adjust their expectations to the actual possibilities and thus choose their NDRAs accordingly to prevent frustration by being forced to interrupt an activity. As this availability duration would be an estimation, the experts suggested displaying an indication of how certain this estimation would be. Hence, transparency could be increased in case the estimated availability duration is not entirely correct. In addition, they considered it likely that suggestions about which NDRAs could be completed during the automated drive, following the logic of a recommender system, could be of help to users. Also, an overview of all sections on which automated driving is available on one's route – displayed during all system states – might help users organise their NDRAs correspondingly. In case of non-availability, a display of when the ADS will be available could add value as well as reasons why the ADS is not available. Table 4.1 summarises the information needs generated.

Anticipated information needs when ADS is available	**Anticipated information needs when ADS is not available**
Estimated availability duration of the ADS before activation	Reasons for non-availability
Certainty of availability duration	Duration until ADS is available
Overview of availability periods on the whole route	
Suggestions of NDRAs feasible during the automated drive	

Table 4.1: Information needs gathered in the focus group interview split by system states.

4.3.2 Driving Simulator Study

Even though the experimental drives were completed in randomised order, they are referred to as scenario 1, scenario 2, and so on, to ensure identification with the routes described above in the methods section of this chapter.

After experiencing scenarios 1 and 2 – which were classified as ideal and non-ideal depending on whether the availability duration was sufficient for completing the NDRA or not– the participants were asked what availability duration they would have initially expected. Eight out of 15 participants stated they would have assumed infinite availability and that they would have manually deactivated the ADS before leaving the highway. Four said they expected short periods of availability, two said they expected to be able to complete the NDRA and one participant did not specify a period but expected availability until a critical event occurred. The frequencies are shown in Figure 4.3.

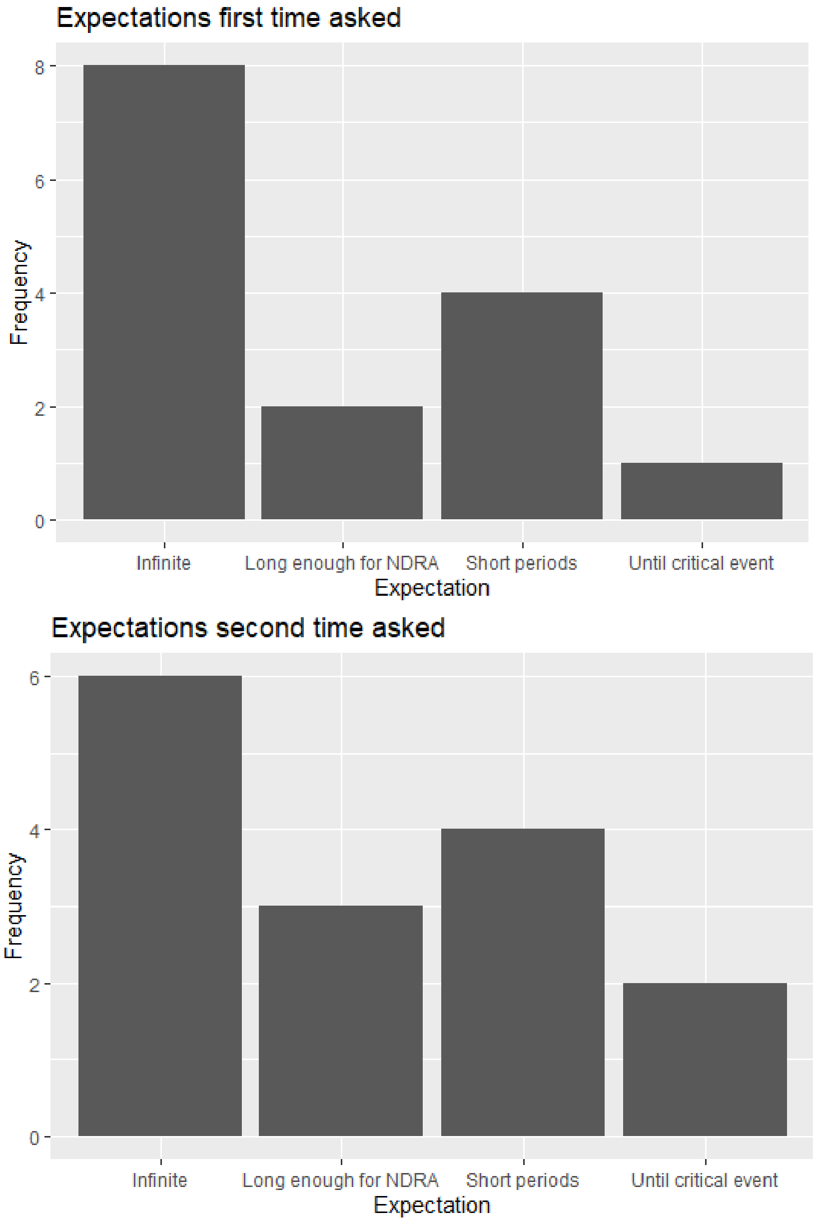


Figure 4.3: Frequency of mentioned expectations regarding availability when asked for the first and the second time (Danner, Pfromm, Limbacher, & Bengler, 2020).

After experiencing scenarios one and two, study participants were also asked what information they would have needed before activating the ADS. Out of 15, ten stated they would have wished for a time duration or distance indication of how long the ADS would be available. Two indicated that they did not need any additional information before activation but wanted a time period of availability after activating the ADS. In addition, two respondents would have liked an indication of whether the ADS would be able to handle the situation. Figure 4.4 shows the frequencies of responses.

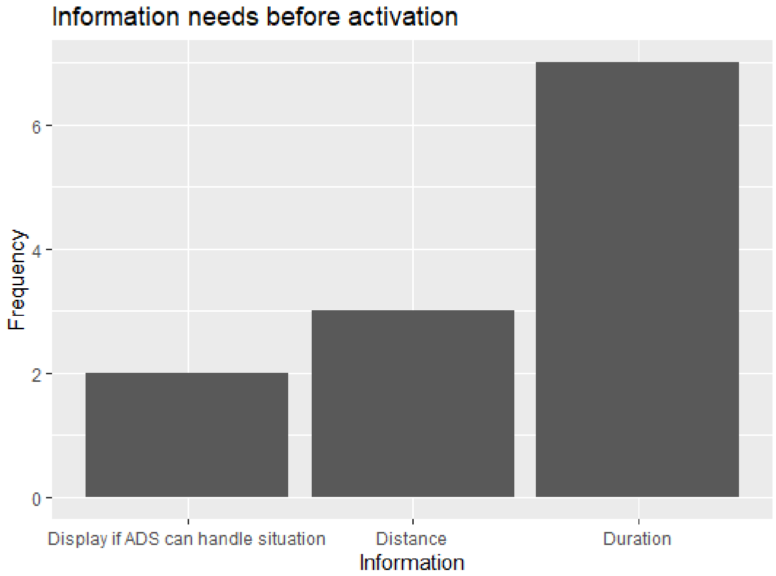


Figure 4.4: Information needs before activating the ADS (Danner, Pfromm, Limbacher, & Bengler, 2020).

To identify information needs when the ADS was not available, participants were asked after scenario three (in which the ADS did not become available) what information they would have found useful during this drive. Out of 15, eight said they would have liked information about when the ADS would be available. Four said they did not want to receive any information regarding the ADS when not available. Also, four stated they would have appreciated a permanent icon to indicate the non-availability. Three stated that they missed a notification about the reasons for unavailability. Figure 4.5 shows all answers along with their frequencies.

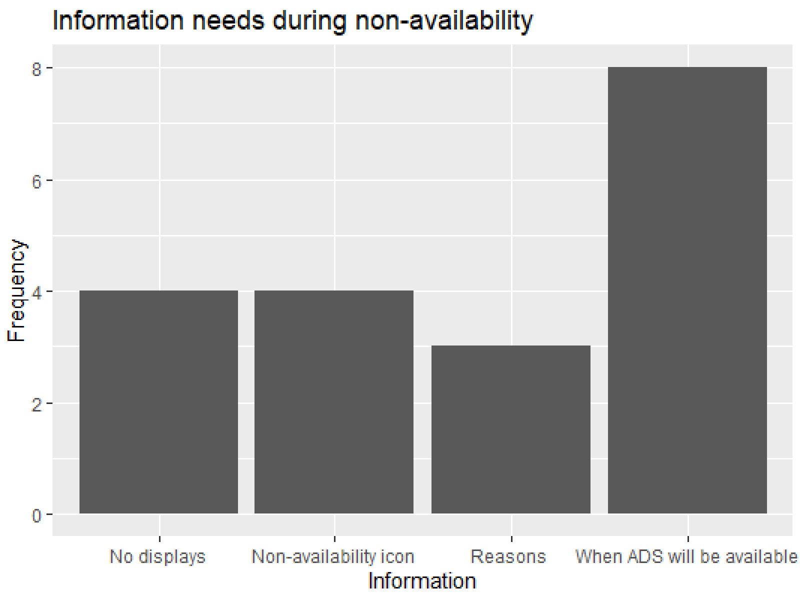


Figure 4.5: Information needs regarding the ADS when not available (Danner, Pfromm, Limbacher, & Bengler, 2020).

Of the 15 participants, six tried to activate the ADS even though it was unavailable and thus saw the above-mentioned notification. The nine remaining participants noticed the notification only after completing scenario 3, after being instructed by the conductor of the experiment to press the activation button. When asked if this hint was sufficient and if not, what else they would have wished for, eleven of the participants stated this feedback was not sufficient and desired more detailed reasons for non-availability. The study participants were also asked what reason for non-availability they assumed during the drive. Six answered they suspected the traffic density had been too high to drive automatically and four thought technical reasons had caused the non-availability. None of the participants gave the correct reason, which was the lack of an emergency lane.

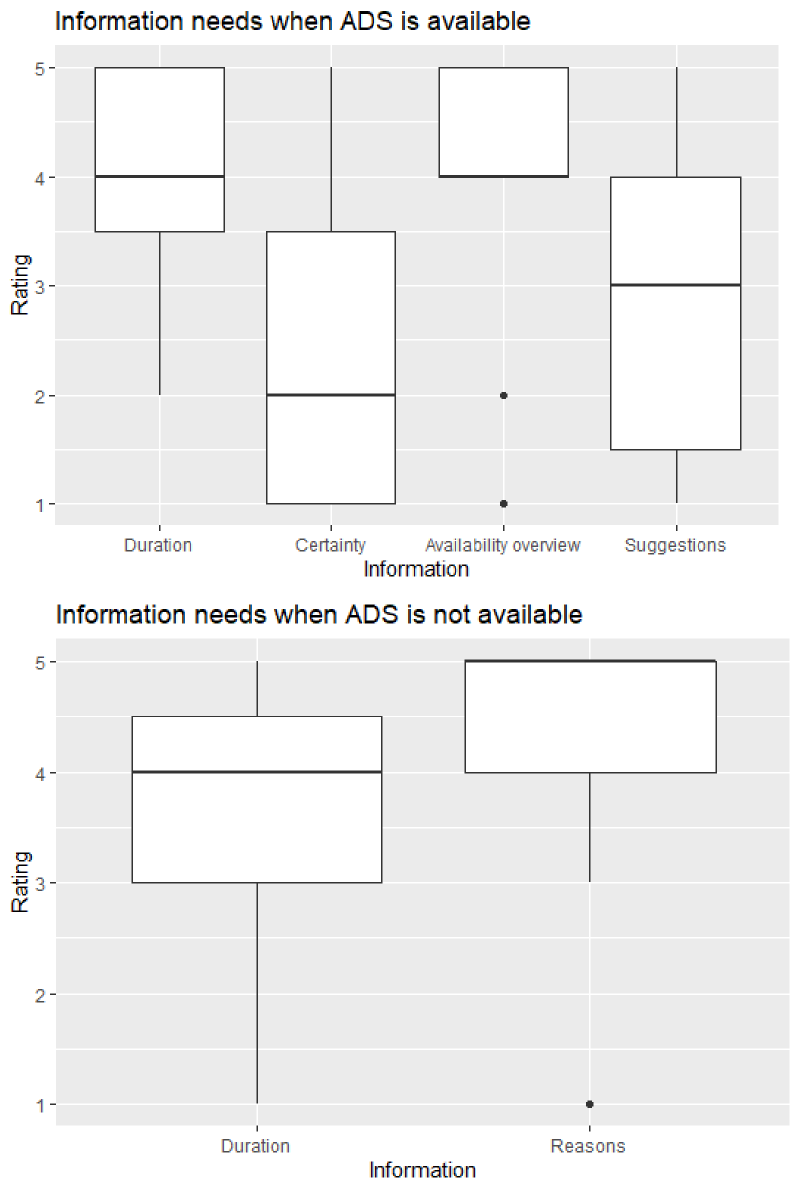


Figure 4.6: Rating of the information needs generated in the focus group discussion (Danner, Pfromm, Limbacher, & Bengler, 2020).

At the end of the four drives, participants were asked to rate how important and how useful they considered the presentation of the information generated in the focus group. The results can be seen in the form of boxplots with medians in Figure 4.6. A higher rating indicates greater perceived importance.

To answer the last research question, participants were asked what reasons they would have to activate the ADS and what reasons they would have to drive manually even though the ADS was available. The most frequently cited reason to drive automatically was to perform NDRAs, and the second most frequently cited reason was when traffic would not allow faster driving anyway. The most common reason for choosing not to activate was the desire to drive faster than the ADS would allow the vehicle to drive. Figure 4.7 shows all mentions for both questions, including their frequency.

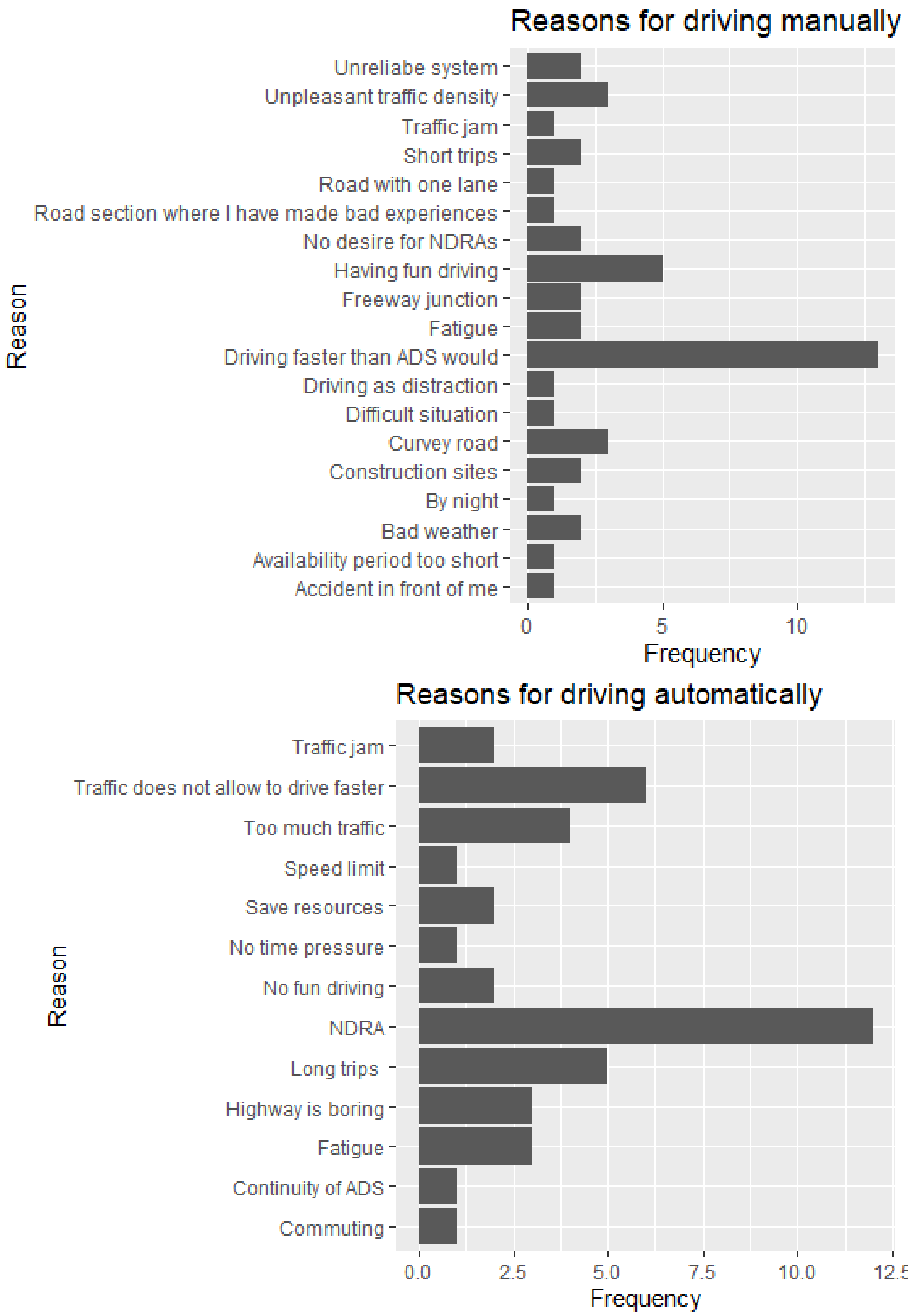


Figure 4.7: Reasons for driving manually and for driving automatically with frequency.

4.4 Discussion and Limitations

In this study, qualitative methods were used to find answers to the five research questions outlined above. Firstly, it was found that the expectations of laypersons regarding the availability of an ADS, including Level 3 automation, are too high. This was reflected in the statements of the experts as well as in the statements of a large part of the laypersons in the driving simulator study. According to the focus group discussants, the expectations of infinite availability times and availability times that are sufficient for the completion of an NDRA in any case are unrealistic. Future users are expected to face interruptions. To prevent frustration that can arise from unmet expectations (Ochs et al., 2008), it would be useful to provide users with information on how long the ADS will be available, both before and after the activation.

Other information needs identified in this study are a certainty indication regarding the estimation of the availability duration, suggestions for NDRAs that can be completed during the automated drive, an overview of all availability sections on any particular route, a time duration until the ADS will be available, and reasons for the non-availability of the ADS. Interestingly, the focus group interview had already revealed all these information needs. The participants of the driving simulator study did not add any information needs related to availability. According to the evaluation of the participants, the certainty indication as well as the suggestions on NDRAs are the least important information. Moreover, these two requirements also show the widest variance as can be seen in Figure 4.6. It can therefore be concluded that the greatest interindividual differences exist here.

In addition, the focus group interview showed that novices sometimes have difficulties accepting non-availability, which was not demonstrated in the driving simulator study. Nonetheless, no participant was able to correctly determine the reason for non-availability, although this was at least visually indicated on the route by the missing emergency lane. This confirms the relevance of the provision of information on reasons for the ADS not being available, especially as this could lead to an increased perceived system understanding, similar to the presentation of reasons for RtIs (Körber et al., 2018).

As stated in the theoretical background of this thesis, one of the greatest benefit of automated driving is the possibility to pursue NDRAs (Pettersson & Karlsson, 2015; Pfleging et al., 2016). This is also reflected in the statements of the driving simulator study participants, most of whom reported that pursuing an NDRA would be their main reason to activate the ADS.

Since this study involved qualitative approaches, all statements are of course subjective, not quantifiable, and difficult to generalise. The quantitative rating of the single information needs and the difference between these ratings can also only be interpreted to a limited extent, especially since these differences were not subject to statistical testing. This was omitted because no hypotheses could be formulated in advance due to a lack of theoretical background for the research questions. Another limitation of this study is the lack of demographic data on the experts in the focus group interview. However, the specifics of the experts' experience were not investigated: it was not surveyed how many trips each expert had made with novices in AVs, how long these trips were, and how mature the technology was at the time of the trips as collection of additional data was not possible since anonymity could not be guaranteed with such a small number of participants all working for the same company.

Nevertheless, the results of the expert discussion were confirmed in the statements of the novices in the driving simulator study. They identified the same information needs, and in part, the needs postulated by the experts were rated as important and useful. Due to the multimethodological approach used for this study, it can be assumed that, despite the importance of interpreting these results with caution, the information needs gathered in this study are an important contribution to research related to the question of how future users can be supported in understanding the ADS.

5 Evaluation of Information Needs

The following subchapters present an evaluation of the information needs gathered in the previous exploratory study. Provided it was methodologically possible, several information needs were tested simultaneously in one study, while other information needs had to be evaluated separately. The evaluation of the information needs was carried out since their relevance for future users of an ADS could not be assessed solely based on the results of the previously described study due to the qualitative nature of their collection. Prior to each experiment, literature research was conducted to gather theoretical background knowledge on the respective requirements, if available, to incorporate it into the conceptualization of the studies and design solutions.

The information needs evaluated in this chapter can be found in Table 5.1. The order of presentation corresponds to the order of the subchapters of the corresponding studies.

Information needs	**Subchapter**
Estimated availability duration of the ADS before activation	5.1
Certainty of the estimation of the availability duration	5.2
Reasons for non-availability Duration until ADS becomes available	5.3
Overview of all availability periods on one’s route	5.4

Table 5.1: Information needs evaluated and corresponding subchapters.

5.1 Displaying the Estimated Availability Duration of the ADS Before Activation

This study focuses on displaying the estimated availability duration of an SAE Level 3 ADS (SAE, 2021), since that is expected to launch earlier than Level 4 systems and the problem statement from the theoretical background applies equally to both levels. In addition, it is easier for participants to only be required to understand the functionality of one ADS and not two different ones. The following gives a brief theoretical overview of the information need of the availability period before activation of the ADS, as well as explaining the methodology and study design, and presenting and discussing the results. This study has been previously published in Danner, Pfromm, and Bengler (2020), and for a more detailed depiction of the theoretical background and statistical analyses, the reader is referred to the publication.

5.1.1 Research Questions and Hypotheses

It can be assumed that transitions from automated to manual driving will, for the most part, not be triggered by critical and unforeseen situations but rather by predictable system limitations (Holländer & Pfleging, 2018). In many cases, these system limits will be determined by route characteristics and therefore giving an estimate of the duration of availability as in Holländer and Pfleging (2018) and Wandtner et al. (2018) will be possible. Since the possibility of conducting NDRAs is among the most important benefits of AVs (Danner, Pfromm, Limbacher, & Bengler, 2020), it can be assumed that information regarding the available amount of time to perform an NDRA will influence the choice of the activity and the behaviour during automated driving. For example, users might only perform short NDRAs on routes with short availability periods because they might not want to be interrupted. Or they might choose an NDRA that takes more time, accepting the possibility of an interruption. Laypersons in particular seem to expect periods of availability that are not considered realistic by experts, and thus might assume that they can perform an NDRA for any length of time (Danner, Pfromm, Limbacher, & Bengler, 2020). This leads to the question of what happens when users activate the ADS for executing a specific NDRA of a duration that can be estimated in advance, but they are interrupted in the middle of the activity and must resume driving themselves.

