6-18-2022

The Hazard Potential of Non-Driving-Related Tasks in Conditionally Automated Driving

Moritz Markgraf
University of Augsburg, moritz.markgraf@fim-rc.de

Marco Schmidt
University of Applied Sciences Augsburg, marco.schmidt@fim-rc.de

Niklas Wolf
University of Augsburg, niklas.wolf@fim-rc.de

Leon Adler
University of Augsburg, leon.adler@fim-rc.de

THE HAZARD POTENTIAL OF NON-DRIVING-RELATED TASKS IN CONDITIONALLY AUTOMATED DRIVING

Research Paper

Moritz Markgraf, FIM Research Center, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, University of Augsburg, moritz.markgraf@fim-rc.de

Marco Schmidt, FIM Research Center, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, University of Applied Sciences Augsburg, marco.schmidt@fim-rc.de

Niklas Wolf, FIM Research Center, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, University of Augsburg, niklas.wolf@fim-rc.de

Leon Adler, FIM Research Center, University of Augsburg, leon.adler@fim-rc.de

Abstract

Today, humans and machines successfully interact in a multitude of scenarios. Facilitated by advancements in artificial intelligence, increasing driving automation may allow drivers to focus on non-driving-related tasks (NDRTs) during the automated ride. However, conditionally automated driving as a transitional state between human-operated driving and fully automated driving requires drivers to take over control of the vehicle whenever requested. Thus, the productive use of driving time might come at the cost of increased traffic safety risks due to insufficient and insecure human-vehicle interaction. This study aims to explore the take-over performance and risk potential of different NDRTs (auditory task, visual task on regular display, visual task with mixed reality hardware) while driving. Our study indicates the hazard potential of visual vs. auditory distraction and multitasking vs. sequential tasking. Our findings contribute to understanding what influences the acceptance and adoption of automated driving and inform the design of safe vehicle-human take-overs.

Keywords: Automated Driving, Human-Vehicle Interaction, Multitasking Behavior.

1 Introduction

Following years of digitalization, we today live in a world in which humans and machines successfully interact in a multitude of scenarios. In recent years, many hybrid human-machine settings, for example, in the contexts of robotics or (mobile) personal support systems, have reached a productive state and greatly impact individuals, organizations, and society. Carried by technological advancements, the socio-technological transformation we currently go through has changed the way humans live and work. New human-machine interfaces let people perform their tasks more efficiently and even allow for the simultaneous execution of multiple tasks. But multitasking is a double-edged sword. Although it can be viewed as an efficiency boost, multitasking may cause reduced performance in each of the single tasks (Miller and Durst, 2014). In some settings, this may entail severe consequences. A prominent example is the current transitional state between fully human-operated driving and fully automated driving (Stephanidis et al., 2019). In this transitional state, commonly referred to as conditionally automated driving, the artificially intelligent vehicle drives autonomously in most scenarios but may request the driver to take over control in unknown or complex traffic situations. Currently, the traffic regulations in most countries require human drivers to fully focus on the road, permitting drivers only a few exceptions for non-driving-related tasks (NDRTs) which cause limited cognitive load, for example, hands-free phone calls or listening to the radio (Charlton, 2009). Due to reduced requirements regarding the driver's attention, increased driving automation creates demand for a wider scope of action for individuals operating conditionally automated vehicles (Stephanidis et al., 2019). Recently, the United Nations (UN) Regulation No. 157 (United Nations, 2021), which has become effective in the European Union, Japan, and Canada as of January 2021, has laid the cornerstone for greater flexibility in the execution of NDRTs while driving. Extending current traffic regulations, the new regulation allows drivers of conditionally automated vehicles during the automated ride to shift their attention away from the road and towards other visually demanding tasks.

While this allows productive use of otherwise unproductive driving time, the risks associated with this regulatory change cannot yet be fully assessed. A decade ago, when driving automation was significantly less advanced, the WHO found that mobile phone use during driving increases the risk of car crashes by a factor of four (World Health Organization, 2011). Although conditionally automated vehicles are equipped with powerful assistance systems that can prevent crashes and increase driving safety in many scenarios, these vehicles cannot yet respond adequately to all conceivable traffic situations. Instead, they have a defined automation zone in which they are confident to work reliably (e.g., on the highway). But whenever the vehicle is about to leave this automation zone, the vehicle requests and initiates a take-over to the human driver, causing a safety-critical interaction between the human and the vehicle. As of now, take-overs between the vehicle and the driver are still a regular occurrence. Therefore, drivers need to be ready to take control of the vehicle whenever requested, as the vehicle may come into a critical situation it cannot solve on its own. Thus, if it is not ensured that drivers can react in time, the productive use of driving time might come at the cost of increased safety risks (Seppelt and Victor, 2016).