Since unmet expectations of a system can lead to frustration (Ochs et al., 2008), adjusting the users' expectations is one way to prevent this negative emotion. Therefore, this study postulated that displaying information on how long the ADS will be available before it is activated – and with this influencing the expectation – might help participants use the ADS more purposefully. The definition of purposefulness in this work follows the Theory of Purposeful Work Behavior [sic] (Barrick et al., 2013) in which it is stated that purposeful behaviour is directed towards achieving a goal. In the case of using an ADS, a behaviour would be purposeful if the activation of the ADS were connected to a specific goal that can be realistically achieved. Hence, purposefulness served as dependent variable in this study. See Chapter 5.1.2.3 for the operationalisation of the same.

NDRAs are among the main reasons to use an ADS, and therefore it is assumed in this study that the successful completion of planned activities might positively influence the perceived usefulness and user satisfaction positively, both of which are components of acceptance (Adell, 2009), which is one further dependent variable tested in this study. As outlined in the general methods section (Chapter 3.3), usability and workload are important constructs to be considered when developing an HMI concept, which is why they served as dependent variables as well. It was assumed that the additional information and thus increased transparency might improve usability and that the knowledge about the possibility of which NDRA can be performed in the upcoming availability period without being interrupted might influence the workload. Since it must be considered that workload could also be impaired by providing additional information visually, the hypothesis regarding this dependent variable was non-directed.

The evaluation of the system and the purposefulness of the activation behaviour could be influenced by the length of the availability periods and the possibility to complete NDRAs. This consideration is taken into account by adding an independent variable regarding different availability durations. Two conditions were defined for this study: a positive condition in which more than half of the availability periods were sufficient for the successful completion of an NDRA and a negative condition in which less than half were sufficient.

As unsuccessful interactions can lead to negative feelings, which in turn can negatively influence future interactions (Brave & Nass, 2007), it is not only relevant to find out what information users need, but also how this information should be conveyed. As part of the preparation for this study, a Design Thinking Workshop was conducted to obtain concepts for the manner of communicating information on availability. The results of this workshop showed that it might be helpful to have the system refer to itself in the first-person singular ("I"), inducing anthropomorphism when interacting with the user. This is supported by the insight that emotions are not only formed by expectations but also by the way a system communicates (Brave & Nass, 2007). Greetings and farewells also seem to have a positive influence on the users' satisfaction with the interaction (Higashinaka et al., 2008). Hence, purely factual communication was compared to a more personal one. The differences in communicating information can be found in Chapter 5.1.2.2.

Based on the above considerations, the following hypotheses were formulated:

Hypothesis 1a (H1a). Acceptance is higher when the availability duration before activation is displayed.

Hypothesis 1b (H1b). Acceptance differs depending on how information on the automation is displayed.

Hypothesis 2a (H2a). Usability is higher when the availability duration before activation is displayed.

Hypothesis 2b (H2b). Usability differs depending on how information on the automation is displayed.

Hypothesis 3a (H3a). Workload differs depending on whether the availability duration before activation is displayed or not.

Hypothesis 3b (H3b). Workload differs depending on how information on the automation is displayed.

Hypothesis 4 (H4). Purposefulness of the activation behaviour differs depending on whether the availability duration before activation is displayed or not.

5.1.2 Methods

5.1.2.1 Study Design and Procedure

A mixed-model study design was chosen to test the hypotheses. Two independent variables were manipulated, one as a between-subject factor and the other as within-subject factor. The display of information regarding the availability of the ADS served as the three-level within-subject factor. A baseline concept (BL) was compared with two different advanced concepts. In one of them, the availability before activation was factually displayed (time before activation: TB) and in the other one the same information was presented in the form of a personal approach (time before activation plus personal approach: TBP). The between-subject factor consisted of the positive and negative condition, which differed in the durations of the availability periods. Figure 5.1 shows the study design.

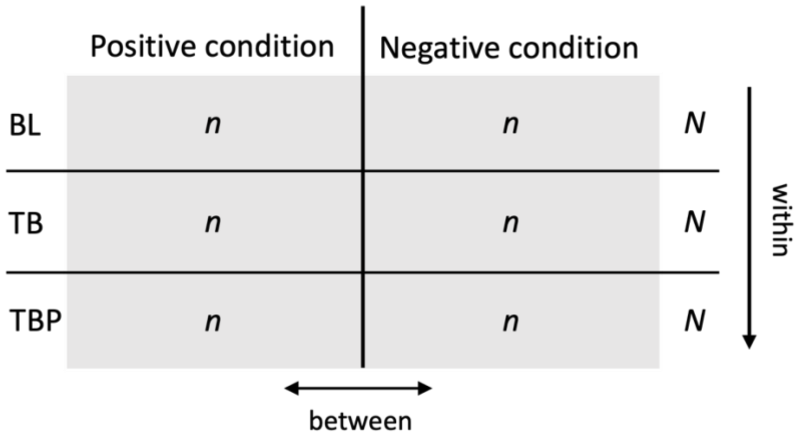


Figure 5.1: Study design with between and within-subject factors (Danner, Pfromm, & Bengler, 2020).

After participants were welcomed and informed about the study and possible side effects of riding in the simulator and their right to discontinue, they gave written consent to participate. Subsequently, they completed a demographic questionnaire and experienced a drive in the driving simulator, during which they learned about manual driving and became familiar with operating the ADS, which was activated and deactivated by pressing a button on the steering wheel. They experienced transitions from manual to automated driving and vice versa. They also familiarised themselves with how to operate the tablet computer, which was mounted in the centre console, simulating the CID. During the experiment, the NDRAs in the form of videos of various lengths were executed on this device. After the familiarisation, participants conducted the experimental drives in which they experienced all three concepts in randomised order. After each drive, they completed the acceptance, usability, and workload questionnaires. Each of these three drives lasted approximately 15 minutes, and, regardless of the condition the participants were assigned to, they experienced six availability periods on each drive. In the positive condition, four of these periods were sufficiently long to successfully complete an NDRA, whereas in the negative condition only two were sufficient. The videos were divided into long and short videos. Long videos lasted about 90 seconds and short videos about 50 seconds. Videos were ordered on the screen of the tablet according to the category to which they were assigned (short vs. long). Short videos were thus fixed on the left side of the screen and long videos on the right side. A sign above the tablet indicated which page corresponded to which video category.

The availability periods were defined to be either long enough for a long or a short video or not long enough for any NDRA. Participants were instructed to finish watching as many videos as possible without interruption during the automated drive. However, they were not instructed to activate the ADS whenever possible, making it the participants' decision whether to activate it or not. To ensure that the participants adhered to the study design, they were told that an intentional interruption of videos was against the rules.

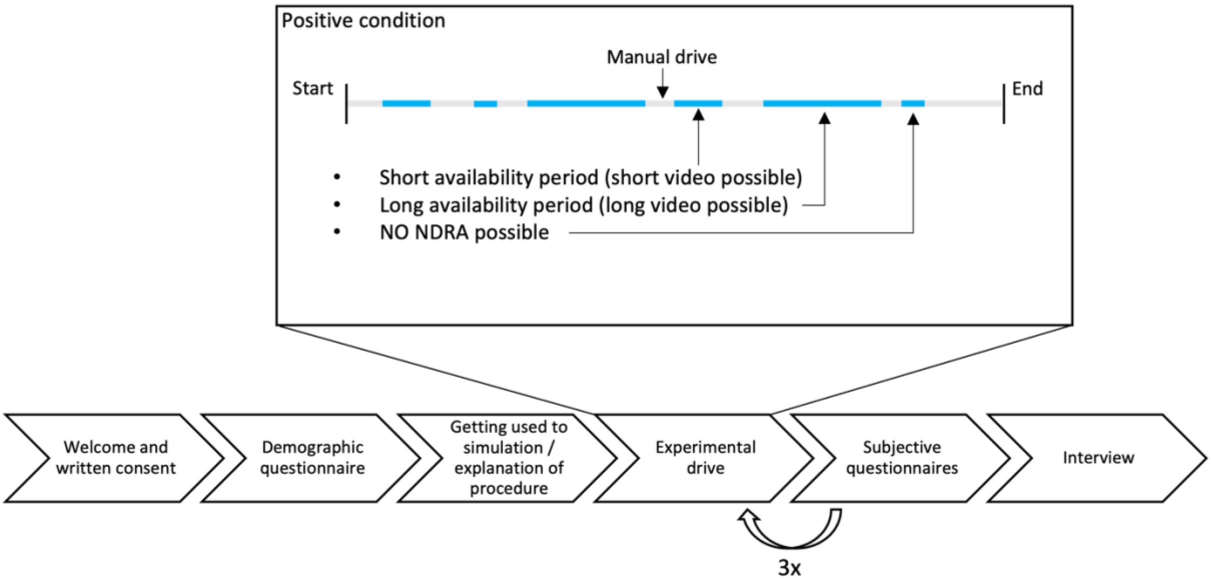


Figure 5.2: Procedure of the study with an exemplary test drive. Periods without available automation are grey; periods with available automation are blue (Danner, Pfromm, & Bengler, 2020).

To avoid sequential effects, the order of the different availability durations was randomised. Figure 5.2 shows the study procedure and an exemplary experimental drive along with the availability durations. After finishing all drives and filling out the last questionnaires, a qualitative semi-structured interview was conducted. The study took place in the same driving simulator as the exploratory study with novices. For further information on the driving simulator please refer to Chapter 4.2 or to Danner, Pfromm, and Bengler (2020).

5.1.2.2 Human-Machine Interface

In the BL concept, the usual information was displayed as in conventional vehicles, including velocity and RPM. When the ADS became available, a pop-up with the text "Automation available" appeared for six seconds, paired with an acoustic signal and a permanent icon symbolising availability. In the TB concept, the same displays were used but the text in the pop-up was extended to include the availability duration: "Automation available for *x* minutes and *y* seconds". In the TBP concept, the text changed to: "Hey! I can drive you now for *x* minutes and *y* seconds.". As explained above, this concept was included because it was shown that a more personal approach of addressing users can have a positive influence on the satisfaction with an interaction. In both the TB and TBP concepts, a horizontal time bar along with a countdown indicating the availability duration was displayed next to the availability icon. All information was displayed in the IC.

The RtI was the same in both the BL and the TB concepts. Twenty seconds before a system limit was reached, another acoustic signal was presented together with a pop-up including the text: “Please take over now”. For the participants to know how much time they had left before the system limit was reached, the RtI was paired with a countdown. In the TBP concept, the RtI text was changed to: “Please take control again. I hope you had a nice ride with me, see you soon!”.

During the automated drive, all concepts were the same, showing surrounding vehicles, velocity, RPM, and the availability duration in form of a time bar and a countdown. Table 5.2 summarises the differences. The difference between TB and TPB was deliberately kept as small as possible and consisted only of differing formulations. Additional factors that would have increased the anthropomorphism of the system, such as an avatar or speech output, were not included, since possible effects could no longer be clearly traced back to their predictors. Also, research shows that text alone can reduce frustration (Hone, 2006).

	Baseline Concept (BL)	**Time before activation concept (TB)**	**Time before activation plus personal approach concept (TBP)**
ADS available	Pop-up + icon displaying availability	Pop-up + icon displaying availability + time bar and countdown displaying availability duration (factual approach)	Pop-up + icon displaying availability + time bar and countdown displaying availability duration (personal approach)
RtI	Request to intervene + countdown	Request to intervene + countdown	Request to intervene with different wording (personal approach) + countdown

Table 5.2: Different information displayed in the different HMI concepts.

5.1.2.3 Dependent Variables

This study made use of the questionnaires described in the general methods (see Chapter 3.3). The Acceptance Scale by van der Laan et al. (1997) was used for measuring acceptance, the SUS (Brooke, 1996) for measuring usability, and for investigating workload, the NASA-rTLX (Hart, 2006) was used.

To be able to assess the participants’ activation behaviour when the ADS became available, it was operationalised as either purposeful or non-purposeful following the definition given above. Since participants were told that their only aim was to watch and finish as many videos as possible without interruptions, activations were defined as purposeful when an NDRA could be finished within the availability period. Furthermore, the decision to not activate when no

NDRA was feasible was also defined as purposeful. To quantify the purposefulness of the activation behaviour a percentage value was calculated as follows:

$$Purposefulness = \frac{Quantity\ of\ purposeful\ decisions}{Quantitiy\ of\ availability\ periods} * 100$$

If a participant experiencing a drive in the positive condition purposefully activated the ADS three times and decided one time to purposefully not activate, the participant made four purposeful decisions. This is then divided by six (total number of availability periods) and multiplied by 100, leading to a purposefulness rating of 66.7%.

5.1.2.4 Participants

The study sample consisted of $N = 33$ participants, 19 of whom were men and 14 of whom were women. The average age was $M = 45.03$ ($SD = 14.50$) years and the average annual mileage was $M = 17{,}891$ ($SD = 18{,}371$). To be able to test the within-subject factor, the participants were divided into two groups. The positive condition consisted of 16 and the negative condition of 17 participants.

5.1.2.5 Statistical Analysis

The statistical analysis was conducted using R Studio. Since more than two groups were compared including between and within-subject factors, mixed ANOVA was chosen as statistical instrument. All dependent variables were considered suitable for parametric testing, since Likert scales are deemed to be interval-scaled, whereas single-item responses are considered to be ordinal (Carifio & Perla, 2008). If the assumption for an ANOVA of normal distribution was violated, the analysis was still conducted and interpreted, since ANOVA was shown robust against this violation (Blanca et al., 2017). For the subjective measures, a correction of the alpha level was conducted using the Bonferroni method, resulting in an alpha level of $\alpha = 0.017$.

5.1.3 Results

The ANOVA conducted to test the acceptance ratings became significant with a strong effect (Greenhouse–Geisser $F(1.23, 37.91) = 9.41$, $p < 0.01$, partial $\eta2 = 0.24$), but only with respect to the HMI concept factor and not for the condition ($F(1, 31) = 0.03$, $p = 0.86$). There was no significant interaction (Greenhouse–Geisser $F(1.23, 37.91) = 1.69$, $p = 0.20$). The post-hoc tests revealed significant differences between the BL and the TB concepts ($t(32) = 3.61$, $p_{bonf} < 0.01$) as well as between the BL and the TBP concepts ($t(32) = 2.79$, $p_{bonf} = 0.03$), in both cases the advanced concepts scored higher. No differences were found between TB and TBP ($t(32) = 1.12$, $p_{\text{bonf}} = 0.82$). Table 5.3 shows the descriptive statistics for all dependent variables.

The same pattern was found when testing usability: significant main effect for the concept (Greenhouse–Geisser $F(1.33, 41.19) = 15.06$, $p < 0.01$, partial $\eta 2 = 0.33$), no effect for the condition ($F(1, 31) = 0.002$, $p = 0.96$), and no interaction (Greenhouse-Geisser $F(1.33, 41.19) = 0.40$, $p = 0.58$). Post-hoc comparisons also showed significant differences between the BL concept and TB ($t(32) = 4.21$, $p_{\text{bonf}} < 0.01$) and between BL and TBP ($t(32) = 4.07$, $p_{\text{bonf}} < 0.001$), but no effect for the comparison between TB and TBP ($t(32) = 0.61$, $p_{\text{bonf}} = 1.00$).

For workload, a significant main effect was also found for the HMI concept ($F(2, 62) = 4.52$, $p = 0.015$, partial $\eta 2 = 0.13$) with no interaction ($F(2, 62) = 0.96$, $p = 0.39$) or effect for the condition ($F(1, 31) = 4.62$, $p = 0.04$). However, only one significant post-hoc comparison was found between BL and TPB ($t(32) = 2.84$, $p\text{bonf} = 0.02$), indicating higher workload for the BL concept. The comparison between BL and TB was not significant ($t(32) = 2.12$, $p_{\text{bonf}} = 0.13$), and neither was the comparison between TB and TBP ($t(32) = 0.82$, $p_{\text{bonf}} = 1.00$).

Due to a violation of prerequisites for an ANOVA, a Friedman-test was conducted to investigate the differences in purposefulness. This was significant for the comparison of HMI concepts (Chi-Square(2) = 41.85, $p < 0.01$), with post-hoc tests revealing significant differences between BL and TB ($t(32) = 6.69$, $p_{\text{bonf}} < 0.01$) and between BL and TBP ($t(32) = 8.58$, $p_{\text{bonf}} < 0.01$). In both cases, the scores of the advanced concepts were superior to those of the BL. No significant difference was found between TB and TBP ($t(32) = 1.87$, $p_{\text{bonf}} = 0.21$).

5.1.4 Discussion

Hypotheses 1A, 2A, 3A, and 4 could be confirmed while Hypotheses 1B, 2B, and 3B had to be rejected. The results of this study show that presenting information on the estimated availability duration before activating the ADS can improve acceptance, usability, and workload. For the aspect workload, the main effect was only significant for one comparison: BL vs. TBP. However, since there was no difference between TB and TBP and on a descriptive level, the workload was also lower for the TB concept than for BL, no direct conclusions can be drawn about the superiority of the TBP over TB. No significant differences depending on the way the information was presented could be found regarding the subjective measures.

The same pattern emerged for purposefulness. However, the comparison of the descriptive statistics between the positive and negative condition is interesting. In the positive condition, purposefulness would have been 66.67% when participants had activated the ADS every time possible (see calculation logic above), whereas in the negative condition, it would have only been 33.33% when activated at every opportunity. Table 5.3 shows that the average purposefulness for the BL concept was in both conditions below the respective values. Here, the difference between positive and negative was the largest. However, the values in the positive and negative conditions converged for TB and TBP concepts, suggesting that providing information about availability duration prior to activation influenced activation behaviour and participants not only activated more frequently, when an activity was feasible but also activated less frequently when availability duration was insufficient for any NDRA.

		Baseline Concept		Time before activation		Time before activation + personal approach	
Acceptance	Positive Condition	0.69 (0.94)	0.57 (1.04)	1.05 (0.50)	1.11 (0.46)	0.92 (0.52)	1.04 (0.46)
	Negative Condition	0.45 (1.14)		1.17 (0.43)		1.15 (0.39)	
Usability	Positive Condition	67.66 (22.37)	66.59 (21.00)	78.75 (11.97)	80.23 (11.72)	79.38 (15.69)	79.32 (14.57)
	Negative Condition	65.59 (20.26)		81.62 (11.66)		79.26 (13.91)	
Workload	Positive Condition	35.23 (20.27)	38.27 (20.22)	29.56 (15.41)	31.25 (15.83)	27.99 (17.17)	28.87 (15.57)
	Negative Condition	38.47 (19.74)		31.53 (15.41)		30.58 (16.48)	
Purposefulness of activation behavior	Positive Condition	58.34% (12.17)	44.44% (17.01)	76.04% (19.21)	74.24% (21.69)	84.38% (14.23)	81.82% (20.98)
	Negative Condition	31.37% (8.08)		72.55% (24.25)		79.41% (26.04)	

Table 5.3: *M*(*SD*) of the dependent variables.

5.2 Certainty of the Estimation of the Availability Duration

The study[1] described in this subchapter has previously been published (Hecht, Danner, et al., 2020) and is summarised below in abbreviated form, focusing on background, results, and discussion. For further details, please refer to the publication.

[1] The study was designed and conducted with the help of Jana Mühlhausen (2020) as part of her master's thesis.

5.2.1 Research Questions and Hypotheses

The requirement of a certainty indication of the estimation of the availability duration was gathered in the explorative study described in Chapter 4. It emerges from the idea that the availability duration that might be presented to users will always be an estimation such as the estimated time of arrival when using navigation systems.

The visualisation of uncertainty has already been studied in aviation (McGuirl & Sarter, 2006), recommender systems (Shani et al., 2013), and also with reference to automated driving. For example, in Stockert et al. (2015), the status of an ACC system was described by such an uncertainty indication, which led to an improvement in the response to critical events as well as in usability ratings, but without influencing acceptance. In another study, the reliability of the ADS was presented using certainty visualisations. Some of the participants indicated that this had an impact on their engagement in NDRAs while others stated that they did not notice any changes in the display and thus did not draw any conclusions from it (Large et al., 2017). In further studies with HMIs presenting information in a different way visually but with corresponding content, it was shown that certainty displays improve take-over performance and the ability of study participants to dedicate themselves to NDRAs (Helldin et al., 2013). Beller et al. (2013) found that certainty displays can increase acceptance and trust while Kunze et al. (2019) also found positive effects on trust and situation awareness but negative effects on workload as well as an impaired ability to execute a visual search task among participants.

It seems certainty displays can improve acceptance and usability and impair workload. Since their influence on these dependent variables has not yet been tested with reference to the estimated availability duration of an ADS, the study presented in this subchapter aims to close this gap. As future users might have difficulties accepting limited availability periods on highways, the influence of the driving context on frustration as a dimension of the workload index (Hart, 2006) was tested as additional factor. Therefore, driving context was used as two-factorial independent variable: highway vs. city. It was investigated if a certainty symbol was able to lead to an interaction with the driving context, since it was assumed that participants would be more likely to accept wrong estimations in the city compared to the highway because more unanticipated events can occur in urban areas. This way it could be explored, if the potential difference of the context can be reduced by providing additional information on the reliability of the availability estimation. The following non-directed hypotheses were formulated:

Hypothesis 1a (H1a). Acceptance differs depending on whether the certainty symbol regarding the estimated automated driving time is displayed or not.

Hypothesis 1b (H1b). Acceptance differs depending on the driving scenario (city vs. highway).

Hypothesis 2 (H2). Usability differs depending on whether the certainty symbol regarding the estimated automated driving time is displayed or not.

Hypothesis 3a (H3a). Participants' frustration differs depending on whether the certainty symbol regarding the estimated automated driving time is displayed or not.

Hypothesis 3b (H3b). Participants' frustration is higher in the highway context compared to the city context.

5.2.2 Methods

5.2.2.1 Study Design and Procedure

The same procedure regarding welcoming participants, familiarisation and experimental drives, and questionnaires after the drives was conducted as described in Chapter 5.1.2.1 with differences in how many drives participants experienced and their respective durations.

The study took place in a mock-up at the Chair of Ergonomics at the Technical University of Munich (see Figure 5.3) and involved the comparison of two HMI concepts, which differed only with respect to whether a certainty symbol was displayed or not. A mixed design was chosen, and thus, there was a within-subject factor and a between-subject factor. The former was the HMI concept, and the latter was the contextual condition comparing city vs. highway.

Even though the routes on the highway and in the city were different, the conditions regarding the availability of the ADS were the same. There were four availability periods of different lengths. In both conditions, leaps in the estimated availability durations occurred in three of them. For example, the availability estimation jumped from five minutes to two minutes. Most time leaps were defined as negative, which means that the availability duration was shortened, but also positive time leaps occurred after which the availability duration was increased.

Hence, participants experienced two drives, one with the baseline and one with the advanced concept and faced non-continuous availability of the ADS. The availability durations varied (three, four, and six minutes) and were only interrupted by short manual driving periods. Participants were instructed to conduct NDRAs on their phone or a tablet computer provided while driving automatically. One group experienced the city scenario in both drives and the other group the highway scenario.

After the test drives and questionnaires were completed, a semi-structured qualitative interview was conducted. Participants were asked whether they noticed the time leaps and if they found the additional information helpful.

Figure 5.3: Mock-up used for this study.

5.2.2.2 Human-Machine Interface

An HMI concept was created for this study through a preliminary survey to represent the certainty of the estimation. The visualisation was adapted from the reception bars of a cell phone and accompanied by an explanatory text ("confidence in time budget estimation"). In the advanced concept (see Figure 5.4), the uncertainty of the estimate was indicated, but not in the baseline concept. Besides that, there was no difference between the concepts. All information was displayed in the IC. After a leap in the estimation occurred, the new availability duration estimation was presented along with a yellow rectangle containing the text "time budget adjustment".

5.2.2.3 Dependent Variables

This study used the dependent variables described in the general methods section (Chapter 3.3). From the NASA-rTLX (Hart, 2006), however, only the dimension of *frustration* was evaluated, since it was assumed that incorrect estimations of the availability duration would have a particular influence on this factor.

5.2.2.4 Participants

The study sample consisted of $N = 32$ participants, 16 per between-subject factor. The average age was $M = 31.56$ ($SD = 7.16$) years, with 25 men and seven women. Thirteen of the participants (40.63%) had taken part in a driving simulator study before.

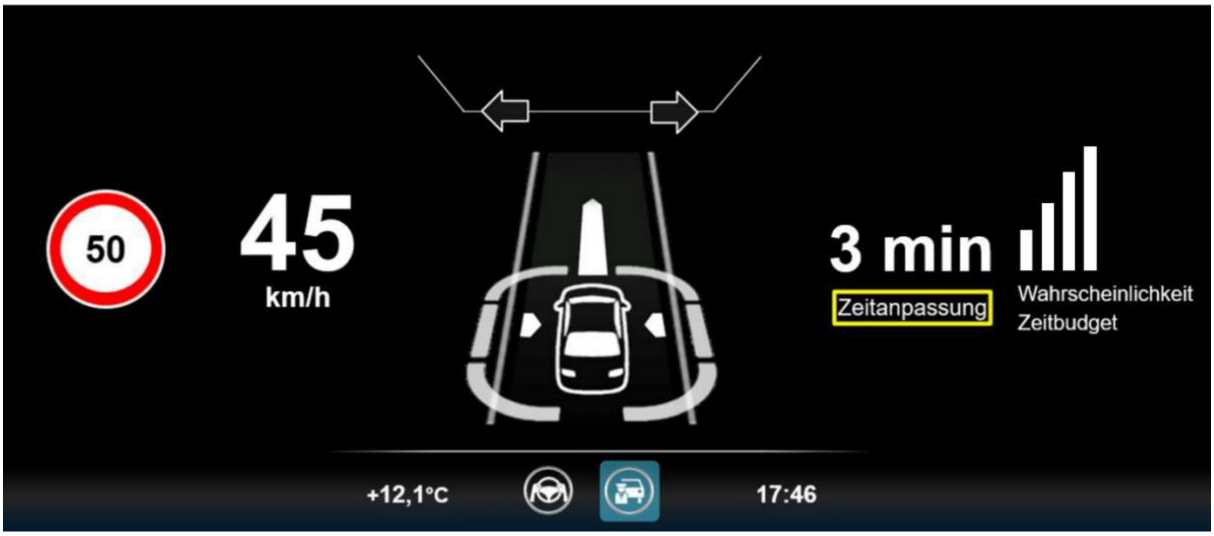


Figure 5.4: Advanced concept including an estimation of the availability duration and an indication of how reliable this estimation is. When a time leap occurred, this was further indicated with a yellow rectangle containing the text "time budget adjustment".

5.2.2.5 Statistical Analysis

IBM SPSS 24 was used for the statistical analysis. As in the previous subchapter, ANOVAs were calculated. For an explanation regarding the prerequisites of an analysis of variances and how these are handled in this thesis, please refer to Chapter 5.1.2.5.

5.2.3 Results

No significant effects regarding acceptance were found using ANOVA for either the within-subject factor HMI concept ($F(1, 30) = 0.00$, $p > 0.99$) nor the between-subject factor city vs. highway ($F(1, 30) = 1.23, p = 0.53$). No significant interaction could be found either ($F(1, 30) = 2.97, p = 0.10$).

Furthermore, no significant effect regarding usability could be found for the within-subject factor ($F(1, 30) = 0.35, p > 0.99$), nor for the between-subject factor ($F(1, 30) = 1.29, p = 0.55$), nor for the interaction ($F(1, 30) = 1.12, p = 0.29$).

Frustration was investigated using non-parametric tests revealing a significant effect showing that frustration was higher in the city condition ($z = -3.17$, $p = 0.01$, $n = 32$, $r = 0.56$), which is not in accordance with the assumption that frustration would be lower in the city condition. No effect regarding the HMI concepts ($z = -0.22$, $p > 0.99$, $n = 32$) could be found. Hence, no hypothesis formulated in Chapter 5.2.1 could be confirmed.

In the post-study interview, 65.63% of the participants stated to have preferred the baseline concept. Only half of them noticed the time leaps and extensions were mostly rated positively while reductions in the availability duration were mostly rated neutral. More than half of the participants – independent of whether they have noticed the time leaps or not – did not feel supported in their NDRA planning by the additional displays in the advanced concept. When taking the subset of participants into account who noticed the time leaps, only 18.75% preferred the advanced concept. Table 5.4 shows the descriptive values for each condition and dependent variable.

		Baseline Concept		**Advanced Concept**	
Acceptance	Highway	0.69 (0.94)	0.57 (1.04)	1.05 (0.50)	1.11 (0.46)
	City	0.45 (1.14)		1.17 (0.43)	
Usability	Highway	67.66 (22.37)	66.59 (21.00)	78.75 (11.97)	80.23 (11.72)
	City	65.59 (20.26)		81.62 (11.66)	
Frustration	Highway	5.56 (3.51)	6.97 (4.16)	5.18 (2.67)	5.91 (3.23)
	City	8.09 (4.44)		6.64 (3.44)	

Table 5.4: $M(SD)$ for dependent variables.

5.2.4 Discussion

The only significant result was the effect of the city vs. highway condition on frustration. As stated above, it was assumed that frustration would be lower in the city when time leaps occur since city traffic is more unforeseeable and therefore changes in availability durations were assumed to be more understandable for participants. Since the results show that frustration was higher in the city condition, the considerations above must be rejected. In addition, the absence of further effects, the fact that a large part of the study population did not even notice the time leaps, and of those who noticed, only a small proportion preferred the advanced concept, it can be concluded that the additional information tested in this study would not lead to any advantages for future users.

One further reason for absent effects could be the design of the certainty icon, but in the post-study interview, only two participants stated that they had problems understanding the display. Another limitation of this study was the short availability durations. The biggest time leap was from five to two minutes, which might not be enough for evoking frustration or a decrease in acceptance. There might be different reactions to time leaps in real usage situations when users conduct NDRAs that are subjectively considered important.

The limited benefit for participants might also result from the fact that the display did not give fully reliable information. If the icon did not indicate 100% certainty, participants could not be sure that no leaps were going to occur, and low certainty did not necessarily mean that the estimation of the availability duration was wrong.

5.3 Reasons for Non-Availability and Duration Until ADS Becomes Available

This study[2] was carried out to investigate whether a display of the reasons for non-availability of the ADS and of when it becomes available is beneficial for users. The study was published in Danner et al. (2021). For further details not described in this subchapter, please refer to the publication.

5.3.1 Research Questions and Hypotheses

Since the success of a layperson's interactions with an ADS is dependent on the correctness of the mental model of the system and these mental models are likely to be at least partly incorrect regarding the availability of an ADS (Carsten & Martens, 2018; Danner, Pfromm, Limbacher, & Bengler, 2020), the quality of the human-machine interaction might be impeded (Beggiato & Krems, 2013). This corresponds to the findings that drivers over-rely on their systems

[2] The study was designed and conducted with the help of Steidl (2020) as part of his master's thesis.

(Naujoks et al., 2016) because they do not fully understand how they work and what their limitations are (Walker et al., 2016). Existing mental models can change over time, either because newly received information update them or aspects of them be erased due to being not relevant for the user – for example, system limitations previously learnt can be erased after not being experienced for several rides (Beggiato & Krems, 2013; Stanton & Young, 2000). This might result in a discrepancy between the users' mental models and actual situations, which can lead to automation surprise (Carsten & Martens, 2018).

As outlined in the previous chapters, mental models can be influenced by giving information using HMI displays. Analogously to giving information on how long the ADS will be available even before the system is activated – which led to positive effects regarding acceptance (Danner, Pfromm, & Bengler, 2020) – information on when the ADS will be available can be displayed to give users the possibility to adapt their mental models. Furthermore, reasons for non-availability – like reasons for RtIs – might not be intuitively understandable and hence, it might be beneficial to display these since it could be shown that giving reasons for RtIs can improve the perceived system understanding (Körber et al., 2018).

Based on the outcome of the exploratory study investigating potential information needs regarding the availability of an ADS (see Chapter 4), this study aimed at exploring how displaying the duration until an ADS becomes available when it is not available as well as the reasons for non-availability affects the dependent variables described below. Potential interaction effects are exploratively tested along with the following non-directed hypotheses:

Hypothesis 1a (H1a). Perceived system understanding differs depending on whether information on when the ADS will be available is displayed.

Hypothesis 1b (H1b). Perceived system understanding differs depending on whether information on reasons for non-availability of the ADS is displayed.

Hypothesis 2a (H2a). Acceptance differs depending on whether information on when the ADS will be available is displayed.

Hypothesis 2b (H2b). Acceptance differs depending on whether information on reasons for non-availability of the ADS is displayed.

Hypothesis 3a (H3a). Usability differs depending on whether information on when the ADS will be available is displayed.

Hypothesis 3b (H3b). Usability differs depending on whether information on reasons for non-availability of the ADS is displayed.

Hypothesis 4a (H4a). Workload differs depending on whether information on when the ADS will be available is displayed.

Hypothesis 4b (H4b). Workload differs depending on whether information on reasons for non-availability of the ADS is displayed.

5.3.2 Methods

5.3.2.1 Study Design and Procedure

The same procedure regarding welcoming participants, familiarisation and experimental drives, and questionnaires after the drives was conducted as described in Chapter 5.1.2.1. The fix-based driving simulator of the Chair of Ergonomics at the Technical University of Munich was used for this study (see Figure 5.5)

A 2x3 mixed design was chosen for this study. The different HMI concepts served as the within-subject factor while the obviousness of the reasons for non-availability served as the between-factor. In one of the between-subject factor conditions, the reasons for non-availability were obvious for the participants and in the other one, the reasons were not directly observable on the road (obvious reasons *OR* vs non-obvious reasons *NOR*). Obvious reasons were for example construction sites along the route and missing lane markings. Non-obvious reasons were sensor errors and missing map data. Apart from these changes, the programmed routes did not differ between the two conditions.

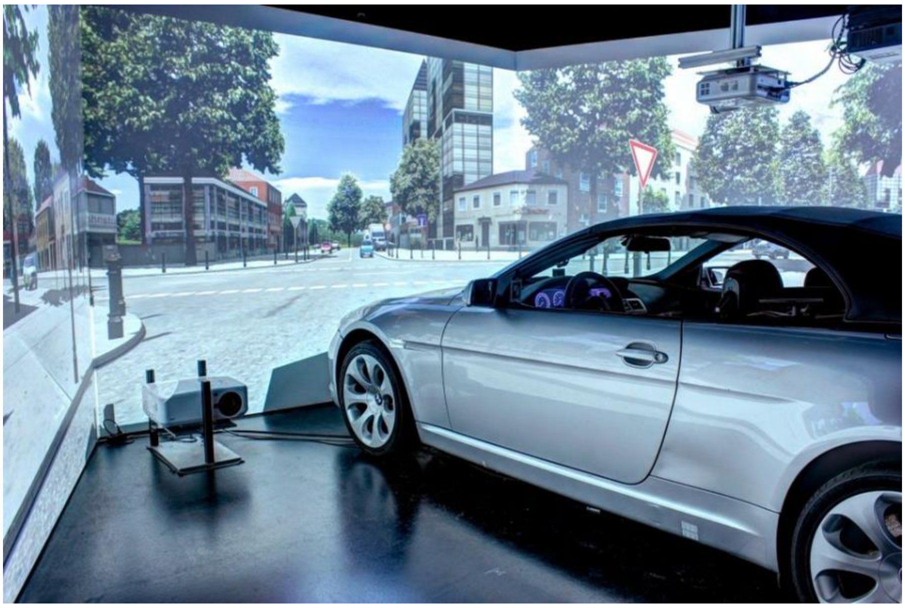


Figure 5.5: Fix-base driving simulator of the Chair of Ergonomics at Technical University of Munich used in this study.

Every participant experienced three rides – one per HMI concept – with a duration of about 8 minutes each. For the first five minutes, the ADS was not available – giving participants the chance to explore the displays regarding the non-available ADS. While driving manually, one half of the participants experienced obvious reasons for non-availability while the other half experienced non-obvious reasons. After five minutes, the ADS became available. Participants were instructed to perform an NDRA when driving automatically, which was a game on a tablet (the name of the game was “2048”). This game was chosen because it was easy to interrupt. After three minutes of driving automatically, a RtI occurred, and the participants had to leave the highway and stop the car at a parking lot. The questionnaires were filled in after each ride rating the experience with the respective HMI concept. The order in which the participants experienced the HMI concepts was randomised.

5.3.2.2 Human-Machine Interface

Three different HMI concepts were created to test the hypotheses. The concepts were presented in the IC. One baseline concept with no display regarding the two additional information (BC), one advanced concept giving information on when the ADS will be available (AC1), and another advanced concept giving information on when the ADS will be available as well as the reasons for non-availability (AC2) were compared. Only one reason for non-availability was permanently displayed in the two advanced concepts. If another reason existed, “+1 further” was displayed. When the user pressed the ADS activation button, a pop-up appeared giving all reasons for non-availability. Figure 5.6 shows AC2 with and without pop-up.

5.3.2.3 Dependent Variables

The dependent variables used in this study were acceptance, measured using the van der Laan Acceptance Scale (van der Laan et al., 1997); usability, measured using the SUS (Brooke, 1996); and workload investigated via the NASA-rTLX (Hart, 2006). For further descriptions of these questionnaires, please refer to the general methods section of this thesis (see Chapter 3.3).

Furthermore, the perceived system understanding was investigated via two statements, which the participants were asked to rate on a five-point Likert scale ranging from 1 to 5 (“I asked myself when the automation will be available again.”; “It was clear at any point of time why the automation was not available.”).

5.3.2.4 Participants

The study was carried out with $N = 34$ participants with an average age of $M = 30.59$ ($SD = 6.96$) years. 21 participants were male, and 13 participants were female (38%). Ten of the participants (29%) had already taken part in a driving simulator study and eight (24%) had already taken part in a driving simulator study concerning automated driving.

Figure 5.6: The picture above shows the HMI display for AC2 with the ADS being non-available. The duration until the ADS became available is in the top left corner with the following text: "Automation available in 5min". In the top right corner is an icon indicating the reason for non-availability along with the following text: "Impediment due to roadworks". AC1 was the same only without the reason for non-availability. BL was the same as AC1 but without the time until availability. The bottom picture shows the pop-up in AC2 when the activation button was pressed.

5.3.2.5 Statistical Analysis

To answer the research questions, mixed ANOVAs were conducted, provided that the assumptions for parametric testing were not violated. If the prerequisite of normal distribution was not fulfilled, the ANOVA was still performed, since this method is considered robust against this violation (Blanca et al., 2017). The alternative for ANOVA was the Friedman-Test. The statistical analysis was conducted using R. Two questions regarding the perceived transparency or system understanding were asked. To test the participants' answers, medians were calculated, and a Friedman-Test was conducted for each question, since single item Likert scales cannot be considered interval scaled. As the statements were not part of a standardised questionnaire, no overall transparency score could be built, and statistical analyses were carried out for the two answer scores separately.

5.3.3 Results

The perceived system understanding was measured using two statements. The first statement referred to whether participants understood when the ADS would be available, the second to whether the participants understood why the ADS was not available. The Friedman-Test for the first statement was significant (Chi-Square(2) = 12.13; $p < 0.01$; Kendall's $W = 0.29$) and post-hoc tests revealed significant differences between BC and AC1 ($p_{holm} = 0.01$; $r = 0.50$) as well as between BC and AC2 ($p_{holm} < 0.01$; $r = 0.63$), but no difference between AC1 and AC2 ($p_{holm} = 0.25$). This is in accordance with the hypothesis that giving information on the remaining time in the system status of unavailability helps participants understand when they can use the ADS. The Friedman-Test for the second statement was also significant (Chi-Square(2) = 19.61; $p < 0.01$; Kendall's $W = 0.54$), with significant post-hoc tests for the comparison of BC and AC2 ($p_{holm} < 0.01$; $r = 0.70$) and for the comparison of AC1 and AC2 ($p_{holm} = 0.02$; $r = 0.47$). These results show that giving information on why the ADS is not available can lead to an improved perceived transparency. The significant difference between AC1 and AC2, indicating more understanding when participants experienced AC2, shows that it is not only the fact of additional information leading to better ratings, but that the content of the display addressed the issue of participants not knowing why they cannot use the ADS.

A mixed ANOVA was conducted to test the hypotheses on acceptance. No significant results could be found regarding the HMI concepts ($F(1, 31) = 0.60$, $p = 0.44$), the conditions (Greenhouse–Geisser $F(1.50, 46.63) = 2.57$, $p = 0.10$) or the interaction effect (Greenhouse-Geisser $F(1.50, 46.63) = 0.94$, $p = 0.38$). Nonetheless, independent of the condition, there was a tendency on a descriptive level indicating better ratings for AC1than for AC2, and better ratings for AC2 than for BC. Descriptive data can be found in Table 5.5.

For testing the usability hypotheses, a mixed ANOVA was calculated leading to no significant results regarding the HMI concepts (Greenhouse–Geisser $F(1.4, 43.8) = 0.392$; $p = 0.65$), the conditions ($F(1, 31) = 3.06$; $p = 0.09$) or the interaction effect (Greenhouse-Geisser $F(1.4, 43.8) = 0.446$; $p = 0.56$). The descriptive data show a different pattern than it does for acceptance. The mean ratings for usability independent of the condition were highest for BC, followed by AC1.

The mixed ANOVA for testing the dependent variable workload did also lead to no significant effect regarding the HMI concepts ($F(2, 62) = 0.27$; $p = 0.76$), the conditions ($F(1, 31) = 0.01$; $p = 0.94$) or the interaction ($F(2, 62) = 1.44$; $p = 0.25$). On a descriptive level, the mean ratings show only small differences (see Table 5.5).

Besides filling in the questionnaires, the participants were also asked to comment on the different concepts to gather not only quantitative but also qualitative information. Eight participants stated they would have wished to know when the ADS will be available after experiencing BC and four would have wanted to know the reasons for non-availability after the ride with AC1. Four participants said that everything was clear after experiencing BC while five stated the same after using AC1. Also, five participants said that everything was clear after the usage of AC2. After the three rides, participants were asked to rank the concepts. For getting ranked first, a concept was given three points, for getting ranked second two points, and for getting ranked third one point. On average, AC1 was ranked highest, closely followed by AC2 (78 vs. 75 points) while BC received the worst ranking (45 points).

	Baseline Concept	**Advanced Concept 1**	**Advanced Concept 2**
Perceived system understanding 1 (when will the ADS become available)	$Mdn = 4$	$Mdn = 2$	$Mdn = 2$
Perceived system understanding 2 (why is the ADS not available)	$Mdn = 3$	$Mdn = 4$	$Mdn = 5$
Acceptance	$M = 1.07$, $SD = 0.63$	$M = 1.30$, $SD = 0.50$	$M = 1.17$, $SD = 0.65$
Usability	$M = 82.35$ $SD = 9.86$	$M = 81.82$, $SD = 18.31$	$M = 79.62$, $SD = 17.40$
Workload	$M = 26.59$ $SD = 13.03$	$M = 26.86$, $SD = 11.67$	$M = 26.30$, $SD = 14.73$

Table 5.5: Medians, means and standard deviations for the dependent variables. Higher values for perceived system understanding 1 indicate less understanding while higher values for perceived system understanding 2 indicate more understanding.

5.3.4 Discussion

Only the hypotheses regarding perceived system understanding were significant, indicating that additional information – when displayed understandably – can increase system transparency, which is associated with trust in automation (Kraus et al., 2020). Since no safety-critical HMI displays were tested in this study and the definition of trust includes a vulnerability aspect (Lee & See, 2004), no data on trust were analysed. Future studies should investigate how this kind of information can influence trust ratings and how perceived system understanding and trust are connected.

Although the additional displays aimed to increase the perceived usefulness of the ADS, no significant effects on acceptance and usability were found – there was only a descriptive tendency towards better acceptance when using AC1 and AC2. Acceptance, and especially usefulness, are influenced by the aims users have and by the belief that a system can help the users to achieve those aims. In this study, participants played a game on a tablet computer which might have been not enough motivation to use the ADS purposefully. Future research should investigate the effects of these information displays in a study setting where participants can choose NDRAs that are important or interesting to them. In this study, a standardised approach was chosen to find initial answers to the question of whether the displays enhance the system understanding, which was the case. The usability was not improved which might be due to the similar visual appearance of the different concepts except for the additional information. The permanent presentation of reasons for non-availability might have negatively affected usability since it is not necessary to see this information during the whole manual drive. Also, there was no effect on workload, which means that displaying more information did not increase the necessary mental effort when interacting with the concepts.

5.4 Overview of All Availability Periods on the Route

In addition to displaying the time until an ADS will be available – as discussed in the previous subchapter – it might be beneficial for users to know all availability sections on their route as well as their durations. This information need is based on the findings presented in Chapter 4. The following study was previously published in Danner et al. (2023).

5.4.1 Research Questions and Hypotheses

The possibility of conducting NDRAs is one of the biggest advantages of ADS (Danner, Pfromm, Limbacher, & Bengler, 2020; König & Neumayr, 2017) and therefore it is important to enable users to perform these activities as satisfactorily as possible. According to Hecht, Darlagiannis, and Bengler (2020), different activities can take up different amounts of time and therefore it is not possible make general statements on how long an availability period needs to be to allow users to conduct NDRAs. Displaying all route sections on which the ADS is available aims at increasing system transparency and decreasing automation surprise (Carsten & Martens, 2018), as this would enable users to choose NDRAs matching the availability duration or start activities that can easily be interrupted.