These safety risks do not remain unnoticed by potential users. Various studies focus on the acceptance and adoption of automated driving list safety concerns and a lack of trust among the most relevant inhibitors of automated vehicle acceptance (Bornholt and Heidt, 2019; Ernst and Reinelt, 2017; Hein et al., 2018). An extended version of the technology acceptance model includes trust and reliability as key factors for the adoption of safety-critical technology applications such as automated driving (Hutchins and Hook, 2017). Although the possibility to take control of the vehicle may increase trust (Lackes et al., 2020; Keller et al., 2021), performant vehicle-human take-overs are one of the major challenges of conditionally automated driving (Casner et al., 2016). While drivers executing an NDRT are typically less attentive to the road complicating the take-over (NHTSA, 2013), different NDRTs may have different effects on the driver's reaction (Dogan et al., 2019; Radlmayr et al., 2019; Zhang et al., 2019). To produce relevant knowledge on the safety risks associated with different NDRTs, the present study contrasts the take-over performance of three NDRTs. These three NDRTs differ in regard to their multitasking potential (sequential tasking vs. multitasking) and sensory input channel (auditory vs. visual NDRT). Our research can unravel the hazard potential of various NDRTs, formatively inform traffic regulations with respect to driving safety, and indicate regulative rules increasing trust in automated vehicles. The study elaborates on the research question:

How do different non-driving-related tasks during conditionally automated driving influence the driver's take-over performance in critical situations?

We conduct a within-subjects behavioral experiment investigating the effects of different NDRTs on the driver's take-over performance in various driving scenarios. Participants operate a simulated automated vehicle that drives autonomously in these scenarios and requests the participant to take control of the vehicle in a critical traffic situation. We evaluate the driver's reaction to the situation comparing auditory and visual tasks at different levels of multitasking potential. The selection of these tasks is aligned with what is permitted by currently effective traffic regulations in the European Union. The remainder of this paper is structured as follows. Section 2 establishes a common understanding of the theoretical background and derives hypotheses. Section 3 presents the experimental design, procedure, and measurements. Section 4 describes the statistical results. Section 5 discusses the findings, contributions, and limitations before section 6 concludes with an outlook on further research.

2 Theoretical Background and Hypotheses Development

Automated driving refers to the idea that vehicles move without the intervention of a human driver (SAE, 2016). For this, automated vehicles need to perceive their surroundings, interact with the environment, handle incomplete information, and adapt to unpredictable changes (Kolodko and Vlacic, 2003). With the constant evolution of technologies such as artificial intelligence, machines are on the verge of effectively performing these functionalities (Biondi et al., 2019). But despite promising developments, fully automated vehicles not requiring a driver ready to take control are still dreams of the future. The SAE J3016 standard (SAE, 2016) defines six maturity levels (0 to 5) as stages of development to fully automated driving (level 5). As of now, conditionally automated driving (level 3) is becoming reality and automated vehicles can perform the entire dynamic driving task within a certain *automation zone*. Level 3 is the first level allowing drivers to reduce the attention devoted to the traffic situation (Marinik et al., 2014; Trimble et al., 2014) and to occupy themselves with an NDRT (Gold et al., 2018b) such as reading a newspaper (Kuehn et al., 2017) or playing a video game (Wu et al., 2020). However, in case of errors or events that leave the defined automation zone, the vehicle cannot guarantee to lead to a state of minimal risk (Tellis et al., 2016). Hence, the driver must be prepared that the vehicle may warn them and hands over control to them with a sufficient time budget to manage the situation (Kuehn et al., 2017). Since the time budget is a crucial factor, critical (Lotz et al., 2020) or unplanned (Lee and Yang, 2020) take-over requests following unforeseeable events like a technical issue or a disabled car ahead are of special interest (Lee and Yang, 2020; Lotz et al., 2020). Consequently, this interaction between driver and vehicle, the *take-over*, is vital for the driver's as well as for the other traffic users' safety. It needs to enable the driver to intervene as soon as possible to prevent a crash, for example, by braking or changing the lane (Radlmayr et al., 2019).