Displaying information on all availability sections might help participants plan and perform their NDRAs throughout the whole route, which could improve the usefulness and ease of use of the system, both being dimensions of acceptance (Davis, 1989). Usability might also be positively influenced by this additional information, since displaying the estimated availability duration of the current section before and after activation was shown to improve the same (Danner, Pfromm, & Bengler, 2020; Holländer & Pfleging, 2018). As it was shown that any discrepancy between users' expectations and reality can increase mental workload – since incorrect mental models have to be adapted to information perceived from the system and the environment (Beggiato & Krems, 2013) – workload might also be improved by providing users with information on availability to adjust their expectations. Incorrect assumptions about the system might even prevent usage entirely (Josten et al., 2018), which makes it relevant to help users form correct assumptions by the means of an HMI that increases system transparency. Since it was assumed that providing information on all availability sections can increase users' ability to plan and conduct NDRAs, task performance in NDRAs was chosen to be one dependent variable in this study. Moreover, the dependent variables described in Chapter 3.3 were tested and hence, the following hypotheses were formulated for this study:

Hypothesis 1 (H1). Acceptance of the ADS is higher when the availability sections of the ADS are displayed.

Hypothesis 2 (H2). Usability of the HMI of the ADS is higher when the availability sections of the ADS are displayed.

Hypothesis 3 (H3). Workload is lower when the availability sections of the ADS are displayed.

Hypothesis 4 (H4). Participants achieve higher scores in NDRA performance when the availability sections of the ADS are displayed.

Further explorative investigation pertained to whether the way of presenting the information influences the results.

5.4.2 Methods

5.4.2.1 Study Design and Procedure

The study consisted of filling in a demographic questionnaire, experiencing a test drive in the driving simulator at AUDI AG (see description of the simulator in Chapter 3), experiencing the experimental drive, and filling in questionnaires after using the concepts. Participants were asked to drive according to the traffic rules when the ADS was not available and to activate the system when possible. Furthermore, the participants were instructed to conduct NDRAs on the tablet computer to gain points (see Figure 5.7) but that it was against the rules to conduct NDRAs while driving manually. The logic of how points were counted was explained to them. Participants were told to conduct the tasks they wanted to do in an order of their choice. They were also made aware that unfinished activities would lead to no points. Therefore, their main aim regarding the NDRAs was to collect as many points as possible without breaking the rules. A between-subject design was chosen with the HMI concepts serving as a factor, three concepts were compared. Hence, every participant only experienced one HMI concept. This way, it was possible to test whether the different HMI concepts led to different outcomes regarding the dependent variables.

Hence, every participant experienced one experimental drive that took about 52 minutes with 34 minutes of available automation and only short interruptions of about three minutes between the sections where automated driving was possible.

5.4.2.2 Human-Machine Interface

Three different concepts were created for this study: one baseline concept (BC) without any information on the distribution of the availability sections; one concept where this information was displayed using an abstract time bar (AC); and one concept where this information was integrated into the navigation map (MC). All concepts contained a map in the CID as well as information on the duration of the current (non-)availability section in the IC. Hence, the additional information on the distribution of the availability sections in the CID served as a preview of upcoming possibilities to conduct NDRAs while participants knew the duration of the current (non-)availability section in all conditions. Figure 5.7 shows the different HMI concepts.

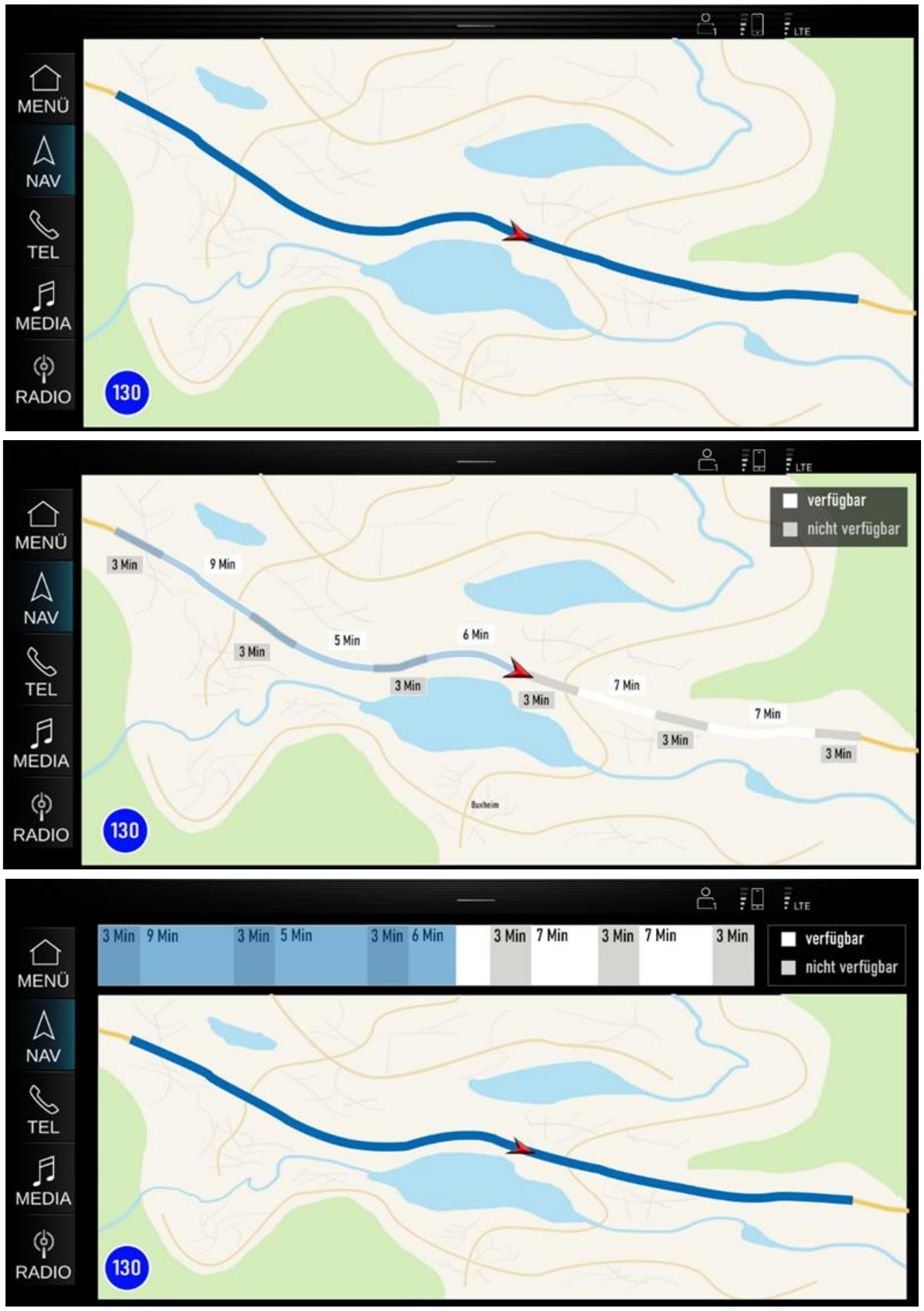


Figure 5.7: From top to bottom: BC, AC and MC. AC and MC providing information on where and how long the ADS will be available. White indicated availability and grey indicates non-availability.

5.4.2.3 Dependent Variables

The dependent variables used in this study were again acceptance, measured using the van der Laan Acceptance Scale (van der Laan et al., 1997), usability, measured using the SUS (Brooke, 1996), and workload, measured using the NASA-rTLX (Hart, 2006). For further descriptions of these questionnaires, please refer to the general methods section of this thesis (Chapter 3.3). Furthermore, task performance in NDRAs was investigated to find out whether the presentation of the availability sections helped participants plan and engage in activities while driving automatically. For this purpose, an app was programmed using Python to simulate a variance of different NDRAs. Participants could choose different NDRAs of different durations from a start screen. The longer an NDRA would take, the more points could be earned by finishing it. Participants had the chance to watch videos, do a quiz, save a phone contact from a card, read emails and news, or check the weather. Participants had to answer questions after most of these tasks to test if they have read the presented text. Therefore, it was assured that the tasks were completed and not skipped. After one task was completed, a green tick appeared next to it. Figure 5.8 shows the start screen of this app.

Figure 5.8: Overview screen of the NDRA.

5.4.2.4 Participants

A total of $N = 54$ participants (30 of them male; 24 of them female) took part in the study, but one participant had to be excluded for technical reasons. The mean age was $M = 37.6$ years ($SD = 14.4$ years). The mean annual mileage was $M = 18.365$ km ($SD = 13.318.2$ km); forty-five (83%) of the participants had already taken part in one or more driving simulator studies.

5.4.2.5 Statistical Analysis

For testing directed hypotheses, *a priori* contrast analyses were calculated using R. This tool allows combined and straightforward comparisons between groups, while more statistical power is achieved than with an ANOVA (Field et al., 2013). The contrast weights were defined as orthogonal, as shown in Table 5.6. The first contrast is a combined comparison between BC and the concepts providing information on the availability sections and the second contrast compares group two and group three (i.e., AC versus MC). The critical alpha value tested against was set at 0.05.

	Baseline concept (BC)	Abstract time bar (AC)	Availability in the map (MC)
Contrast 1	-2	1	1
Contrast 2	0	-1	1

Table 5.6: Orthogonal contrast weights used for contrast analysis.

5.4.3 Results

The first contrast was significant for acceptance values, showing higher ratings for AC and MC ($p = 0.02$; $d = 0.58$). The second contrast was not significant ($p = 0.78$). For usability, no significant effects were found for the first contrast ($p = 0.33$), nor for the second contrast ($p = 0.09$). Also, neither the first contrast ($p = 0.49$) nor the second contrast ($p = 0.30$) was significant for the measures of workload. The descriptive values can be found in Table 5.7.

The first contrast was significant for task performance ($p < 0.01$; $d = 0.55$), indicating significantly more achieved points in AC and MC compared to BC. The second contrast for task performance was not significant ($p = 0.78$). Therefore, the way of presenting the availability sections did not influence the task performance. Hence, Hypotheses 1 and 4 could be confirmed.

Concept	BC	AC	MC	AC + MC
Acceptance	1.16 (0.41)	1.33 (0.40)	1.44 (0.32)	1.38 (0.36)
Usability	87.08 (8.84)	82.22 (14.55)	88.97 (10.94)	85.50 (13.18)
Workload	27.44 (14.52)	27.84 (14.43)	27.38 (14.69)	27.84 (14.43)
Task performance	225.72 (86.85)	262.78 (73.12)	275.82 (78.62)	269.11 (75.00)

Table 5.7: *M*(*SD*) for the different concepts and dependent variables.

5.4.4 Discussion

These results indicate that presenting the availability sections can lead to increased acceptance towards the ADS when it is not continuously available, independent of how the sections are displayed. As no significant differences in usability were found, it can be concluded that this additional information did not lead to any improvements in how well and easily participants were able to use the HMI displays. Further, this also shows that the two HMI concepts containing more information than the BL did not decrease the ease of use of the HMI. On a descriptive level, the AC received the worst ratings, but is still in the range of good usability, close to excellent according to Bangor et al. (2008). The same observations apply to the workload ratings. Additional displays could have led to more cognitive workload since participants had to process more information, but no significant differences were found. The task performance was significantly higher in AC and MC indicating that presenting the availability sections helped participants to plan and conduct their activities purposefully. No significant differences were found between AC and MC showing that the information itself and not the changed design led to increased ratings.

The NDRAs were chosen to resemble activities users of an ADS might conduct in a natural driving context. Still, the emails, for example, were dummy emails with no personal relevance for the study participants. None of these activities was tailored to the participants' needs and therefore the commitment to conducting them might have been different than had they been able to choose their own activities.

A further limitation of this study was the between-subject design. While it allowed longer driving sessions and enabled the participants to get used to the respective HMI concepts, choosing a within-subject design would have allowed the participants to directly compare the different concepts. Nevertheless, significant effects were found despite this limitation showing the impact of the different concepts.

6 Development and Evaluation of the Final HMI concept

After all information needs collected in the explorative study were transformed into information displays, their effects on participants' perception of the ADS and its HMI were evaluated (see Chapter 5). These singular information elements were then combined into one final HMI concept. A state-of-the-art HMI for manual and automated driving was used and enhanced with the additional information displays on availability that were generated in this thesis. In this chapter, the development of the final HMI concept is described, as well as the study in which it was evaluated.

6.1 The Final HMI concept

Only those information elements that were shown to be beneficial for participants were considered for the final concept. According to the results of the studies described in Chapter 5, these include the following:

- Estimated availability duration of the ADS before activation
- Reasons for non-availability
- Duration until ADS becomes available
- Overview of all availability periods on the route

The estimated availability duration of the ADS before activation was used for the final concept as it was used for the evaluation study described in Chapter 5.1, since no negative aspects were revealed. The only adjustment made on the content level was removing the time bar from the concept, since it conveyed the same information as the time display itself and required extra space and both pieces of information were displayed in the IC. Design-wise, the duration display was bigger and in the centre of the display (see Figure 6.1).

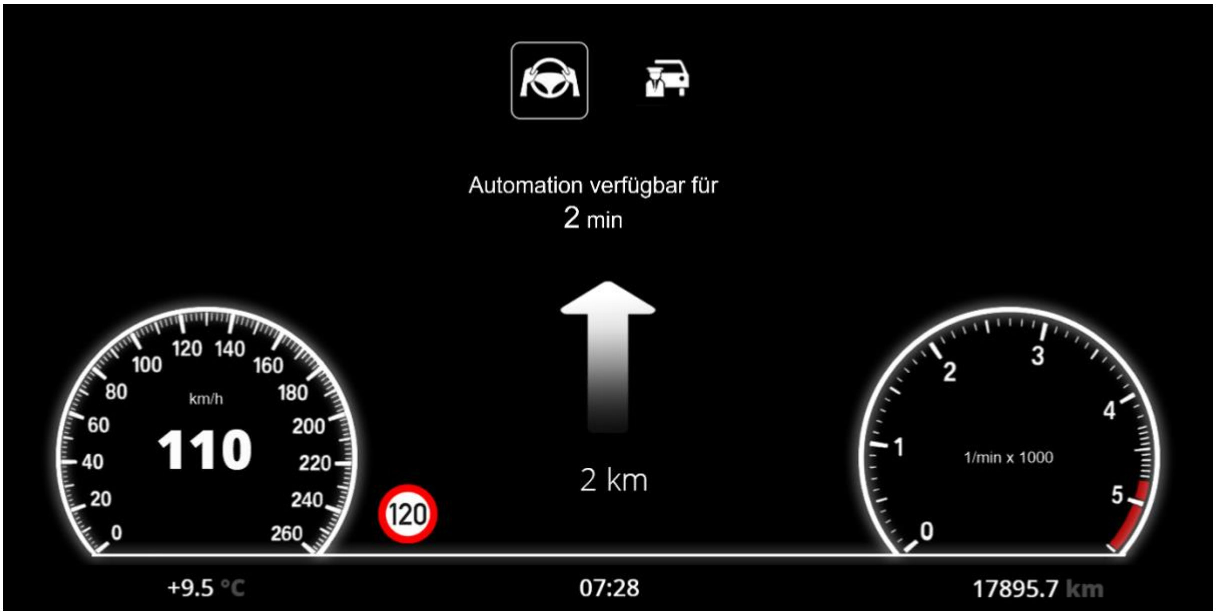


Figure 6.1: IC when ADS is available but not activated showing the availability duration: "Automation is available for 2 min". The right icon at the top of the display indicates the available automation.

The reasons for non-availability were integrated into the final concept in a similar visual way as in Chapter 5.3 but without being permanently visible in the IC. According to the results discussed, a permanent reminder of a system that is currently not working might not be beneficial for usability and acceptance. Furthermore, this extra display would make the IC appear too crowded. Therefore, the reasons for non-availability were only displayed in a pop-up window in the IC when participants tried to activate the ADS when it was not available (see Figure 6.2). Moreover, the reasons were also permanently presented in the CID along with the display of the availability sections (see Figure 6.3).

The duration until the ADS becomes available was displayed analogously to the information of how long the ADS will be available. The wording changed from "Automation available in *x* minutes" to "Automation not available for *x* minutes" since the first two words of the former formulation were "Automation available" which could irritate the users. The latter formulation contained a "not" which aimed at clarifying the system state. Furthermore, the icon symbolising automated driving was greyed out, indicating non-availability (see Figure 6.2).

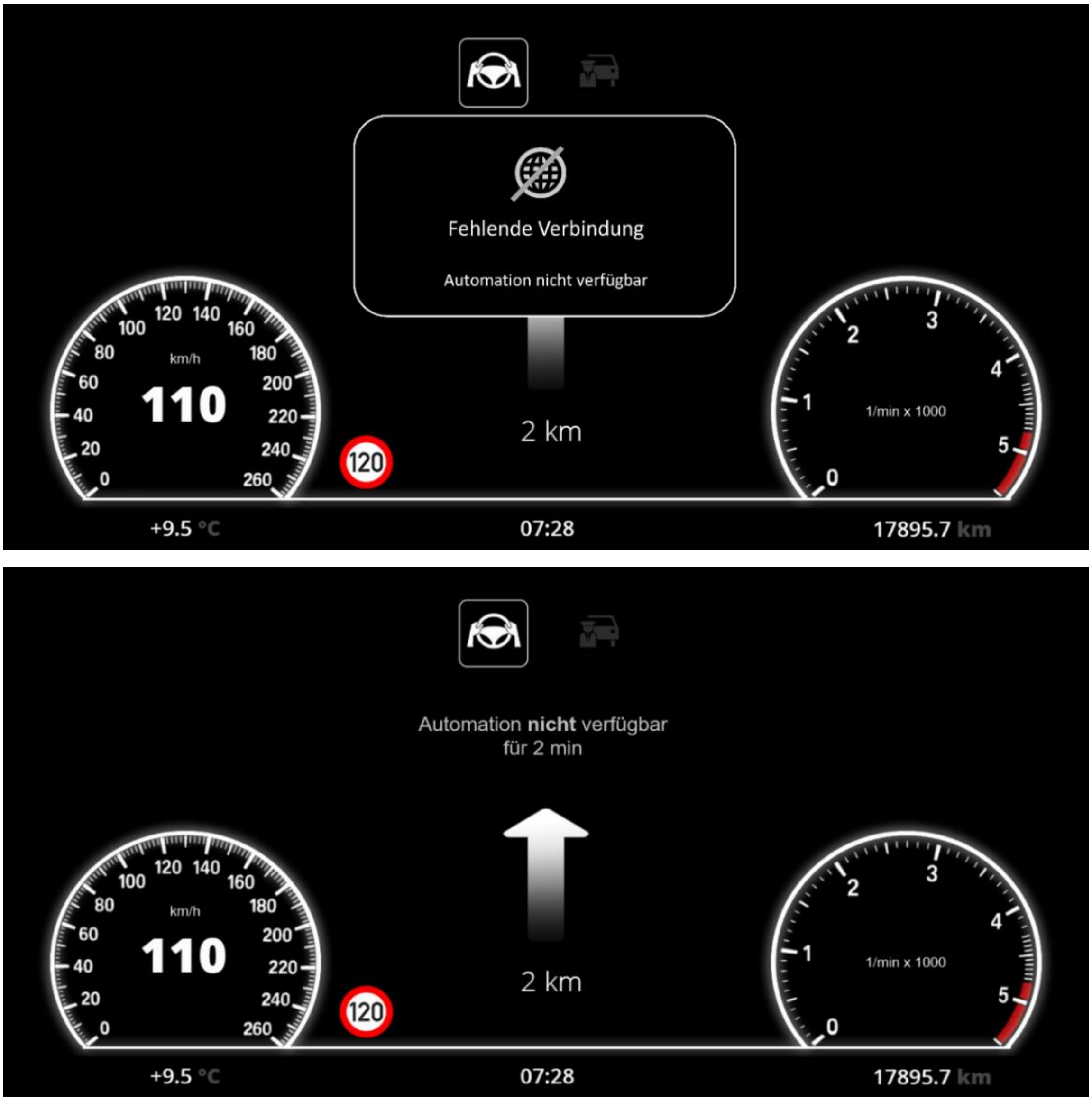


Figure 6.2: IC when ADS is not available showing the reason for non-availability after participant tries to activate the ADS: “Missing connection – Automation not available”. The second figure shows duration when ADS will become available: “Automation is not available for 2 min”. The right icon in the top of the display is grey, indicating non-availability.

The display of the availability sections of the ADS in the CID was built analogously to the abstract time bar concept in Chapter 5.4, since the map concept was not superior in the study. Furthermore, the navigation map display is already crowded with traffic and route information in current vehicles, which might make additionally displaying the availability sections lead to high complexity. Hence, it was decided to choose a design that focused solely on the ADS. As can be seen in Figure 6.3, the time bar was stylistically changed and an icon symbolising the car’s position within the bar was added to make it easier for participants to see how much time was left in the current system state. Reasons for non-availability were also added to the single sections using icons. Only the icons of the current section were displayed in white, the other ones (of previous and upcoming sections) were in grey to minimise confusion. When driving in a section where the ADS was available, an icon standing for a mobile phone and one standing for videos were shown in blue to symbolise the possibility to drive automatically and conduct NDRAs. The permanent icons explaining the non-availability were moved from the IC to the CID since these reasons are non-driving related information and hence might be too distracting in the IC.

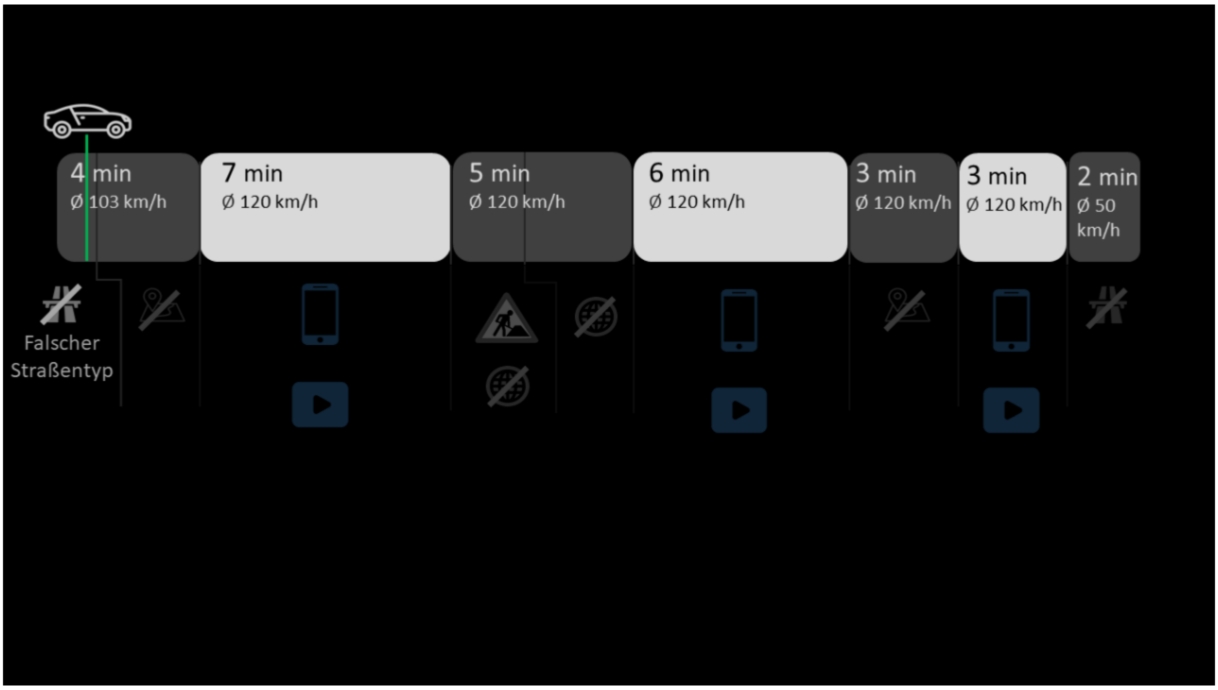


Figure 6.3: CID showing all availability and non-availability sections and the current reasons for non-availability.

6.2 Dependent Variables and Research Questions

Just like in the studies described in the previous sections, acceptance, usability, and workload were used as dependent variables. For further information regarding these aspects, please refer to Chapter 3.3. Since the Acceptance Scale (van der Laan et al., 1997) is not based on the TAM theoretically (Davis, 1989), but the TAM was used to explain acceptance in this thesis (see Chapter 2.2.2), another acceptance measurement – based on the TAM and the UTAUT (Venkatesh et al., 2003) – was used to validate the results of the Acceptance Scale. This additional measure was the Technology Usage Inventory (TUI) (Kothgassner et al., 2012), which was created in Austria and consists of eight dimensions: *curiosity*, *anxiety*, *interest*, *usability*, *immersion*, *usefulness*, *scepticism*, and *accessibility*. Moreover, it includes the scale *intention to use* (ITU). The dimensions *immersion* and *accessibility* were not investigated in this study as they are described as optional. *Immersion* is not a measure that can be used to describe an HMI concept but rather the driving simulation whose immersive quality was not subject of this work. *Accessibility* cannot be measured regarding ADSs of higher levels, since these systems are not available for purchase yet.

In addition to the standardised questionnaires, the subjective ability to plan the ride and the NDRAs was investigated using items that were created for this study: “I had all the information I would need to plan my non-driving related activities”. Perceived system understanding was investigated using the following items: “I always knew why the automation was not available” and “I was wondering when the automation will be available”. The items were answered on a Likert scale ranging from -2 (I do not agree at all) to 2 (I totally agree).

Furthermore, since the information displayed in this study was not considered safety critical and was therefore not crucial for the safe usage of an ADS but rather served the users as a comfort feature and should help improve the rating and understanding of the system, the subjective importance of the single information displays was evaluated using the Kano Model (Fundin & Nilsson, 2003; Hölzing, 2008; Löfgren & Witell, 2005; Ma et al., 2019). According to this model, the features of a system can be divided into different categories:

- Attractive quality: users do not expect this feature; its existence leads to positive reactions
- Must-be quality: users expect this feature; its non-existence leads to negative reactions
- Reverse quality: users are not interested in this feature; its non-existence leads to satisfaction

- One-dimensional quality: this feature can diminish unsatisfaction and increase satisfaction depending on the degree of manifestation
- Indifferent quality: its existence or non-existence has no influence

Participants rated every single HMI feature and assigned them to the categories described above. They answered two questions per feature: the *functional* question “How would you feel if this feature *xy* would exist?” and the *dysfunctional* question “How would you feel if this feature *xy* would not exist?”. Pictures of the respective HMI features were presented along with the questions to ensure every participant understood which element was referred to in the question. The answer options were:

1. I like it.
2. I expect it.
3. I am neutral.
4. I can tolerate it.
5. I dislike it.

Based on the combination of the answers to the *functional* and *dysfunctional* questions, the features were assigned to categories (see Table 6.1). Contradictory answers are not counted.

Functional	Dysfunctional	Quality
I like it	I dislike it	One-dimensional
I like it	I am neutral	Attractive
I like it	I can tolerate it	Attractive
I expect it	I dislike it	Must-be
I am neutral	I dislike it	Must-be
I am neutral	I am neutral	Indifferent
I dislike it	I expect it	Reverse
I can tolerate it	I like it	Reverse
I dislike it	I can tolerate it	Reverse

Table 6.1: Categorisation of the different product qualities according to Kano (1984).

Since the information needs that were used for creating the HMI were all shown to be superior to baseline concepts in previous studies, no HMI comparison was conducted in this study. The usefulness of the single information elements was already tested, and therefore, this aspect was not central to this study; instead, the focus was on whether the combination of the information pieces to one final concept still leads to ratings that are as satisfactory as in the previous studies. Therefore, the outcomes of the dependent variables were compared to threshold values. The threshold for the Acceptance Scale, which ranges from -2 to 2, was zero, and analysis involved testing whether the average rating in this study was superior to this neutral acceptance score. An average rating that is significantly above zero can be considered to be in a positive range of acceptance. Furthermore, confidence intervals were calculated to investigate the distance between the lower bound and the threshold. The same approach was chosen for perceived system understanding and the ability to plan NDRAs, since these items were also tested using a scale ranging from -2 to 2. The threshold for usability was chosen to be 72, since according to Bangor et al. (2008), good usability starts with this score. The threshold for workload was taken from the previously mentioned studies, intended to test whether the workload of the final concept was not significantly higher than the lowest workload in studies testing the single information needs. The threshold was an overall workload value of 26 (see Chapter 5.3.2) with scores ranging from 0 to 100. Hence, the hypotheses tested in this study were:

Hypothesis 1 (H1). The single acceptance scores for usefulness and satisfaction that can range from -2 to 2 are significantly higher than 0 as well as the overall score for acceptance.

Hypothesis 2 (H2). The usability score that can range from 0 to 100 is significantly higher than 72.

Hypothesis 3 (H3). The workload score that can range from 0 to 100 is not significantly higher than 26.

Hypothesis 4 (H4). The single scores for perceived system understanding items that can range from -2 to 2 are significantly higher than 0.

The categorisation of the HMI components into Kano categories was conducted to attain further insight into the importance of the single information displays. Moreover, the results of the investigated TUI dimensions were analysed on a descriptive level. Tests were done to determine whether the single TUI dimensions were correlated with the ITU score from the same questionnaire, as ITU is reported to be a crucial indicator for overall acceptance according to the TAM (Davis, 1989). Correlation tests were also conducted for the two factors of the van der Laan Acceptance Scale (van der Laan et al., 1997) and ITU of the TUI questionnaire to explore whether these two different acceptance measures show sufficient construct validity.

Besides these hypotheses, qualitative research questions were formulated. Especially the understandability of the HMI concept and how it might help with using ADSs and planning NDRAs was investigated by means of a short interview.

6.3 Methods

6.3.1 Study Procedure

To investigate the hypotheses, a driving simulator study was conducted, using the same simulator at the Chair of Ergonomics at the Technical University of Munich as in Chapter 5.3. Since the HMI concept was not compared to another one and no hypothesis regarding sample differences was formulated, the experiment consisted of only one group. Therefore, all study participants experienced the final HMI concept.

After giving written consent to take part in this study, the participants filled out a demographic questionnaire. Subsequently, the participants experienced an initial drive of about ten minutes to get used to the driving simulator and the operation of the automation. They were instructed to activate and deactivate the ADS using a button on the steering wheel. During this familiarisation drive, they saw a basic HMI concept in the CID. It looked the same as the final concept but without all the information regarding the availability of the ADS, except for the icon indicating whether the ADS was available or not.

Thereafter, the experimental drive began. Participants were instructed to drive according to the navigation in the CID and to use the ADS when possible. When the ADS was active, they were allowed to conduct NDRAs of their own choice. They could use their phones or a tablet computer that was available in the vehicle with various games and videos on it. No explanation regarding the HMI displays and the principle of CAD was given, in order to explore their ease to understand. Shortly after the first automated driving section, the participants were asked to stop at a highway service area. This short break was used for conducting a semi-structured qualitative interview to investigate whether the participants understood the different displays and the functionalities and limitations of the ADS. The following questions were asked:

- Which different bits of information were presented to you in the in-vehicle displays?
- How would you explain this system to a person you know who is not interested in technical details but would like to use it?
- Is there anything in the HMI concept that was not clear?
- Have you understood all the displays?

This conversation was followed by an explanation of the HMI concept, making sure that all participants had the same understanding of the concept before the second half of the study. Only the HMI display was explained, not the principles of CAD.

After the interview, the participants continued the experimental drive with the same instructions as before the first half of the drive. Subsequently, they answered questions in another qualitative interview to gather information on how well they understood the ADS and what they thought about the HMI concept. The following open questions were asked:

- How would you explain this system to a person you know who is not interested in technical details but would like to use it?
- Is there anything you would change about the concept?
- Is there anything you considered unnecessary about this concept?
- Is there anything you missed in this concept?

Moreover, the participants were asked to answer with ratings from one to five the following questions:

- How much would this HMI concept help you understand the automated driving system (one: not at all – five: very much)?
- How important do you rate the individual pieces of information (one: not important at all – five: very important)?

The last question was accompanied by a list of the presented information. The interview was followed by the standardised questionnaires described above.

6.3.2 Study Population

The study was conducted with $N = 30$ participants taking part. The mean age was $M = 28.03$ ($SD = 4.48$). 20 of the participants were men and ten were women. The average time that the participants had held a driving licence was $M = 10.45$ years ($SD = 4.31$) and the average distance travelled per year was $M = 6846.97$ km ($SD = 8523.12$). More than half of the participants ($n = 17$) have already participated in driving simulator studies in the past.

6.3.3 Statistical Analysis

One-sample t-Tests were conducted to test the hypotheses. In case the prerequisites were not met, the non-parametric alternative, the one-sample Wilcoxon-test, was used. The analysis was conducted using R-Studio.

6.4 Results

6.4.1 Quantitative Hypotheses Testing

The first hypothesis referred to the acceptance ratings and whether these were significantly higher than 0 and therefore in a positive range. This analysis was conducted for the two different dimensions acceptance consists of according to van der Laan et al. (1997) and for the total acceptance score.

The mean value was $M = 1.23$ ($SD = 0.34$) for the dimension of *usefulness* and $M = 1.22$ ($SD = 0.44$) for the dimension of *satisfaction*. The mean of the total acceptance based on all items of the questionnaire was $M = 1.22$ ($SD = 0.33$). Boxplots are depicted in Figure 6.4.

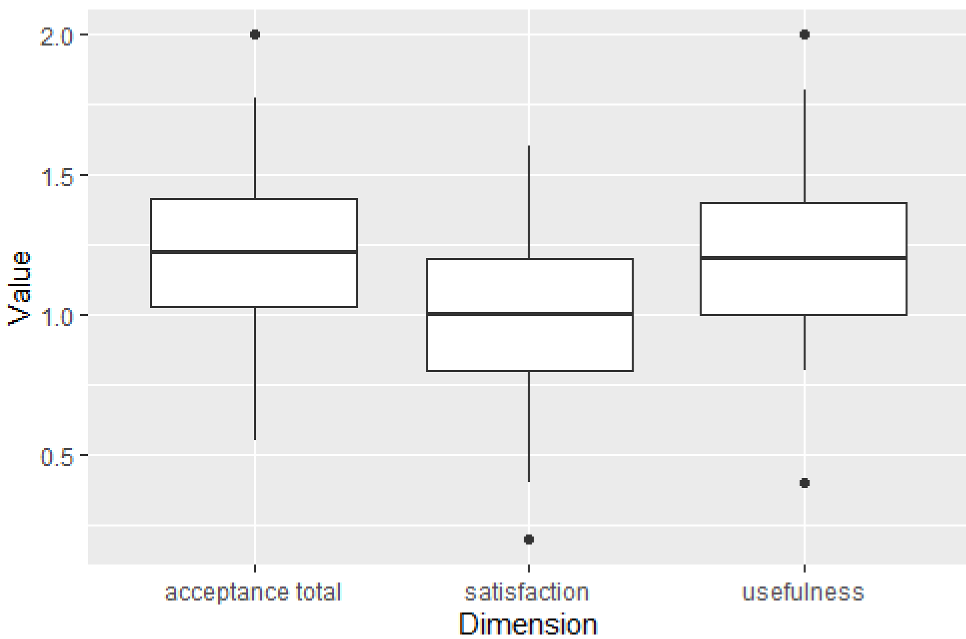


Figure 6.4: Boxplots of total acceptance and the two dimensions *satisfaction* and *usefulness*.

The assumption of normal distribution was tested for *usefulness*, *satisfaction*, and the total acceptance score using the Shapiro-Wilk-Test, which was not significant for all three dependent variables ($p > 0.05$). Hence, normal distribution could be assumed and one-sample t-Tests were calculated. All three one-tailed t-Tests were conducted with the hypothesis that the mean value of the respective dependent variable was significantly greater than 0 (while the range of the acceptance score is -2 to 2). The t-Test for *usefulness* was significant with a large effect ($t(29) = 19.80$, $p < 0.01$, $d = 3.62$), as were the t-Test for *satisfaction* ($t(29) = 15.18$, $p < 0.01$, $d = 2.77$) and the test for the total acceptance score ($t(29) = 20.44$, $p < 0.01$, $d = 3.73$). The 95% confidence intervals were calculated for the three acceptance scores to explore their distance from the neutral value of 0. The confidence interval for *usefulness* was 1.09 – 1.34, for *satisfaction* 1.03 – 1.36 and for the total acceptance score 1.08 – 1.33. Therefore, Hypothesis 1 could be confirmed.

Usability was tested analogously to acceptance. The mean value was $M = 89.00$ ($SD = 8.11$) and the Shapiro-Wilk-Test was significant ($p < 0.05$) and therefore normal distribution could not be assumed. Hence, the non-parametric alternative to a one-sample t-Test, the one-sample Wilcoxon-Signed-Rank-Test, was conducted with a reference value of 72 to test against. Figure 6.5 shows the boxplot depicting descriptive statistics of the usability ratings. The Wilcoxon-Signed-Rank-Test was significant ($p < 0.01$; $r = 0.87$) with a large effect and hence, Hypothesis 2 could be confirmed.

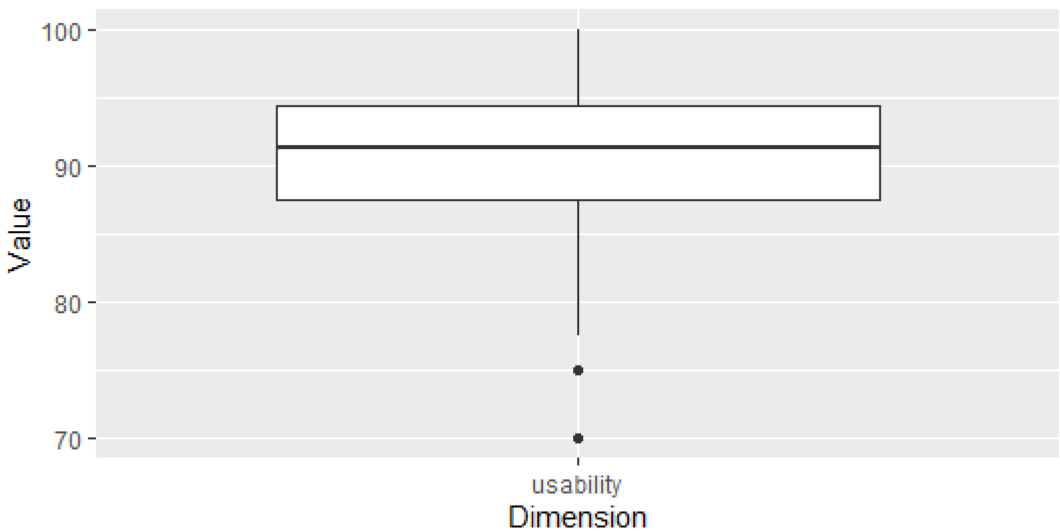


Figure 6.5: Boxplot of usability ratings.

The research question regarding workload was tested analogously against the reference value of 26 (see Figure 6.6 for boxplot). The mean value was $M = 13.39$ ($SD = 10.36$) with a significant Shapiro-Wilk-Test ($p < 0.05$). The Wilcoxon-Signed-Rank-Test was significant ($p < 0.01$; $r = 0.78$) with a large effect size. Hence, Hypothesis 3 could be confirmed.

The first item regarding system transparency created for this study was “I had all the information I would need to plan my non-driving related activities”, which was – like the other items on system transparency – answered on a 5-point Likert-scale ranging from -2 to 2. The mean value for this item was $M = 1.83$ ($SD = 0.38$). The Shapiro-Wilk-Test was significant ($p < 0.01$) and therefore a Wilcoxon-Signed-Rank-Test against the reference value of 0 was conducted. It was significant ($p < 0.01$; $r = 0.94$) with a large effect size.

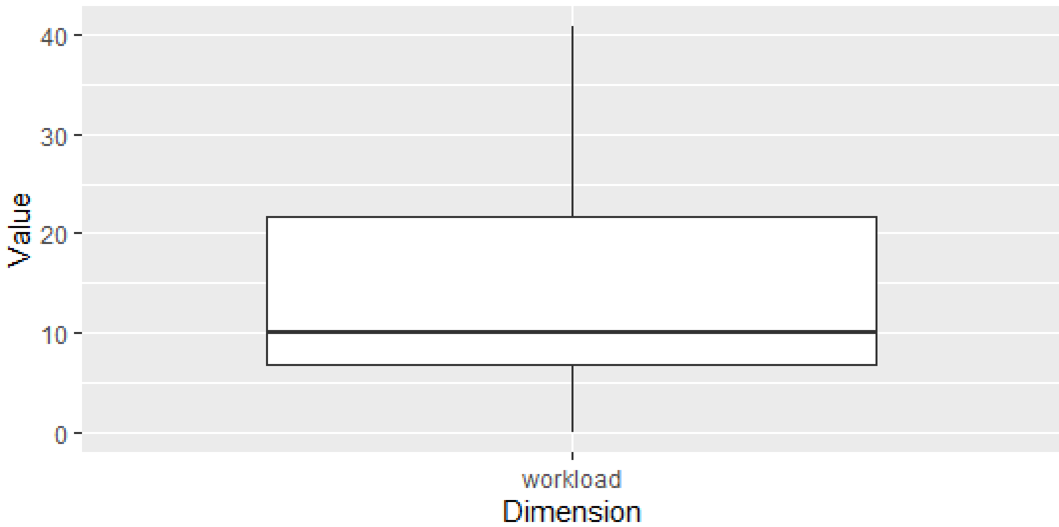


Figure 6.6: Boxplot of workload ratings.

The second item on transparency was “I always knew why the automation was not available”. The mean value was $M = 1.70$ ($SD = 0.79$). The Shapiro-Wilk-Test was significant ($p < 0.01$) and therefore normal distribution was not assumed and a Wilcoxon-Signed-Rank-Test was conducted against the reference value of 0. It was significant ($p < 0.01$; $r = 0.92$) with a large effect size.

The last item on transparency was “I was wondering when the automation will be available”, which was recoded so that bigger values would stand for better understanding. The mean value was $M = 1.43$ ($SD = 1.07$). No normal distribution could be assumed (Shapiro-Wilk-Test significant with $p < 0.01$) and therefore a Wilcoxon-Signed-Rank-Test was conducted against the reference value of 0. It was significant ($p < 0.01$; $r = 0.80$) with a large effect size.

6.4.2 TUI Dimensions

This study surveyed the TUI dimensions *curiosity*, *anxiety*, *interest*, *usability*, *usefulness*, and *scepticism*, and assessed the ITU score regarding this system. The mean values and standard deviations can be found in Table 6.2. The sums of the single dimensions were transformed to stanine scores ranging from one to nine with a mean of five and a standard deviation of two. This transformation was performed according to the test manual which also provides percentile ranks for the reference population of the studied sample that was used for validating the questionnaire (Kothgassner et al., 2012). The mean stanine scores were rounded off to be able to associate them to percentile ranks, which are not provided for decimal numbers. Note that the percentile ranks were taken from the questionnaire manual.

				Correlation with ITU	
	M	***SD***	**Percentile rank**	***r***	***p***
Curiosity	7.10	1.16	77- 89	0.40*	0.03
Anxiety	5.07	1.53	40 - 60	-0.38*	0.04
Interest	6.23	1.51	60 - 77	0.28	0.13
Usability	7.17	1.65	77 - 89	0.40*	0.03
Usefulness	8.69	0.76	96 - 100	0.60*	<0.01
Scepticism	3.79	1.54	23 - 40	-0.59*	<0.01
Intention to use (ITU)	7.76	1.06	89 - 96	-	-

Table 6.2: Means and standard deviations for the stanine scores of the TUI dimension along with the according percentile ranks. The Pearson's *r* values indicating the correlation between the TUI dimensions and ITU are presented along with the respective *p*-values. Significant correlations are marked with *.

Table 6.2 also shows the correlation between the single dimensions and ITU to investigate if the factors of the questionnaire were connected to the participants' intention to use the technology. All dimensions but *interest* significantly correlated with ITU.

It was also analysed whether the two dimensions *usefulness* and *satisfaction* of the van der Laan Acceptance Scale were correlated with ITU from the TUI questionnaire. The results showed that *usefulness* was not significantly correlated ($p = 0.91$, $r = 0.02$). The test for *satisfaction* also showed non-significant results ($p = 0.09$, $r = 0.31$).

6.4.3 Kano Categories

As described above, the respective information pieces were rated by participants according to Hölzing (2008) and Ma et al. (2019).

Estimated availability duration of the ADS before activation was categorised as must-be quality 15 times. For eight participants it was a one-dimensional quality, for six participants an attractive one, and for one participant it was an indifferent quality.

Reasons for non-availability was found to be an attractive quality, as it was put in this category eleven times. It was categorised as one-dimensional quality nine times and as indifferent eight times. Two times it was seen as a must-be quality.

Duration until ADS becomes available was categorised as attractive quality as well, since the answers of 12 participants led to this categorisation. It was further categorised as must-be quality nine times and as one-dimensional quality eight times, and only once as indifferent quality.