The take-over from an automated vehicle to its driver is one of automated driving's utmost challenges (Kuehn et al., 2017; Radlmayr et al., 2019; Zhang et al., 2019) and is subject to thorough scientific investigation. Thus, it requires a carefully designed human-vehicle interface that can enable an effective take-over (Lee and Yang, 2020). In the event of a take-over, the driver must build up the necessary awareness of the situation as quickly as possible and act accordingly. This process can be divided into three steps, where first the driver perceives visual objects, then processes them based on the individual goals and make a prediction for future states and events, and finally acts based on the individual goals and predictions (Wolfe et al., 2019). Taking this process into account, a central concept to validate the take-over is the *take-over performance* (TOP). Adhering to Gold et al. (2018a), we refer to TOP as the interplay of various factors of an individual's response to a take-over request. This includes the *take-over time*, the *minimum time to collision*, brake application, and crash incidence. Previous research has examined influencing factors on the TOP such as human factors like age (Wu et al., 2020), trust in automation (Hergeth et al., 2016), or driving experience (e.g., Lotz et al., 2020).

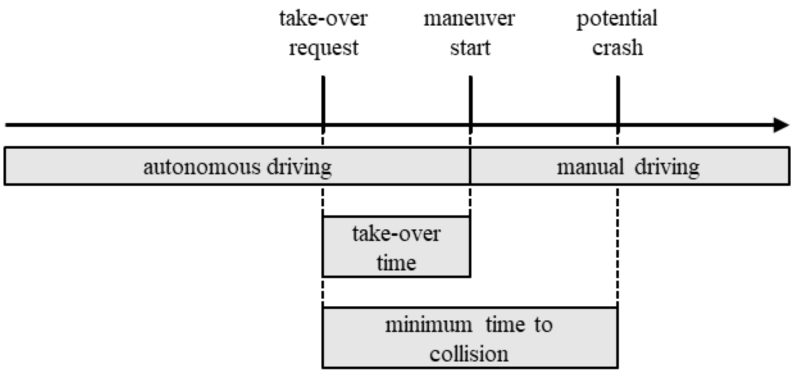

Figure 1. Visual illustration of take-over (inspired by Zhang et al., 2019))

Previous studies investigating the TOP (e.g., Kuehn et al., 2017; Marberger et al., 2018; Lee and Yang, 2020) mutually agree that the time budget available for the driver to acquire control of the vehicle after a take-over request is crucial to prevent crashes. Thus, they demand that the vehicle needs to inform the driver as soon as possible. However, the time budgets tested in previous studies differ greatly (Zhang et al., 2019) ranging from three (Vogelpohl et al., 2016) to 15 seconds (Kuehn et al., 2017). The time budget is indeed essential for the safety in level 3 automated driving, yet the take-over time depends on multiple factors (Zeeb et al., 2015), specifically, the driver, the environment, the vehicle, and the human-vehicle interface (Vogelpohl et al., 2016). In practice, the available time budget is limited naturally as a take-over request can only start as early as an event occurs that makes the vehicle leave its specified automation zone. Therefore, independently of the specified threshold, there will be situations in which there is not enough time to prevent a crash. Hence, conditionally automated driving needs to strive for the quickest possible reaction to prevent severe outcomes.