Overview of all availability periods on the route was an attractive quality for twelve participants, a one-dimensional quality for nine, and a must-be quality for five participants. Four participants categorised it as indifferent quality.

Pop-up with reasons for non-availability was categorised as must-be quality by fourteen and as attractive quality by eight participants. Four times, it was categorised as one-dimensional and another four times as indifferent quality.

6.4.4 Subjective Rating of Importance of Single Displays

Before answering questions regarding the importance of the single pieces of information, participants were asked how this HMI concept would help them understand the ADS (*How much would this HMI concept help you understanding the automated driving system; one: not at all – five: very much?)*. The mean value was M = 4.55 (SD = 0.50) indicating a high perceived system understanding. The ratings of the different information displays can be found in Table 6.3. The question asked was: *How important do you rate the individual pieces of information (1: not important at all – 5: very important)?*

	M	*SD*
Duration until ADS becomes available (displayed in the IC)	4.28	0.69
Pop-up with reasons for non-availability (displayed in the IC)	4.17	0.79
Icons indicating reasons for non-availability (displayed in the CID)	3	0.91
Estimated availability duration of the ADS before activation (displayed in the IC)	4.55	0.56
Overview of all availability periods on one's route (displayed in the CID)	3.55	1.00

Table 6.3: Means and standard deviations of the importance ratings of the single information displays.

6.4.5 Qualitative Interview

After the first automated driving section, the participants were asked how they would explain the system to a friend or family member who would like to try the ADS without any previous knowledge. This made it possible to find out how well they understood the principle of CAD only by using the HMI displays without any further explanation. It was of particular interest if participants understood the fact that the ADS was not available infinitely and not under all circumstances, without stating negative emotions or attitudes towards this system.

All participants mentioned in their explanation that the ADS was only available on specific route sections or under specific circumstances and that the HMI displays gave the relevant information on (non-)availability durations. They stated that switching between availability and non-availability was a possibility, but that the HMI displays let the users know when this occurred and that it was also possible to see when the next switch would happen. Therefore, it can be concluded that all participants understood the information on when the ADS would be available and the availability duration.

However, not all participants understood the display of the reasons for non-availability. Only 18 participants (60%) noticed and understood the HMI concept in the CID, while the rest of the participants either saw it without being able to retrieve relevant information or did not notice this display at all. Three (10%) of the participants stated that it took them some time to understand the HMI concept in the CID. The participants were asked to rate the importance of displaying the reasons for non-availability. The median value was calculated for the participants who have understood the displays and for those who did not. The median was *Mdn* = 4 for the first group and *Mdn* = 3 for the latter one, indicating that participants who rated the importance of this display higher were better at understanding the display.

The participants were also asked if there was anything they would want to change about the HMI concept. Four (13%) stated that the icons symbolising the possibility to conduct NDRAs were unnecessary or misleading, since the icons were a smartphone and a video symbol, which ignored the fact that participants were also allowed to read a book or just look out the window.

Three participants (10%) said that the HMI concept in the CID was overloaded and therefore too distracting for manual driving. Also, 10% stated that the HMI concept in the CID should not be displayed permanently even though it was important. Three participants noted that the pop-up window showing the reason for non-availability when participants tried to activate the system was not shown long enough. This pop-up was only shown for the duration the button was pressed and participants stated that it would be better if this window would appear for a certain amount of time even if the button was only pressed once without being held. 10% of the participants stated that they wished for better highlighting of the current driving section in the CID.

One participant wished that the time bar moved instead of the car icon. This would lead to the effect that past sections, which were considered unnecessary, would not be displayed anymore. Another participant wished for the possibility to switch between a time and a distance display when giving the availability durations.

6.5 Discussion

This study aimed to investigate whether the HMI concept consisting of all information displays tested beforehand fulfilled pre-defined criteria. Therefore, no tests were conducted to explore the content-related advantages of the concept components. Instead, tests were carried out to examine how well the individual components were combined into an overall concept and whether this overall concept fulfilled requirements for acceptance, usability, and workload as well as other aspects described above.

All hypotheses could be confirmed and hence the final HMI concept received an acceptance rating that was significantly above zero, i.e., in the positive range (for the single dimensions of the van der Laan et al. (1997) acceptance scale and for an overall acceptance value).

Furthermore, the usability rating was significantly higher than the threshold of good usability and was in fact in the range of excellent usability (Bangor et al., 2008). The workload was not significantly higher than the lowest workload score achieved in previous studies in which the single information components were tested, which leads to the conclusion that the final HMI concept did not increase workload, even though more information was displayed simultaneously in comparison to the previous studies. Furthermore, the transparency ratings were also significantly above zero and therefore in the hypothesised range.

As acceptance is a construct with widely varying definitions and measurement methods, two different scales, the TUI (Kothgassner et al., 2012) and the Acceptance Scale (van der Laan et al., 1997), were used in this study to investigate potential correlations or differences. The TUI results showed high values – especially for *usefulness* and ITU – indicating good acceptance. All but one TUI dimensions correlated significantly with the ITU dimension, which shows that these dimensions really are associated with ITU, which is an important indicator of technology acceptance (Davis, 1989; Venkatesh et al., 2003). On the other hand, the two dimensions *usefulness* and *satisfaction* of the Acceptance Scale (van der Laan et al., 1997) did not correlate with the ITU measure of the TUI. The Acceptance Scale is not theoretically based on the TAM or the UTAUT and therefore the behavioural intention to use the technology is not explicitly mentioned as a mediating factor. Nonetheless, this may simply mean that the Acceptance Scale is not a suitable instrument for measuring acceptance according to the TAM or following the definition of Adell (2009), who postulated that acceptance is the willingness to use technology when it is available. As stated by van der Laan et al. (1997), their questionnaire aims at investigating acceptance as an attitudinal factor. However, the attitudinal factor at least partly mediates the influence of PEoU and PU on ITU in the original TAM and therefore an association of attitudinal acceptance and behavioural acceptance should be present, which was not the case in this study. A more detailed discussion on acceptance and measurement methods can be found in Chapter 7.3.

Two out of the five information components, *Estimated availability duration of the ADS before activation* and *Pop-up with reasons for non-availability*, were categorised as a must-be quality according to the classification described by C. C. Yang (2005), which means that the absence of this information in the display would decrease customers' satisfaction while its presence would not increase satisfaction. The other components were categorised as attractive qualities. This means that the presence of these components could lead to an increase in satisfaction (also dependent on how they are achieved), but their absence would not lead to an increase in dissatisfaction.

The participants were also asked to rate the importance of the HMI components directly. It appeared that, on a descriptive level, the information displayed in the IC was rated as more important than the information bits displayed in the CID. This might result from participants thinking that the most important information is displayed in the IC. Nonetheless, this could also indicate that the information on when the ADS will be available, how long it will be available as well as information on why it does not work when an activation attempt failed are considered more critical. The overview of the availability sections was somehow redundant since the duration of the current section was always displayed in the IC. The same applies to the icons indicating the reasons for non-availability in the CID. The reasons were also available in the IC when participants tried to activate the ADS in a section of non-availability. Hence, all information in the CID aimed to enable participants to foresee how long automated driving was possible during the entire ride, helping with planning NDRAs and understanding the limitations of the system, while the information in the IC aimed to help participants understand the current situation and system status.

About half of the participants did not initially understand or notice the information in the CID in the first half of the experimental test drive which became apparent due to the mid-drive interview. Nevertheless, all participants appeared to have understood how the ADS worked and that limitations were part of the system. None of them expressed dissatisfaction or lack of understanding even though the way an SAE Level 3 ADS (SAE, 2021) worked was not explained to them before using it in this experiment. This stands in contrast to results from Danner, Pfromm, Limbacher, and Bengler (2020), where some participants thought the ADS should work without limitations – at least on a highway – even after experiencing RtIs on a highway drive. Considering the adequate answers when asked how they would explain the system, it can be concluded that knowing that the ADS is currently not available, when it will be available, and for how long, is sufficient information to form a mental model that allows participants to use the ADS without irritation – which is in accordance with Moser (2003), who stated that even incomplete mental models can lead to correct decisions when solving problems using a system.

The CID was too full and overloaded for some participants which might have influenced their answers in the questionnaires. Future studies should investigate, how to display all availability sections on one's route together with reasons for non-availability since crowded displays might lead to distraction. Furthermore, the icons indicating the possibility to conduct NDRAs should be changed to something more general, since the way they were designed in this study might implicate that only using the phone or watching videos were feasible activities. Another limitation was that the pop-up window was only displayed for the duration that the participants kept the activation button pressed down. This should be changed to a default minimum duration in future studies.

7 General Discussion

This chapter contains a summary of the results as well as a discussion of this work's limitations and an outlook for future research. For a summary of recommendations, please refer to Chapter 8.

7.1 Summary of the Results

Before testing whether presenting information on the availability of an ADS is advantageous for users, information needs had to be collected. In the first exploratory study presented in this thesis, these information needs were gathered through two different qualitative approaches. The first was a focus group discussion with automated driving experts and the second was a driving simulator study with laypersons as participants. The most important information needs regarding the ADS that emerged in this study were:

- Estimated availability duration of the ADS before activation when currently available
- Certainty of estimated availability duration when currently available
- Duration until ADS is available when currently not available
- Reasons for non-availability when currently not available
- Overview of all availability sections on the whole route

The aim of the second study was to investigate whether presenting the estimated availability duration before activation helps participants activate the ADS more purposefully and whether this additional information might positively influence the rating of the system and the HMI. It was shown that the presentation of this information can lead to more purposeful activation behaviour as well as to better acceptance and usability ratings.

The third study investigated whether the presentation of an indication of how certain the estimation of the availability duration could positively influence subjective ratings of the system. No significant effects were found and on a descriptive level, the baseline concept without the certainty symbol was preferred by participants. These results led to the decision to exclude this information from the concepts further on.

The fourth study aimed to test whether presenting information on when the ADS will be available when currently not available, as well as reasons for non-availability, might improve perceived system understanding as well as subjective ratings of the system and the HMI, like acceptance, usability, and workload. It was shown that perceived system understanding was significantly better in conditions where additional information was given but no significant effects regarding the standardised measures were found in this study.

The fifth study conducted within the scope of this work aimed to investigate how presenting a distribution of all availability sections of the ADS might influence the perceived system understanding, participants' performance in NDRAs, as well as subjective ratings of the system and the HMI, like acceptance, usability, and workload. Results showed that the presence of this additional information significantly influenced the performance in NDRAs positively, which might be a consequence of an improved ability to plan the activities. Acceptance ratings were also significantly higher in conditions where the additional information was displayed. The way of presenting the distribution of the availability did not have any effects (displaying the distribution on a map was compared to a more abstract but simpler approach). Furthermore, no effects were found on usability and workload.

For the sixth study, the HMI elements from previous studies were combined to create a final HMI concept. All information pieces that emerged from the first exploratory study were included except for the symbol indicating the certainty of the estimation of the availability duration since this information did not lead to any positive effects in the evaluation study. For this final study, no baseline concept was created since the single information pieces were already shown to be superior to not presenting the respective information. Instead, threshold values for the dependent variables were defined beforehand and tested against. The results indicate that the final HMI concept leads to acceptance ratings that are in a significantly positive range, excellent usability ratings, a lower workload score than in the previous studies, as well as positive ratings of perceived system understanding. Performance expectancy is seen as the strongest predictor of ITU (Adell et al., 2011) which is contently related to usefulness, describing how well a system helps users achieve their goals. Usefulness was the TUI dimension with the highest rating in this study, indicating that the HMI concept created in this work might have a strong positive impact on acceptance and usage of this technology. All hypotheses could be confirmed, and the goal of this concept can be considered achieved.

With regard to the research questions from Chapter 3.1 asking for which information regarding the availability of an ADS should be displayed in which system state and where, it can be summarised that the information letting the user know when the ADS will become available should be presented when the ADS is not available along with the reasons for the non-availability. The reasons should not be presented permanently without an option for users to hide this information. The time until the ADS becomes available can be displayed in the IC as this does not increase workload and indicates a potential transition of control, which is driving-relevant information. The reasons for non-availability should not be displayed in the IC but rather in the CID as it was shown to tendentially decrease usability in the IC and was considered less important by study participants. Furthermore, the system should present the reason for any unsuccessful activation attempts. The information of how long the ADS is available should be displayed when the ADS is available but not activated as well when it is activated. It can be displayed in the IC analogously to the information of when it will become available. An overview of all availability sections should be available to users in all system states, but they should also have the possibility to hide this information. As this display might contain more complex information, and therefore needs more space, and is also considered less important than the availability information regarding the current section, it should be displayed in the CID.

7.2 Conceptual Implications

Several information needs have been identified in the literature on automated driving (see Chapter 2.3.2), but there has been little to no understanding of how to present the availability of an ADS in different system states. In studies where availability times or reasons for non-availability were displayed (without the occurrence of an RtI), they were presented while driving automatically (Hecht, Kratzert, & Bengler, 2020; Holländer & Pfleging, 2018; Richardson et al., 2018). If this piece of information was part of a research question, it was mostly tested with regard to RtIs and take over performance (Holländer & Pfleging, 2018; Körber et al., 2018; Richardson et al., 2018; Wandtner et al., 2018). While these insights are very important, especially when considering the safety of automated driving, further questions must be answered to investigate the purposefulness of an automated driving function. Reasons why drivers might want to use such a system have to be explored to find out if an ADS is able to fulfil their needs. If the system is not able to meet their expectations but activation would be still considered beneficial – for example due to increased traffic safety – drivers' intention to use an ADSs must be considered. The present work makes a valuable contribution to the task of closing this gap in the literature.

Users' information needs regarding the availability of an ADS were collected, as were reasons why they would activate this system (Danner, Pfromm, Limbacher, & Bengler, 2020). One of the main reasons to activate was shown to be the possibility to engage in NDRAs. It was also shown that availability periods that are too short for conducting a planned NDRA can decrease subjective ratings of the system such as acceptance, which is considered a predictor of actual usage, and that by displaying the availability period before activation, this effect can be buffered by preventing users from activating when they cannot expect any advantage from it (Danner, Pfromm, & Bengler, 2020).

Furthermore, task performance in NDRAs was shown to be increased when providing information on all availability periods, which also led to increased acceptance ratings (Danner et al., 2023). Providing information on the reasons why the ADS is not available and when it will become available can positively influence system transparency (Danner et al., 2021) which on the other hand can increase trust in the ADS (Butakov & Ioannou, 2015). Initial incorrect mental models can converge over time with correct mental models through usage of a system (Beggiato & Krems, 2013). Nonetheless, for this to be the case users need input information to update their mental models. As shown in Chapter 4, users might have incorrect assumptions about the reasons for non-availability. Since assumptions that are not falsified will be considered valid (Stanton & Young, 2000), it is important to provide information on actual reasons for non-availability to prevent the consolidation of incorrect mental models.

Moreover, presenting the reasons for non-availability when users try to activate the ADS received high importance ratings by participants, which is in accordance with the finding that every user input should lead to an action of the system (Shneiderman et al., 2018).

Hence, this research adds important information needs to the list of requirements that must be considered when developing HMIs for ADS. Recommendations can be found in Chapter 8.

7.3 Methodological Implications

In the second chapter of this thesis, acceptance was defined mainly through the TAM (Davis, 1989), which was not created for explaining acceptance in the driving context. The corresponding questionnaire is therefore seldom used in studies investigating automated driving. The van der Laan et al. (1997) Acceptance Scale on the other hand is the most used one with regard to surveying whether needs and requirements are satisfied by ADSs (Adell, Nilsson, & Várhelyi, 2018) and was therefore used in this work. In general, the measurement of acceptance is dependent on how acceptance is defined. Regan et al. (2018) summarise different acceptance definitions as well as different ways of investigating this construct in the driving context. It becomes apparent that acceptance is surveyed very differently in different studies. A number of studies try to measure it directly using words like “acceptability”. Others use the participants’ willingness to use a system or pay for it to explain acceptance (Regan et al., 2018). The car TAM (Osswald et al., 2012) is an extension of the TAM and the UTAUT and aims to adapt the models to the driving context using similar predictors for acceptance. Other researchers see the importance of a system as an indicator of acceptance even though this is not based on any particular definition of acceptance (Regan et al., 2018). Nonetheless, the subjectively perceived importance somehow reflects an attitude towards the system and could be an indicator of general acceptance (Adell, Nilsson, & Várhelyi, 2018).

The results of the final study in this work showed that different acceptance questionnaires can lead to different results. The Acceptance Scale dimensions were not correlated with the ITU dimension of the TUI, even though the former aims to measure the attitudinal component of acceptance (van der Laan et al., 1997), which is actually considered a predictor of the behavioural component ITU in the original TAM (Davis, 1989). Hence, it can be concluded that the definition and measurement of acceptance are not consistent in research on ADS and that a broader approach of using different measures associated with it may be necessary to capture it, while not obtaining a singular value of acceptance itself. Furthermore, the results of this work indicate that the Acceptance Scale by van der Laan et al. (1997) might have to be newly validated and that it should be tested if it is an appropriate tool for investigating acceptance as a behavioural or attitudinal construct. As the acceptance definition in this thesis is based on the TAM, which identifies ITU as one of the main predictors and the Acceptance Scale was not associated with the ITU, it must be deduced that the Acceptance Scale was not the right tool to measure acceptance – at least in the studies conducted for this thesis. Nonetheless, since different instruments were used besides the Acceptance Scale – measuring transparency, usability, workload, and the importance of the HMI concept – the need for a broad approach of testing new systems is satisfied in this work.

The TAM2 and the UTAUT include predictors that would not have been applicable in the studies presented in this thesis, as factors like *social influence* cannot be taken into account when investigating systems that are not available on the market yet. Moreover, a simplistic approach to this measure was preferred. Nevertheless, considering the fact that all but one of the dimensions of the TUI questionnaire significantly correlated with the ITU, it can be concluded that the simplistic approach of using only two factors – as in the TAM as well as in the Acceptance Scale – might not be sufficient to investigate such a complex construct.

7.4 Limitations

This work aimed to investigate which information on the availability of an ADS users might need to be able to understand and accept the same. In the literature review presented in Chapter 2, no papers were found focusing on this topic and therefore all information needs had to be collected firstly as part of this thesis. The process described in DIN EN ISO 9241-210 (International Organization for Standardization, 2019) served as a guide for the methodological approach as well as the approach in Beggiato, Hartwich, et al. (2015) where a focus group discussion was combined with a driving simulator study. The ISO norm does not provide guidelines for a research project including several studies with inference statistical analyses but is more focused on a quick and iterative user research and prototyping process. According to this standard, the overall HMI concept would have been prototyped very roughly after investigating the users' information needs, iteratively testing, and improving the same prototype, while simultaneously developing it with constantly increasing fidelity. In this thesis, this process was adjusted to an approach more suitable for thoroughly investigating the potential effects of the singular information needs instead of continuously testing an HMI concept containing all the information, which would prevent finding out which information piece causes which effect. Nonetheless, the general process chosen for this thesis deviated from a standardised procedure, which might have led to biases.

The samples in most of the studies conducted for this thesis consisted mainly of younger male participants. This demographic group is likely to be more enthusiastic about automated driving than other groups (Nielsen & Haustein, 2018), which could have biased the results. The ADSs might have been rated better than they would have been with more heterogenic samples. However, all but the last studies tested different HMI concepts and therefore comparisons between ratings were used to describe effects. This approach counteracts the impact of positive bias, as this would have occurred in every experimental group.

One assumption made for some of the studies was that users or participants respectively have specific activities on their minds that they want to finish while driving automatically. In reality, this will not always be true, since study participants stated that they would also like to talk to other passengers or watch out of the window while driving automatically (Danner, Pfromm, Limbacher, & Bengler, 2020). Therefore, the results achieved in this thesis cannot be generalised to every situation of driving automatically but rather to specific situations where the user of an ADS plans on conducting NDRAs of certain, at least partly predictable durations.

Moreover, in some cases, the NDRAs in the studies were chosen to resemble naturalistic activities to increase the participants' commitment to the tasks. Nevertheless, it must be considered that participants would have conducted completely different NDRAs for themselves if they would not have been observed in a study context. Hence, the effects achieved in this work might have been different with NDRAs participants really would have wanted to conduct.

Furthermore, perceived system understanding, and system transparency were treated like the same construct in this work. In studies where this construct was investigated, no standardised questionnaire was used to collect information on the same but items that were created for the purpose of this work. They have not been validated beforehand and hence it cannot be ruled out that these items were not the right ones to survey system transparency.

The HMI concepts described in this work mainly addressed the visual channel, which is the most important one while driving (Meixner & Müller, 2017). Since most information relevant to driving is usually conveyed visually and drivers also perceive the driving environment using their optic senses, adding additional information to this sensory channel should be avoided (Bruder & Didier, 2015). Nonetheless, while the tactile channel is not suitable for conveying detailed information like presented in this work, using speech output conveying at least some of the information, for example, would have been possible. While investigating how this would affect the ratings of the system and how well this would support users is important, the focus of the present work was primarily on investigating the information needs and their impact on the content level.

Due to the limitations stated in this chapter, it is necessary to carefully interpret the results of the studies and refrain from over-generalising the gained insights. Still, the implications and statements made in the discussion and summary of this work were formulated taking these limitations into account and are therefore still considered valid.

7.5 Outlook and Further Research Questions

Future studies should systematically investigate how different types of NDRAs might lead to different results in subjective ratings of the ADS and the HMI concept. In the final study of this thesis, the participants were free to choose their NDRAs for periods of automated driving in the driving simulator but the study context – including being monitored – might prevent participants from behaving as they would in a more naturalistic driving context. Therefore, studies should focus on exploring how to create a study environment in which participants conduct NDRAs that are important for themselves to test how helpful the displayed information is. Moreover, it should be investigated how participants rate the HMI concept and its importance in situations in which they do not want to conduct any NDRAs at all.

Another aspect that could not be covered in this work, but seems crucial for understanding usage behaviour regarding limited automated driving functions like SAE Level 3 and 4 (SAE, 2021), is to study how much availability users will face on real routes. By knowing this, it would be possible to estimate how often changes of the system state will occur, which might have an impact on the relevance of some segments of the HMI concept developed in this work. Future work could explore how different lengths of availability sections and different distributions of the same influence usage behaviour and deduct which of the presented information displays are necessary in which situation.

Furthermore, future research should investigate how to measure system transparency in a standardised but still flexible way. This construct is measured in this work using self-created items and is referred to as perceived system understanding, but no data were collected on actual or objective system understanding. It would be possible to study this by testing participants' knowledge of the system which would be similar to the measurement of transparency. Moreover, it should be investigated how transparency affects acceptance and its factor *perceived usefulness*, since the mere presentation of additional information affected acceptance in the studies presented in this work even though transparency is mainly reported to influence trust (Oliveira et al., 2020). As described above, measuring acceptance is complicated and sometimes handled ambiguously, and therefore future research should test existing acceptance questionnaires on their validity and systematically create a catalogue of the different instruments and when to use them.