The TOP depends on various factors (Borojeni et al., 2017; Lee and Yang, 2020; Kujala et al., 2021). For instance, the environment inside (Horberry et al., 2006) and outside of the vehicle such as the criticality of the traffic situation (Lotz et al., 2020) affect the driving performance. Additionally, physiological factors like age (Wood, 2002), fatigue (Philip et al., 2005), or the driver's speed in processing information (Roenker et al., 2003) may influence driving. More generally, the TOP depends on the driver's mental model as there is the need to evaluate the environment including several sources of uncertainty like the intentions of other road users (Brechtel et al., 2014). There are different approaches to model the overall task difficulty of both single tasks and multiple tasks. One of these theories is the mental workload theory, which assumes that a person has limited mental resources available, which can be partially or fully demanded by single or multiple tasks depending on the task difficulty. Comparing the available mental resources with the demand for mental resources for one or multiple tasks there is a critical region if the demand of the tasks is higher than the available resources. In this case, the performance of the tasks will degrade, but it is uncertain which task performance will degrade and to what extent (Wickens, 2008). When focusing solely on driving, the driver continuously – in part subconsciously – assesses the environment, for instance, knowing automatically that they pass a traffic light. Therefore, the driver possesses some situational awareness which refers to the driver's awareness "of the current and future driving conditions facing the vehicle" (Petersen et al., 2019, p. 5). In the case of performing an NDRT during automated driving level 3, the driver does not focus on the environment and, hence, has a lower situational awareness (Atwood et al., 2019; Marberger et al., 2018). Therefore, the driver needs to process more data to obtain the same level of awareness as a driver that mainly focuses on driving. In the context of the mental workload theory (Wickens, 2008), the execution of an NDRT while driving will lead to an increase in the demand for mental resources at the moment of the take-over request. If the demand for mental resources is higher than the available mental resources, it leads to a degradation of the respective task performance (i.e., the TOP). With respect to situational awareness, the task of generating situational awareness may compete with the NDRTs for the supplied mental resources, which means that a high mental workload may lead to worse situational awareness (Tsang and Vidulich, 2006). Therefore, we argue that the evaluation of the environment will not be of the same quality. Consequently, performing an NDRT likely affects the TOP. Following previous research (e.g., Dogan et al., 2019; Kun et al., 2017) and to establish a common ground, we hypothesize as follows.

H1: Performing an NDRT negatively affects the TOP.

Previous empirical studies have investigated the effects on the TOP utilizing various NDRTs such as different distraction tests (Lee et al., 2021), writing emails (Dogan et al., 2019), reading a newspaper (Kuehn et al., 2017), playing a video game (Kuehn et al., 2017; Wu et al., 2020), or watching a video clip (Dogan et al., 2019; Wu et al., 2020). Performing those NDRTs requires the driver to primarily focus on the NDRT instead of the traffic. In the case of a take-over request, the driver stops the NDRT and focuses on the vehicle and the traffic environment. Hence, those NDRTs follow the paradigm of sequential tasking as the driver either focuses on the NDRT or the driving. By definition, level 3 automated driving does not require the driver to devote attention to the automated driving at any time and hence, sequential tasking is presumably compatible with level 3 automated vehicles. However, presently with vehicles up to level 2 automation, drivers may only perform NDRTs that allow parallel monitoring of the driving, for example, listening to a podcast or talking to another passenger. These tasks refer to the paradigm of multitasking which means that both tasks (i.e., NDRT and driving) are executed simultaneously (Goldstein, 2015). We, thus, call such NDRTs multitasking-compatible. While it is scientifically disputable if this is real multitasking (Goldstein, 2015), in general, multitasking has an arguable bad reputation as it leads to worse performances of each task when compared to a sequential execution (Buser and Peter, 2012). Nevertheless, in the context of manual driving, previous research (Villalobos-Zúñiga et al., 2016) provides evidence that the integration of a head-up display, a mixed reality device commonly used in vehicles, decreases the distraction of the investigated NDRT and has a rather small effect on the driving performance. Adapting this to the context of conditionally automated driving, multitasking-compatible NDRTs might allow the driver to evaluate the traffic environment to some extent and hence, obtain situational awareness. This information head start might be an advantage in the evaluation process at a take-over request. Therefore, we pose that it leads to a less time-consuming or a more precise evaluation ultimately yielding a better TOP.

H2: Performing a multitasking-compatible NDRT has a weaker negative effect on the TOP than performing a sequential NDRT.

The literature distinguishes four forms of driving distraction (Pettitt et al., 2009; World Health Organization, 2011): visual distraction, which refers to a shift of the visual focus away from the road, auditory distraction caused by focusing on auditory signals, cognitive distraction, denoting a shift of the cognitive focus away from the road, and biomechanical distraction, which is the case when the driver takes his hands off the human-vehicle interface (e.g., the steering wheel). Some studies suggest that there are no significant differences in the TOP between NDRTs that allow the driver to visually monitor the driving environment, for example, using a HoloLens for video calls during the automated ride (Kun et al., 2017), and such that also demand their full visual attention (Radlmayr et al., 2019). However, this is disputable because eyesight is a crucial factor for driving performance (Roenker et al., 2003; Wood, 2002). As people usually cannot focus their visual sight on two different things simultaneously (Duncan, 1984), the driver cannot glimpse at the environment when performing an NDRT demanding visual attention. Even if he could see and process these two different visual sources at the same time, for example, with the use of a HoloLens, it is likely that the task performances will degrade, since both tasks inherit both a visual and a perceptual-cognitive activity causing a higher mental workload because of the similarity of the tasks in means of the resource demands (Wickens, 2008; Tsang and Vidulich, 2006). Hence, at a take-over request, the driver must start out of the blue to evaluate the environment leading to an increase in the demanded mental resources. Consequently, the take-over time rises when performing a visual NDRT (Zhang et al., 2019). The lab experiment of Radlmayr et al. (2019) confirms this by showing a significant deterioration of driver availability when performing a visual-motoric NDRT. Furthermore, the driver availability does not decrease for cognitive (e.g., listening to a podcast) or physical NDRT (Radlmayr et al., 2019). Hence, we hypothesize as follows.