As described in the previous chapter, this work focuses on conveying information using the visual channel. Future studies should explore how participants would rate speech outputs giving information on the availability of the ADS. Especially in periods of manual driving, the overview of all availability sections along with the reasons for non-availability might be too distracting if displayed like in Chapter 6.

8 Recommendations for the Design of the HMI Concept

In the following, recommendations are made on the requirements to be considered when designing HMI concepts for ADSs. These requirements are only valid under the following circumstances:

1. The concept is for an in-vehicle HMI display accompanying an ADS of SAE Level 3 or Level 4.
2. The concept was mainly tested on highway drives and is therefore only applicable to these.
3. The usefulness of some of the displayed information is based on the assumption that users want to conduct specific NDRAs.

Furthermore, the requirements only refer to the availability of an ADS and do not include an exhaustive list of all necessary HMI components for automated driving. It is also presumed that an indication of whether an ADS is generally available or active exists (which was the case in all cited studies in this work that included automated driving).

It must be noted that the following requirements are based on study results presented in this work and therefore reflect the average opinions and attitudes of study populations. This does not mean that these requirements cover the needs of every single individual but rather that interindividual differences exist and should be accounted for. In addition, the designs of the concepts, as shown in the images in the respective studies, were not a direct subject of this research. Thus, the recommendations are detached from detailed design decisions and only concern the content, duration, and location of the concept components.

Following the requirements outlined below while designing HMI concepts can help increase users' system understanding in relation to ADSs. Furthermore, acceptance – which is a predictor of usage (Adell, Várhelyi, & Nilsson, 2018a) and could be negatively affected by non-permanent availability of an ADS – can be positively influenced by displaying the recommended information. Interindividual differences that might emerge from different expectations towards an ADS can be decreased by increasing transparency by means of the additional HMI displays shaping the users' mental models:

- When the ADS is not available – hence, the vehicle is not within an ODD – the duration of how long it will take until an ODD is reached should be displayed.

- When the ADS is not available and users try to activate the same, information should be presented on why it is not available.
- When the ADS is available but not activated, the duration of how long it will be available should be displayed.
- When the ADS is available and activated, the duration of how long it will be available should be displayed.
- Users should be able to inform themselves about the distribution of all availability sections on their route.
- Users should be able to inform themselves about the reasons for non-availability for each section.

The first four points are considered more important than the latter two (see Chapter 6.4.3) and should be displayed as part of the automation HMI in the IC, since the IC is considered suitable to convey system relevant information and serves as display for the automation HMI in several studies (Bengler et al., 2020). The last two points can be displayed in the CID, as it was done in the final study of this work and should be presented as optional. That means that users should have the possibility to either retrieve or hide this display. Since the information on the distribution of the availability sections and the accumulated reasons for non-availability have no direct influence on the immediate driving situation, the IC would not be suitable.

The outcomes summarised in this chapter have not been discussed in research before. Considering the positive impact of the concept on users' acceptance and perceived system understanding, it is advisable that the recommendations resulting from this research are taken into account by developers of HMIs for automated driving.

References

Adell, E. (2009). *Driver experience and acceptance of driver support systems - a case of speed adaption* [Dissertation]. Lund Institute of Technology, Sweden.

Adell, E., Nilsson, L [Lena], & Várhelyi, A. (2018). How Is Acceptance Measured? Overview of Measurement Issues, Methods and Tools. In M. A. Regan, T. Horberry, & A. Stevens (Eds.), *Driver Acceptance of New Technology* (pp. 73–88). CRC Press.

Adell, E., Várhelyi, A., & Fontana, M. d. (2011). The effects of a driver assistance system for safe speed and safe distance – A real-life field study. *Transportation Research Part C: Emerging Technologies*, *19*(1), 145–155. https://doi.org/10.1016/j.trc.2010.04.006

Adell, E., Várhelyi, A., & Nilsson, L [Lena]. (2018a). The Definition of Acceptance and Acceptability. In M. A. Regan, T. Horberry, & A. Stevens (Eds.), *Driver Acceptance of New Technology* (pp. 11–22). CRC Press. https://doi.org/10.1201/9781315578132-2

Adell, E., Várhelyi, A., & Nilsson, L [Lena]. (2018b). Modelling Acceptance of Driver Assistance Systems: Application of the Unified Theory of Acceptance and Use of Technology. In M. A. Regan, T. Horberry, & A. Stevens (Eds.), *Driver Acceptance of New Technology* (pp. 23–34). CRC Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Albers, D., Radlmayr, J., Loew, A., Hergeth, S., Naujoks, F., Keinath, A., & Bengler, K. (2020). Usability Evaluation—Advances in Experimental Design in the Context of Automated Driving Human–Machine Interfaces. *Information*, *11*(5), 240. https://doi.org/10.3390/info11050240

Bach, N. (2000). *Mentale Modelle als Basis von Implementierungsstrategien*. Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-663-08757-1

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Barrick, M. R., Mount, M. K., & Li, N. (2013). The Theory of Purposeful Work Behavior: The Role of Personality, Higher-Order Goals, and Job Characteristics. *Academy of Management Review*, *38*(1), 132–153. https://doi.org/10.5465/amr.2010.0479

BaSt. (2012). *Rechtsfolgen zunehmender Fahrzeugautomatisierung: Gemeinsamer Schlussbericht der Projektgruppe ; [Bericht zum Forschungsprojekt F 1100.5409013.01 ...]. Berichte der Bundesanstalt für Strassenwesen : F, Fahrzeugtechnik: Vol. 83*. Wirtschaftsverlag NW.

Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J., Othersen, I., & Petermann-Stock, I. (2015). What would drivers like to know during automated driving? Information needs at different levels of automation. *7th Conference on Driver Assistance.* Advance online publication. https://doi.org/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. https://doi.org/10.1016/j.trf.2012.12.006

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 75–84. https://doi.org/10.1016/j.trf.2015.10.005

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver-automation interaction: An approach using automation uncertainty. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(6), 1130–1141. https://doi.org/10.1177/0018720813482327

Bengler, K. (2017). Driver and Driving Experience in Cars. In G. Meixner & C. Müller (Eds.), *Human–Computer Interaction Series. Automotive User Interfaces* (pp. 79–94). Springer International Publishing. https://doi.org/10.1007/978-3-319-49448-7_3

Bengler, K., Bubb, H., Lange, C., Aringer, C., Trübswetter, N., Conti, A., & Zimmermann, M. (2015). Messmethoden. In H. Bubb, K. Bengler, R. E. Grünen, & M. Vollrath (Eds.), *Automobilergonomie* (pp. 617–662). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-8348-2297-0_11

Bengler, K., Dietmayer, K., Farber, B., Maurer, M [Markus], Stiller, C., & Winner, H. (2014). Three Decades of Driver Assistance Systems: Review and Future Perspectives. *IEEE Intelligent Transportation Systems Magazine*, *6*(4), 6–22. https://doi.org/10.1109/MITS.2014.2336271

Bengler, K., Pfromm, M., & Bruder, R. (2015). Bedienelemente. In H. Winner, S. Hakuli, F. Lotz, & C. Singer (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 647–657). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-05734-3_35

Bengler, K., Rettenmaier, M., Fritz, N., & Feierle, A. (2020). From HMI to HMIs: Towards an HMI Framework for Automated Driving. *Information*, *11*(2), 61. https://doi.org/10.3390/info11020061

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Borojeni, S. S., Flemisch, F., Baltzer, M., & Boll, S. (2018). Automotive UI for Controllability and Safe Transitions of Control. In Unknown (Ed.), *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (pp. 23–29). ACM Press. https://doi.org/10.1145/3239092.3239559

Brave, S., & Nass, C. (2007). Emotion in Human-Computer Interaction. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook.* CRC Press.

Brooke, J. (1996). Sus: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability Evaluation in Industry.* Chapman and Hall/CRC.

Bruder, R., & Didier, M. (2015). Gestaltung von Mensch-Maschine-Schnittstellen. In H. Winner, S. Hakuli, F. Lotz, & C. Singer (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 633–645). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-05734-3_34

Bubb, H. (2003). Fahrerassistenz primär ein Beitrag zum Komfort oder für die Sicherheit? In VDI (Ed.), *Der Fahrer im 21. Jahrhundert : Anforderungen, Anwendungen, Aspekte für Mensch-Maschine-Systeme.* VDI-Verlag.

Bubb, H., & Bengler, K. (2015). Fahrerassistenz. In H. Bubb, K. Bengler, R. E. Grünen, & M. Vollrath (Eds.), *Automobilergonomie* (pp. 525–582). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-8348-2297-0_9

Bubb, H., Bengler, K., Breuninger, J., Gold, C., & Helmbrecht, M. (2015). Systemergonomie des Fahrzeugs. In H. Bubb, K. Bengler, R. E. Grünen, & M. Vollrath (Eds.), *Automobilergonomie* (pp. 259–344). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-8348-2297-0_6

Bubb, H., Grünen, R. E., & Remlinger, W. (2015). Anthropometrische Fahrzeuggestaltung. In H. Bubb, K. Bengler, R. E. Grünen, & M. Vollrath (Eds.), *Automobilergonomie* (pp. 345–470). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-8348-2297-0_7

Bubb, H., Vollrath, M., Reinprecht, K., Mayer, E., & Körber, M. (2015). Der Mensch als Fahrer. In H. Bubb, K. Bengler, R. E. Grünen, & M. Vollrath (Eds.), *Automobilergonomie* (pp. 67–162). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-8348-2297-0_3

Bullinger, H.-J. (Ed.). (1985). *Software-Ergonomie: Vol. 2. Software-Ergonomie '85: Mensch-Computer-Interaktion*. Teubner.

Butakov, V., & Ioannou, P. (2015). Driving Autopilot with Personalization Feature for Improved Safety and Comfort. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems* (pp. 387–393). IEEE. https://doi.org/10.1109/ITSC.2015.72

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Carroll, J. M., & Olson, J. R. (1987). *Mental Models in Human-Computer Interaction: Research Issues About What the User of Software Knows*. National Academy Press.

Carsten, O., & Martens, M. H. (2018). How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work*, *29*(6), 415. https://doi.org/10.1007/s10111-018-0484-0

Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692–702. https://doi.org/10.1080/10447318.2015.1070549

Danner, S., Feierle, A., Gary, S., & Bengler, K. (2023). The effects of displaying availability sections of an automated driving function. *Frontiers in Future Transportation*, *4*, Article 1153599. https://doi.org/10.3389/ffutr.2023.1153599

Danner, S., Hecht, T., Steidl, B., & Bengler, K. (2021). Why is the Automation Not Available and When Can I Use It? In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Lecture Notes in Networks and Systems. Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021)* (Vol. 221, pp. 623–632). Springer International Publishing. https://doi.org/10.1007/978-3-030-74608-7_76

Danner, S., Pfromm, M., & Bengler, K. (2020). Does Information on Automated Driving Functions and the Way of Presenting It before Activation Influence Users' Behavior and Perception of the System? *Information*, *11*(1), 54. https://doi.org/10.3390/info11010054

Danner, S., Pfromm, M., Limbacher, R., & Bengler, K. (2020). Information needs regarding the purposeful activation of automated driving functions - an exploratory study. In D. de Waard, A. Tofetti, L. Pietrantoni, T. Franke, J.-F. Petiot, C. Dumas, . . . F. Mars (Chairs), *HFES Europe,* Nantes, FR.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

DIN (1953). *DIN 73001:1953-10, Bedienung von Kraftfahrzeugen mit Verbrennungsmotor* (73001). Berlin. Beuth Verlag GmbH.

DIN ISO (2017). *Ergonomie der Mensch-System-Interaktion:* Teil 112: Grundsätze der Informationsdarstellung (9241-112:2017).

Donges, E. (1982). Aspekte der Aktiven Sicherheit bei der Führung von Personenkraftwagen. *Automobil-Industrie*(27), 183–190.

Donges, E. (2015). Fahrerverhaltensmodelle. In H. Winner, S. Hakuli, F. Lotz, & C. Singer (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 17–26). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-05734-3_2

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7

Ebnali, M., Hulme, K., Ebnali-Heidari, A., & Mazloumi, A. (2019). How does training effect users' attitudes and skills needed for highly automated driving? *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 184–195. https://doi.org/10.1016/j.trf.2019.09.001

Eckstein, L. (2001). *Entwicklung und Überprüfung eines Bedienkonzepts und von Algorithmen zum Fahren eines Kraftfahrzeugs mit aktiven Sidesticks*. Zugl.: Stuttgart, Univ., Diss (Als Ms. gedr). *Fortschritt-Berichte VDI Reihe 12, Verfahrenstechnik, Fahrzeugtechnik: Vol. 471*. VDI Verl.

Feierle, A., Bücherl, F., Hecht, T., & Bengler, K. (2020). Evaluation of Display Concepts for the Instrument Cluster in Urban Automated Driving. In T. Ahram, W. Karwowski, S. Pickl, & R. Taiar (Eds.), *Advances in Intelligent Systems and Computing. Human Systems Engineering and Design II* (Vol. 1026, pp. 209–215). Springer International Publishing. https://doi.org/10.1007/978-3-030-27928-8_32

Feierle, A., Danner, S., Steininger, S., & Bengler, K. (2020). Information Needs and Visual Attention during Urban, Highly Automated Driving—An Investigation of Potential Influencing Factors. *Information*, *11*(2), 62. https://doi.org/10.3390/info11020062

Feierle, A., Holderied, M., & Bengler, K. (2020). Evaluation of Ambient Light Displays for Requests to Intervene and Minimal Risk Maneuvers in Highly Automated Urban Driving. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1–8). IEEE. https://doi.org/10.1109/ITSC45102.2020.9294645

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the Duration of Automated Driving Influences Take-Over Performance and Gaze Behavior. In C. M. 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). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-53305-5_22

Feldhütter, A., Segler, C., & Bengler, K. (2018). Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness? In N. A. Stanton (Ed.), *Advances in Intelligent Systems and Computing. Advances in Human Aspects of Transportation* (Vol. 597, pp. 730–741). Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_70

Field, A., Miles, J., & Field, Z. (2013). *Discovering statistics using R* (Reprint). Sage.

Forster, Y., Hergeth, S., Naujoks, F., Beggiato, M., Krems, J. F., & Keinath, A. (2019). Learning and Development of Mental Models during Interactions with Driving Automation: A Simulator Study. In *Proceedings of the 10th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design: driving assessment 2019* (pp. 398–404). University of Iowa. https://doi.org/10.17077/drivingassessment.1724

Forster, Y., Hergeth, S., Naujoks, F., & Krems, J. F. (2018). How Usability Can Save the Day - Methodological Considerations for Making Automated Driving a Success Story. In Unknown (Ed.), *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (pp. 278–290). ACM Press. https://doi.org/10.1145/3239060.3239076

Forster, Y., Hergeth, S., Naujoks, F., Krems, J., & Keinath, A. (2019a). Tell Them How They Did: Feedback on Operator Performance Helps Calibrate Perceived Ease of Use in Automated Driving. *Multimodal Technologies and Interaction*, *3*(2), 29. https://doi.org/10.3390/mti3020029

Forster, Y., Hergeth, S., Naujoks, F., Krems, J., & Keinath, A. (2019b). User Education in Automated Driving: Owner's Manual and Interactive Tutorial Support Mental Model Formation and Human-Automation Interaction. *Information*, *10*(4), 143. https://doi.org/10.3390/info10040143

Forster, Y., Hergeth, S., Naujoks, F., Krems, J. F., & Keinath, A. (2020). Self-report measures for the assessment of human–machine interfaces in automated driving. *Cognition, Technology & Work*, *22*(4), 703–720. https://doi.org/10.1007/s10111-019-00599-8

Forster, Y., Naujoks, F., & Neukum, A. (2016). Your Turn or My Turn? In P. Green, S. Boll, G. Burnett, J. Gabbard, & S. Osswald (Eds.), *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - Automotive'UI 16* (pp. 253–260). ACM Press. https://doi.org/10.1145/3003715.3005463

Fundin, A., & Nilsson, L [Lars] (2003). Using Kano's Theory of Attractive Quality to Better Understand Customer Experiences with E-Services. *Asian Journal on Quality*, *4*(2), 32–49. https://doi.org/10.1108/15982688200300018

Gasser, T. M., Seeck, A., & Smith, B. W. (2015). Rahmenbedingungen für die Fahrerassistenzentwicklung. In H. Winner, S. Hakuli, F. Lotz, & C. Singer (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 27–54). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-05734-3_3

Giesen, C. (2016, October 15). Auf Autopilot. *Süddeutsche Zeitung,* 2016. https://www.sueddeutsche.de/wirtschaft/tesla-auf-autopilot-1.3163214

Gold, C. (2016). *Modeling of Take-Over Performance in Highly Automated Vehicle Guidance.* Technische Universität München, München.

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*(1), 1938–1942. https://doi.org/10.1177/1541931213571433

Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis & Prevention*, *29*(4), 495–506. https://doi.org/10.1016/S0001-4575(97)00029-8

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(50), Article 9, 904–908.

Hartwich, F., Hollander, C., Johannmeyer, D., & Krems, J. F. (2021). Improving Passenger Experience and Trust in Automated Vehicles Through User-Adaptive HMIs: "The More the Better" Does Not Apply to Everyone. *Frontiers in Human Dynamics*, *3*, Article 669030. https://doi.org/10.3389/fhumd.2021.669030

Hartwich, F., Schmidt, C., Gräfing, D., & Krems, J. F. (2020). In the Passenger Seat: Differences in the Perception of Human vs. Automated Vehicle Control and Resulting HMI Demands of Users. In H. Krömker (Ed.), *Lecture Notes in Computer Science. HCI in Mobility, Transport, and Automotive Systems. Automated Driving and In-Vehicle Experience Design* (Vol. 12212, pp. 31–45). Springer International Publishing. https://doi.org/10.1007/978-3-030-50523-3_3

Hecht, T., Danner, S., Feierle, A., & Bengler, K. (2020). Does a Confidence Level for Automated Driving Time Estimations Improve the Subjective Evaluation of an Automation HMI? *Multimodal Technologies and Interaction*, *4*(3), 36. https://doi.org/10.3390/mti4030036

Hecht, T., Darlagiannis, E., & Bengler, K. (2020). Non-driving Related Activities in Automated Driving – An Online Survey Investigating User Needs. In T. Ahram, W. Karwowski, S. Pickl, & R. Taiar (Eds.), *Advances in Intelligent Systems and Computing. Human Systems Engineering and Design II* (Vol. 1026, pp. 182–188). Springer International Publishing. https://doi.org/10.1007/978-3-030-27928-8_28

Hecht, T., Kratzert, S., & Bengler, K. (2020). The Effects of a Predictive HMI and Different Transition Frequencies on Acceptance, Workload, Usability, and Gaze Behavior during Urban Automated Driving. *Information*, *11*(2), 73. https://doi.org/10.3390/info11020073

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In J. Terken (Ed.), *ACM Digital Library, Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 210–217). ACM. https://doi.org/10.1145/2516540.2516554

Higashinaka, R., Dohsaka, K., & Isozaki, H. (2008). Effects of Self-Disclosure and Empathy in Human-Computer Dialogue. In IEEE (Ed.), *2008 IEEE Spoken Language Technology Workshop.* I E E E.

Hock, P., Kraus, J., Babel, F., Walch, M., Rukzio, E., & Baumann, M. (2018). How to Design Valid Simulator Studies for Investigating User Experience in Automated Driving. In Unknown (Ed.), *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (pp. 105–117). ACM Press. https://doi.org/10.1145/3239060.3239066

Holländer, K., & Pfleging, B. (2018). Preparing Drivers for Planned Control Transitions in Automated Cars. In S. Abdennadher & F. Alt (Eds.), *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia - MUM 2018* (pp. 83–92). ACM Press. https://doi.org/10.1145/3282894.3282928

Hölzing, J. A. (Ed.). (2008). *Die Kano-Theorie der Kundenzufriedenheitsmessung*. Gabler. https://doi.org/10.1007/978-3-8349-9864-4

Hone, K. (2006). Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interacting with Computers*, *18*(2), 227–245. https://doi.org/10.1016/j.intcom.2005.05.003

Howard, D., & Dai, D. (2013). Public Perceptions of Self-driving Cars: The Case of Berkeley, California. *Paper Presented at the 93rd Annual Meeting TRB*.

International Organization for Standardization (2019). *ISO 9241-210:2019:* Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems (9241-210:2019).

ISO (2003). *Road vehicles - Ergonomic aspects of transport information and control systems Procedure for assessing suitability for use while driving* (17287:2003).

ISO (2009). *4040:2009-02:* Road vehicles - Location of hand controls, indicators and tell-tales in motor vehicles.

Itoh, M. (2012). Toward overtrust-free advanced driver assistance systems. *Cognition, Technology & Work*, *14*(1), 51–60. https://doi.org/10.1007/s10111-011-0195-2

Jing, P., Xu, G., Chen, Y., Shi, Y., & Zhan, F. (2020). The Determinants behind the Acceptance of Autonomous Vehicles: A Systematic Review. *Sustainability*, *12*(5), 1719. https://doi.org/10.3390/su12051719

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 Intelligent Systems and Computing. Advances in Human Aspects of Transportation* (Vol. 597, pp. 606–617). Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_59

Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design*, *45*(3), 339. https://doi.org/10.1504/IJVD.2007.014909

Knoll, P. (2015). Anzeigen für Fahrerassistenzsysteme. In H. Winner, S. Hakuli, F. Lotz, & C. Singer (Eds.), *Handbuch Fahrerassistenzsysteme* (pp. 659–673). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-05734-3_36

König, M., & Neumayr, L. (2017). Users' resistance towards radical innovations: The case of the self-driving car. *Transportation Research Part F: Traffic Psychology and Behaviour*, *44*, 42–52. https://doi.org/10.1016/j.trf.2016.10.013

Körber, M., Prasch, L., & Bengler, K. (2018). Why Do I Have to Drive Now? Post Hoc Explanations of Takeover Requests. *Human Factors*, *60*(3), 305–323. https://doi.org/10.1177/0018720817747730

Kothgassner, O. D., Felnhofer, A., Hauk, N., Kastenhofer, E., Gomm, J., & Kryspin-Exner, I. (2012). *TUI: Technology Usage Inventory.*

Kraus, J., Scholz, D., Stiegemeier, D., & Baumann, M. (2020). The More You Know: Trust Dynamics and Calibration in Highly Automated Driving and the Effects of Take-Overs, System Malfunction, and System Transparency. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *62*(5), 718–736. https://doi.org/10.1177/0018720819853686

Kunze, A., Summerskill, S. J., Marshall, R., & Filtness, A. J. (2019). Automation transparency: Implications of uncertainty communication for human-automation interaction and interfaces. *Ergonomics*, *62*(3), 345–360. https://doi.org/10.1080/00140139.2018.1547842

Lansdown, T. C., Brook-Carter, N., & Kersloot, T [Tanita] (2004). Distraction from multiple in-vehicle secondary tasks: Vehicle performance and mental workload implications. *Ergonomics*, *47*(1), 91–104. https://doi.org/10.1080/00140130310001629775

Large, D [D.], Burnett, G [G.], Morris, A., Muthumani, A [A.], & Matthias, R. (2017). Design Implications of Drivers'Engagement with Secondary Activities During Highly-Automated Driving—A Longitudinal SimulatorStudy. In *Proceedings of the Road Safety and Simulation International Conference (RSS),* The Hague.