H3: Performing a non-visual NDRT has a weaker negative effect on TOP than performing a visual NDRT.

Overall, in H1, H2, and H3, we pose that greater situational awareness can lead to a better TOP and NDRTs reduce situational awareness. However, previous empirical research is inconsistent regarding the effect of NDRT. In some studies, there are significant effects (Zhang et al., 2019) whereas others cannot identify a clear role for NDRTs (e.g., Dogan et al., 2019; Radlmayr et al., 2019). The results of Kun et al. (2017), for example, might be influenced by the limited use of the HoloLens in the experiment since in this study the NDRT was to participate in a video conference. Therein, it is arguable whether the participants in their experiment might have attended the video conference without visually focusing on the video stream but focused primarily on the driving task instead. Hence, we argue that attention to the NDRT influences the effects of NDRT and explains empirical inconsistencies. We state that an increase of attention to the NDRT while maintaining the same attention towards the driving task leads to a higher demand for mental resources. In mental workload theory, this can lead to a degradation of one or both task performances (Wickens, 2008). Accordingly, a high mental workload can have a negative impact on situational awareness which can cause a decrease in quality in decision making (Endsley, 1995) and thus also influence the TOP.

H4: High attention to the NDRT has a stronger negative effect on the TOP than low attention to the NDRT.

3 Research Methodology

Our within-subject behavioral laboratory experiment investigates the effect of different NDRTs on the driver's TOP in automated driving level 3. The participants operate a level 3 automated driving simulator and undergo three experimental conditions, each of which draws the participants' attention to a different NDRT – auditory multitasking-compatible task, visual multitasking-compatible task, visual sequential task – and one control condition with no NDRT. Corresponding to our hypotheses, our study aims to investigate if there are differences in the TOP (as an indicator for traffic safety) across the four conditions. Our experimental design is inspired by previous studies (e.g., Blanco et al., 2015; Kun et al., 2017; Radlmayr et al., 2019). It combines their ideas of investigating the TOP in automated driving level 3 and testing the applicability of a HoloLens in automated driving. We recruited participants from the authors' families and friends. Alike Underwood et al. (2013), we selected only participants with a valid driver's license and at least 2 years of driving experience whose job does not predominantly involve driving. Additionally, to assure the validity of the driving and NDRTs, we ensured that none of the participants possessed any kind of driving restriction or unresolved visual or auditory impairment. We did not provide any incentive for participation but asked the participants to support our scientific inquiry.

3.1 Procedure

As displayed in Figure 2, at the beginning of the experiment, each participant performs a test run of the three NDRTs: the visual sequential task which requires the participants to execute a visual attention test on an additional display, the visual multitasking-compatible task which requires the participant to perform the same test while having a HoloLens on allowing them to keep an eye on the road, and the auditory multitasking-compatible task which involves an auditory attention test. These test runs allow the participants to get accustomed to the tasks they later need to execute while driving and their controls. We refer to this phase as the learning phase. Subsequently, the participants traverse twelve driving scenarios (three scenarios for four different NDRT conditions). While the conditions appeared in randomized order, the driving scenarios for each condition were preserved in a fixed order. Following UN Regulation No. 157 (United Nations, 2021), NDRTs automatically stop as soon as the vehicle requests the driver to take control. Then, we capture and evaluate the driver's reaction to the traffic situation. After the scenarios, in the benchmarking phase, the participants only perform the NDRT without the driving task. Since all mental resources are available, this enables us to determine the individual maximum task performance and compare it to the task performances within scenarios with an NDRT. Finally, the participants answer a survey focusing on socio-demographics.

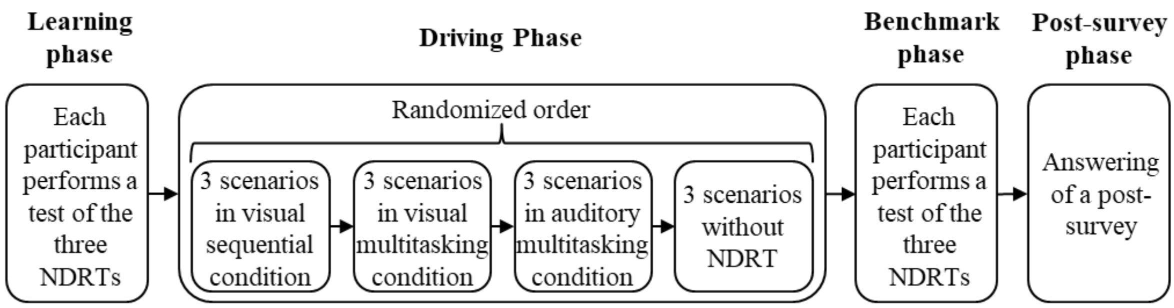

Figure 2. Experimental Procedure

3.2 Experimental Setup

In the course of the experiment, each participant operated a simulated automated driving level 3 vehicle through 12 driving scenarios (four conditions with three scenarios each). The participants operated the simulation visible on a 55" screen with the help of a steering wheel and brake, and accelerator pedals. At the beginning of the experiment, the participant was instructed in detail to only intervene when a take-over signal sounds. Until this signal, they should focus on the test and complete it as well as possible, whereby they are informed that after each scenario they will be asked questions about the traffic situation after the take-over signal sounds. In this way, we achieve that the drivers divided their attention between the driving scenario and a second task, just as in reality.

To create the driving scenarios, we implemented a technical backbone based on the CARLA Urban Driving Simulator from Dosovitskiy et al. (2017). In each of these scenarios, an automated vehicle steers autonomously through a city's scenery and traffic situations until the vehicle requests the driver to take over control. This take-over request triggers at some point in time which is fixed for each scenario but differs between the various scenarios. To reduce experimental bias, we expand this time slot of requesting the take-over compared to alike studies (e.g. Borojeni et al., 2017: 30 to 40 seconds) and request the take-over between 30 and 60 seconds after the start of each scenario. The vehicle does not require any interaction from the driver before the take-over request. All participants were presented with the same twelve scenarios in the same sequence.

When the participant is requested to take control, the vehicle keeps moving with unchanged velocity and direction for three seconds. Then the simulator captures the driving reaction of the participant which is made up of braking, accelerating, and steering. While all scenarios are unique, each block of three scenarios comprises similar traffic situations with one or two critical scenarios with a time to a collision of 3 seconds at the time of the take-over request and one or two scenarios with a lower criticality where no reaction of the participant was needed, so that the participants could not get used to simply braking hard after a take-over request but had to react validly. We decided to include these “fake scenarios” to address a limitation from Blanco et al. (2015) that having only scenarios which require an action might limit the participants’ cognitive effort and, thus, could distort the results. In selecting the critical scenarios, we follow the description of scenarios for automated driving level 3 of Gold et al. (2018b) and select scenarios allowing “maximum driver performance” by exposing the driver to a dangerous situation requiring a fast reaction like driving through a red light which requires immediately braking. While the participants were informed that not all scenarios required an intervention, they had no information about other patterns or the frequency of the scenarios. To prevent learning effects across similar scenarios, the participants did neither see the effect of their intervention nor receive any feedback on their performance. We informed the participants about this design decision before the experiment.

In three of the four NDRT conditions, the participant needs to perform a task based on an attention test. This task had to be done on the numeric keypad of a wireless keyboard. While the visual sequential condition used the same keyboard and input logic as the visual multitasking condition, instead of the HoloLens, an additional monitor was placed slightly diagonally below the large screen with the simulation in a similar ratio as a car’s navigation system to the windscreen. In the visual attention test, the participants had the opportunity to revise the input by repeatedly pressing the same key or to continue with the next test field by pressing the “0” key if the test field took less than 7 seconds to complete.