Larsson, A. F. L. (2012). Driver usage and understanding of adaptive cruise control. *Applied Ergonomics*, *43*(3), 501–506. https://doi.org/10.1016/j.apergo.2011.08.005

Larue, G. S., Rakotonirainy, A., Haworth, N. L., & Darvell, M. (2015). Assessing driver acceptance of Intelligent Transport Systems in the context of railway level crossings. *Transportation Research Part F: Traffic Psychology and Behaviour*, *30*, 1–13. https://doi.org/10.1016/j.trf.2015.02.003

Lau, M., Wilbrink, M., Dodiya, J., & Oehl, M. (2020). Users' Internal HMI Information Requirements for Highly Automated Driving. In C. Stephanidis, M. Antona, & S. Ntoa (Eds.), *Communications in Computer and Information Science. HCI International 2020 – Late Breaking Posters* (Vol. 1294, pp. 585–592). Springer International Publishing. https://doi.org/10.1007/978-3-030-60703-6_75

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Lemmer, K. (2016). *Automatisierter Straßenverkehr der Zukunft: Automatisierter Straßenverkehr der Zukunft. acatech Studie: [I]*. Herbert Utz Verlag.

Löfgren, M., & Witell, L. (2005). Kano's Theory of Attractive Quality and Packaging. *Quality Management Journal*, *12*(3), 7–20. https://doi.org/10.1080/10686967.2005.11919257

Lu, Z., Happee, R., Cabrall, C. D., Kyriakidis, M., & Winter, J. C. de (2016). Human factors of transitions in automated driving: A general framework and literature survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, *43*, 183–198. https://doi.org/10.1016/j.trf.2016.10.007

Ma, M.-Y., Chen, C.-W., & Chang, Y.-M. (2019). Using Kano model to differentiate between future vehicle-driving services. *International Journal of Industrial Ergonomics*, *69*, 142–152. https://doi.org/10.1016/j.ergon.2018.11.003

Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. C. Presmeg (Eds.), *Advances in Mathematics Education. Approaches to qualitative research in mathematics education: Examples of methodology and methods* (pp. 365–380). Springer. https://doi.org/10.1007/978-94-017-9181-6_13

McGuirl, J. M., & Sarter, N. B. (2006). Supporting trust calibration and the effective use of decision aids by presenting dynamic system confidence information. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *48*(4), 656–665. https://doi.org/10.1518/001872006779166334

Meixner, G., Häcker, C., Decker, B., Gerlach, S., Hess, A., Holl, K., Klaus, A., Lüddecke, D., Mauser, D., Orfgen, M., Poguntke, M., Walter, N., & Zhang, R. (2017). Retrospective and Future Automotive Infotainment Systems—100 Years of User Interface Evolution. In G. Meixner & C. Müller (Eds.), *Human–Computer Interaction Series. Automotive User Interfaces* (pp. 3–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-49448-7_1

Meixner, G., & Müller, C. (Eds.). (2017). *Human–Computer Interaction Series*. *Automotive User Interfaces*. Springer International Publishing. https://doi.org/10.1007/978-3-319-49448-7

Mercedes-Benz. (2022). *DRIVE PILOT: Zukunft erfahren: Mit dem aktuellen Höchstmaß an Assistenz und Automatisierung Ihre Ziele sicher und entspannter erreichen.* Mercedes-Benz. https://www.mercedes-benz.de/passengercars/technology/drive-pilot.html

Mirnig, A. G., Meschtscherjakov, A., Wurhofer, D., Meneweger, T., & Tscheligi, M. (2015). A Formal Analysis of the ISO 9241-210 Definition of User Experience. In B. Begole, J. Kim, K. Inkpen, & W. Woo (Eds.), *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 437–450). ACM. https://doi.org/10.1145/2702613.2732511

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). Visual attention and steering wheel control: From engagement to disengagement of Tesla Autopilot. *Proceedings of the Human*

Factors and Ergonomics Society Annual Meeting, *65*(1), 1390–1394. https://doi.org/10.1177/1071181321651118

Moser, K. S. (2003). Mentale Modelle und ihre Bedeutung: Kongintionspsychologische Grundlagen des (Miss-)Verstehens. http://www.symbolforschung.ch/files/pdf/karin_moser.pdf

Mühlhausen, J. (2020). *Erstellung und Evaluierung eines Informationskonzepts für das automatisierte Fahren* [Masterarbeit]. Technische Universität München, Garching bei München.

Muthumani, A [Arun], Diederichs, F., Galle, M., Schmid-Lorch, S., Forsberg, C., Widlroither, H., Feierle, A., & Bengler, K. (2020). How Visual Cues on Steering Wheel Improve Users' Trust, Experience, and Acceptance in Automated Vehicles. In N. A. Stanton (Ed.), *Advances in Intelligent Systems and Computing. Advances in Human Aspects of Transportation* (Vol. 1212, pp. 186–192). Springer International Publishing. https://doi.org/10.1007/978-3-030-50943-9_24

Naujoks, F., Forster, Y., Wiedemann, K., & Neukum, A. (2017a). A Human-Machine Interface for Cooperative Highly Automated Driving. In N. A. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Eds.), *Advances in Intelligent Systems and Computing. Advances in Human Aspects of Transportation* (Vol. 484, pp. 585–595). Springer International Publishing. https://doi.org/10.1007/978-3-319-41682-3_49

Naujoks, F., Forster, Y., Wiedemann, K., & Neukum, A. (2017b). Improving Usefulness of Automated Driving by Lowering Primary Task Interference through HMI Design. *Journal of Advanced Transportation*, *2017*(3), 1–12. https://doi.org/10.1155/2017/6105087

Naujoks, F., Hergeth, S., Wiedemann, K., Schömig, N., Forster, Y., & Keinath, A. (2019). Test procedure for evaluating the human-machine interface of vehicles with automated driving systems. *Traffic Injury Prevention*, *20*(sup1), S146-S151. https://doi.org/10.1080/15389588.2019.1603374

Naujoks, F., Hergeth, S., Wiedemann, K., Schömig, N., & Keinath, A. (2018). Use Cases for Assessing, Testing, and Validating the Human Machine Interface of Automated Driving Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1873–1877. https://doi.org/10.1177/1541931218621426

Naujoks, F., Kiesel, A., & Neukum, A. (2016). Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance. *Accident; Analysis and Prevention*, *97*, 162–175. https://doi.org/10.1016/j.aap.2016.09.009

Naujoks, F., Wiedemann, K., & Schömig, N. (2017). The Importance of Interruption Management for Usefulness and Acceptance of Automated Driving. In S. Boll, B. Pfleging, I. Politis, D. Large, & B. Domnez (Eds.), *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17* (pp. 254–263). ACM Press. https://doi.org/10.1145/3122986.3123000

Nielsen, T. A. S., & Haustein, S. (2018). On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transport Policy*, *66*, 49–55. https://doi.org/10.1016/j.tranpol.2018.03.004

Nordhoff, S., van Arem, B., & Happee, R. (2016). Conceptual Model to Explain, Predict, and Improve User Acceptance of Driverless Podlike Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, *2602*(1), 60–67. https://doi.org/10.3141/2602-08

Norman, D. A. (1987). Some observations on mental models. In R. M. Baecker & W. A. S. Buxton (Eds.), *Readings in human-computer interaction: A multidisciplinary approach* (pp. 241–244). Kaufmann.

Ochs, M., Pelachaud, C., & Sadek, D. (2008). An Empathic Virtual Dialog Agentto Improve Human-Machine Interaction. In L. Padgham (Ed.), *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 1: Volume 1* (Vol. 7, pp. 89–96). International Foundation for Autonomous Agents and Multiagent Systems.

Oliveira, L., Burns, C., Luton, J., Iyer, S., & Birrell, S. (2020). The influence of system transparency on trust: Evaluating interfaces in a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *72*, 280–296. https://doi.org/10.1016/j.trf.2020.06.001

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car. In A. L. Kun, L. Boyle, B. Reimer, A. Riener, J. Healey, W. Zhang, B. Pfleging, & M. Kurz (Eds.), *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 51–58). ACM. https://doi.org/10.1145/2390256.2390264

Panagiotopoulos, I., & Dimitrakopoulos, G. (2018). An empirical investigation on consumers' intentions towards autonomous driving. *Transportation Research Part C: Emerging Technologies*, *95*, 773–784. https://doi.org/10.1016/j.trc.2018.08.013

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Park, E., & Kim, K. J. (2014). Driver acceptance of car navigation systems: integration of locational accuracy, processing speed, and service and display quality with technology acceptance model. *Personal and Ubiquitous Computing*, *18*(3), 503–513. https://doi.org/10.1007/s00779-013-0670-2

Pettersson, I., & Karlsson, I. M. (2015). Setting the stage for autonomous cars: a pilot study of future autonomous driving experiences. *IET Intelligent Transport Systems*, *9*(7), 694–701. https://doi.org/10.1049/iet-its.2014.0168

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. In J. Häkkila & T. Ojala (Eds.), *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia - MUM '16* (pp. 91–99). ACM Press. https://doi.org/10.1145/3012709.3012735

Rahman, M. M. (2016). *Driver acceptance of advanced driver assistance systems and semi-autonomous driving systems* [Dissertation]. Mississippi State University, Mississippi.

Rahman, M. M., Lesch, M. F., Horrey, W. J., & Strawderman, L. (2017). Assessing the utility of TAM, TPB, and UTAUT for advanced driver assistance systems. *Accident; Analysis and Prevention*, *108*, 361–373. https://doi.org/10.1016/j.aap.2017.09.011

Regan, M. A., Horberry, T., & Stevens, A [Alan] (Eds.). (2018). *Driver Acceptance of New Technology*. CRC Press. https://doi.org/10.1201/9781315578132

Richardson, N. T [Natalie T.], Flohr, L., & Michel, B. (2018). Takeover Requests in Highly Automated Truck Driving: How Do the Amount and Type of Additional Information Influence the Driver–Automation Interaction? *Multimodal Technologies and Interaction*, *2*(4), 68. https://doi.org/10.3390/mti2040068

Rödel, C., Stadler, S., Meschtscherjakov, A., & Tscheligi, M. (2014). Towards Autonomous Cars. In E. Miller (Ed.), *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1–8). ACM. https://doi.org/10.1145/2667317.2667330

Rühmann, H. (1993). Schnittstellen in Mensch-Maschine-Systemen. In H. Schmidtke (Ed.), *Ergonomie.* Hanser.

Saad, F. (2004). Behavioural adaptations to new driver support systems: some critical issues. In P. Wieringa (Ed.), *2004 IEEE international conference on systems, man & cybernetics theme: Impacts of emerging cybernetics and human-machine systems : conference proceedings* (pp. 288–293). IEEE. https://doi.org/10.1109/ICSMC.2004.1398312

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016_201806).*

Schmidtke, H. (Ed.). (1993). *Ergonomie*. Hanser.

Schömig, N., Wiedemann, K., Naujoks, F., Hergeth, S., Keinath, A., & Neukum, A. (2020). How Important is the Plausibility of Test Scenarios Within Usability Studies for AV HMI? In N. A.

Stanton (Ed.), *Advances in Intelligent Systems and Computing. Advances in Human Aspects of Transportation* (Vol. 1212, pp. 77–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-50943-9_10

Schömig, N., Wiedemann, K., Naujoks, F., Neukum, A., Leuchtenberg, B., & Vöhringer-Kuhnt, T. (2018). An Augmented Reality Display for Conditionally Automated Driving. In Unknown (Ed.), *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (pp. 137–141). ACM Press. https://doi.org/10.1145/3239092.3265956

Shani, G., Rokach, L., Shapira, B., Hadash, S., & Tangi, M. (2013). Investigating confidence displays for top- N recommendations. *Journal of the American Society for Information Science and Technology*, *64*(12), 2548–2563. https://doi.org/10.1002/asi.22934

Shneiderman, B., Cohen, M., Diakopoulos, N., Elmqvist, N., Jacobs, S., & Plaisant, C. (2018). *Designing the user interface: strategies for effective human-computer interaction* (Sixth edition). *Global edition*. Pearson Educación. https://doi.org/utor

Siebert, F. W., Oehl, M., Höger, R., & Pfister, H.-R. (2013). Discomfort in Automated Driving – The Disco-Scale. In C. Stephanidis (Ed.), *Communications in Computer and Information Science: Vol. 374. HCI International 2013 - Posters' Extended Abstracts* (Vol. 374, pp. 337–341). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39476-8_69

Stanton, N. A., & Young, M. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, *1*(4), 315–331. https://doi.org/10.1080/14639220052399131

Steidl, B. (2020). *Auswirkungen der Anzeige der Dauer bis zur Verfügbarkeit und der Gründe für die aktuelle Nichtverfügbarkeit automatisierter Fahrfunktionen während manueller Fahrt* [Masterarbeit]. Technische Universität München, Garching bei München.

Stevens, A [A.], Quimby, A., Board, A., Kersloot, T [T.], & Burns, P. (Eds.). (2002). *DESIGN GUIDELINES FOR SAFETY OF IN-VEHICLE INFORMATION SYSTEMS.*

Stockert, S., Richardson, N. T [Natalie Tara], & Lienkamp, M. (2015). Driving in an Increasingly Automated World – Approaches to Improve the Driver-automation Interaction. *Procedia Manufacturing*, *3*, 2889–2896. https://doi.org/10.1016/j.promfg.2015.07.797

Timpe, K.-P., & Baggen, R. (Eds.). (2000). *Mensch-Maschine-Systemtechnik: Konzepte, Modellierung, Gestaltung, Evaluation* (1. Aufl., Stand: Oktober 2000). Symposion Publ.

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

Venkatesh, Morris, & Davis (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vorm, E. S., & Combs, D. J. Y. (2022). Integrating Transparency, Trust, and Acceptance: The Intelligent Systems Technology Acceptance Model (ISTAM). *International Journal of Human-Computer Interaction*, 1–18. https://doi.org/10.1080/10447318.2022.2070107

Wachenfeld, W., Winner, H., Gerdes, C., Lenz, B., Maurer, M [M.], Beiker, S. A., Fraedrich, E., & Winkle, T [T.]. (2015). Use-Cases des autonomen Fahrens. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomes Fahren* (pp. 9–37). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45854-9_2

Walker, G. H., Stanton, N. A., & Salmon, P. (2016). Trust in vehicle technology. *International Journal of Vehicle Design*, *70*(2), Article 74419, 157. https://doi.org/10.1504/IJVD.2016.074419

Wallentowitz, H., & Reif, K. (Eds.). (2011). *ATZ/MTZ-Fachbuch. Handbuch Kraftfahrzeugelektronik: Grundlagen - Komponenten - Systeme - Anwendungen* (2., verb. und aktualisierte Aufl.). Vieweg + Teubner.

Wandtner, B., Schömig, N., & Schmidt, G. (2018). Secondary task engagement and disengagement in the context of highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 253–263. https://doi.org/10.1016/j.trf.2018.06.001

Warg, F., Ursing, S., Kaalhus, M., & Wiik, R. (2020). Towards Safety Analysis of Interactions Between Human Users and Automated Driving Systems. In *10th European Congress on Embedded Real Time Software and Systems (ERTS 2020),* Toulouse, France.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Winkle, T [Thomas]. (2016). Safety Benefits of Automated Vehicles: Extended Findings from Accident Research for Development, Validation and Testing. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous Driving* (pp. 335–364). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-48847-8_17

Wintersberger, P., Frison, A.-K., Riener, A., & Sawitzky, T. von (2018). Fostering User Acceptance and Trust in Fully Automated Vehicles: Evaluating the Potential of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *27*(1), 46–62. https://doi.org/10.1162/pres_a_00320

Yang, C. C. (2005). The refined Kano's model and its application. *Total Quality Management & Business Excellence*, *16*(10), 1127–1137. https://doi.org/10.1080/14783360500235850

Yang, Y., Karakaya, B., Dominioni, G. C., Kawabe, K., & Bengler, K. (2018). An HMI Concept to Improve Driver's Visual Behavior and Situation Awareness in Automated Vehicle. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 650–655). IEEE. https://doi.org/10.1109/ITSC.2018.8569986

Yuen, K. F., Cai, L., Qi, G., & Wang, X. (2021). Factors influencing autonomous vehicle adoption: an application of the technology acceptance model and innovation diffusion theory. *Technology Analysis & Strategic Management*, *33*(5), 505–519. https://doi.org/10.1080/09537325.2020.1826423

Zhang, T., Zeng, W., Zhang, Y., Da Tao, Li, G., & Qu, X. (2021). What drives people to use automated vehicles? A meta-analytic review. *Accident; Analysis and Prevention*, *159*, 106270. https://doi.org/10.1016/j.aap.2021.106270

Zoellick, J. C., Kuhlmey, A., Schenk, L., Schindel, D., & Blüher, S. (2019). Assessing acceptance of electric automated vehicles after exposure in a realistic traffic environment. *PloS One*, *14*(5), e0215969. https://doi.org/10.1371/journal.pone.0215969

Appendix

Demographic questionnaire

Persönliche Angaben

Probanden-ID:	______________________
Datum:	______________________
Alter:	______________________
Geschlecht:	______________________
Höchster Bildungsabschluss:	______________________
Wie lange sind Sie in Besitz Ihres Führerscheins:	______________________

Haben Sie bereits an einer oder mehreren Studien in einem Fahrsimulator teilgenommen? ☐ja ☐nein

- Falls ja, wie oft? ____________________________
- Wie viele standen mit automatisiertem Fahren in Verbindung? ____________________________

Fahrgewohnheiten

Wie viele Kilometer fahren Sie in etwa jährlich? ______________ Kilometer pro Jahr

	Täglich	Mehrmals pro Woche	Einmal pro Woche	Einmal im Monat	Seltener
Wie regelmäßig fahren Sie Autobahn?	☐	☐	☐	☐	☐

Wie gern oder ungern fahren Sie auf der Autobahn? Aus welchem Grund

sehr gern ☐ ☐ ☐ ☐ ☐ sehr ungern

Grund:

____

____

Ihre Durchschnittgeschwindigkeit, wenn unbegrenzt _________________ km/h

Questions created for assessing Kano (Fundin & Nilsson, 2003) categories

Wie würden Sie sich fühlen, wenn das Feature "Zeit bis die Automation verfügbar ist" vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Zeit bis die Automation verfügbar ist" NICHT vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature „PopUp mit Grund für Nichtaktivierbarkeit“ vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "PopUp mit Grund für Nichtaktivierbarkeit" NICHT vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Verfügbarkeitsdauer" vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Verfügbarkeitsdauer" NICHT vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Anzeige aller Verfügbarkeitsabschnitte im Mitteldisplay" vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Anzeige aller Verfügbarkeitsabschnitte im Mitteldisplay" NICHT vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Gründe für Nichtverfügbarkeit im Mitteldisplay" (Icons und Text) vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Wie würden Sie sich fühlen, wenn das Feature "Gründe für Nichtverfügbarkeit im Mitteldisplay" (Icons und Text) NICHT vorhanden wäre?

Das würde mich sehr freuen	Das setze ich voraus	Das ist mir egal	Das könnte ich akzeptieren	Das würde mich sehr stören

Acceptance Score (van der Laan et al., 1997)

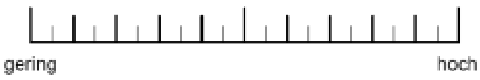
Bitte beurteilen Sie jetzt das System. Lesen Sie hierfür aufmerksam jedes Wortpaar und machen Sie jeweils ein Kreuz pro Zeile.

Nützlich						Nutzlos
Angenehm						Unangenehm
Schlecht						Gut
Nett						Nervig
Effektiv						Unnötig
Ärgerlich						Erfreulich
Hilfreich						Wertlos
Nicht wünschenswert						Wünschenswert
Aktivierend						Einschläfernd

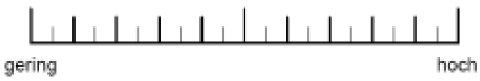
Nasa rTLX (Hart, 2006)

Geben Sie jetzt für jede der unten stehenden Dimensionen an, wie hoch die Beanspruchung war. Markieren Sie dazu bitte auf den folgenden Skalen, in welchem Maße Sie sich in den sechs genannten Dimensionen von der Aufgabe beansprucht oder gefordert gesehen haben.

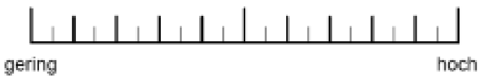
Geistige Anforderungen: Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen…)? War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erforderte sie hohe Genauigkeit oder war sie fehlertolerant?



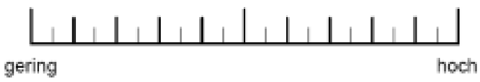
Körperliche Anforderungen: Wie viel körperliche Aktivität war erforderlich (z.B. Ziehen, Drücken, Drehen, Steuern, Aktivieren,…)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?



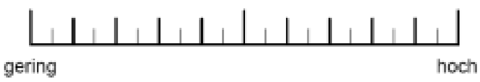
Zeitliche Anforderungen: Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt, mit dem Aufgaben oder Aufgabenelemente auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?



Leistung: Wie erfolgreich haben Sie Ihrer Meinung nach die vom Versuchsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?



Anstrengung: Wie hart mussten sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?



Frustration: Wie unsicher, entmutigt, irritiert, gestresst und verärgert (versus sicher, bestätigt, zufrieden, entspannt und zufrieden mit sich selbst) fühlten Sie sich während der Aufgabe?

System Usability Scale (Brooke, 1996)

1. Ich denke, dass ich das System gerne häufig benutzen würde.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

2. Ich fand das System unnötig komplex.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

3. Ich fand das System einfach zu benutzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

4. Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

5. Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

6. Ich denke, das System enthielt zu viele Inkonsistenzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

7. Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

8. Ich fand das System sehr umständlich zu nutzen.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

9. Ich fühlte mich bei der Benutzung des Systems sehr sicher.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5

10. Ich musste eine Menge lernen, bevor ich anfangen konnte das System zu verwenden.

Stimme überhaupt nicht zu 1	2	3	4	Stimme voll zu 5