3.3 Measures

We determine the driver’s TOP in line with Gold et al. (2018a) from the measures of take-over time, brake application, and crash incidence. Take-over time refers to the “time *between* the take-over request and the maneuver start” (Gold et al., 2018a, p.3). It is measured as the time that elapses between the request and the participant’s first intervention. The brake application is a binary measure denoting whether the intervention includes braking. The crash incidence is a binary measure indicating whether the vehicle would have been involved in an accident following the participant’s intervention or not. Their fourth measure, the minimum time to collision, represents the minimum time remaining until a traffic accident happens in case all road users continue their current trajectories. Since we set the take-over request to be 3 sec in advance of a collision, the minimum time to collision is the result of 3 seconds minus the take-over time. Thus, we do not include the minimum time to a collision as a measure.

To contrast the TOP with the attention dedicated to the NDRTs, each NDRT implements an attention test. In the two conditions with visual NDRTs (visual sequential and visual multitasking), participants perform the d2 attention test (Schmidt-Atzert, 2004) which is a task that requires a visual focus and is a valid measure of individuals’ attention. The test contains a grid in which the letters ‘d’ and ‘p’ are arranged in rows and marked with one or two strokes. The task is to cross out all 'd's with two strokes but none else. For better readability in the HoloLens, we modified the test for the visual multitasking and sequential tasks in the way that we use fewer rows per test grid. In the condition of auditory multitasking, the participants perform a self-constructed audio-based task. To resemble the procedure in the d2 attention test, participants listen to an audiotape consisting of German words with minimal phonemic contrast, that is, words that differ only in small phonetic differences (Swadesh, 1936). Therefore, we selected pairs or tuples of words that start with one of the letters 'b', 'p', 'd', or 't' and can easily be mistaken with another word from the list (e.g., "Dart" and "Part"). The participants are instructed to press a specific key whenever they hear a word starting with one of the letters 'd' or 'b' and another key when they hear a word that starts with 'p' or 't'. This task is similar to the visual NDRTs in the way that it also requires the participant to be attentive to small visual, respectively audible, differences, alters stimuli that require different actions and involves a tactile element. For each NDRT run, we measure the maximum points achievable in that run and the points achieved. Points are given in the visual NDRT for all 'd's that should have been crossed out and in the auditory NDRT for every correctly classified word. In the control condition with no NDRT, there is no additional attention test so that the driver may dedicate their full attentive capacity to the traffic.

4 Results

In the lab experiment, we collected valid scenario data sets for a total of 14 participants – 4 females and 10 males. A preliminary power analysis assuming a twenty percent increase of take-over time in the NDRT conditions compared to the no NDRT condition with a baseline of 1 second, a power of 0.80 and an alpha of 0.05 suggested a minimum sample size of 10 participants. The age span of our sample is 19 to 52 years with an average of 25 years. In terms of driving experience, all participants have a driving license for a minimum of 2 and a maximum of 32 years (on average, 7 years). Overall, we collected 168 scenario data points (14 participants with 12 scenarios each). Among these data points, 13 data points across 8 different participants have been excluded from further analysis because the participant intervened before the take-over request. Of the remaining 155 valid scenario data points, 67 data points belong to low-criticality scenarios not demanding an immediate reaction from the driver. Hence, we focus on the remaining 88 scenario data points when analyzing the TOP for critical situations. Each data point contains the steering wheel angle as well as the brake and accelerator pedal position every 30 milliseconds. We derive the TOP based on these data. Additionally, each data point contains the attention devoted to the NDRT (i.e., the score achieved in the NDRT).

Table 1 illustrates the averages by the scenarios of the gathered data. Simplifying the comparison of all NDRTs combined and no NDRT, the column 'all NDRTs' shows the average of the three different NDRTs. Table 1 shows that the average take-over time rises when performing an NDRT. For critical scenarios, the participants applied the brake less frequently when performing an NDRT, yet vice versa for non-critical scenarios. The crash incidence rises by 4 % indicating that in 42 % of the scenarios with an NDRT a crash would have happened. Despite the trend in the aggregated measures for all NDRTs, the picture is rather ambiguous for the different NDRTs.

		no NDRT	all NDRTs	visual sequential	visual multitask.	auditory multitask.
TOP critical	take-over time	0.87 sec	1.45 sec	1.55 sec	1.31 sec	1.48 sec
	brake application	81 %	73 %	77 %	64 %	78 %
	crash incidence	38 %	42 %	36 %	55 %	35 %
TOP non-critical	take-over time	1.21 sec	1.73 sec	1.85 sec	1.89 sec	1.42 sec
	brake application	41 %	38 %	41 %	18 %	56 %
Measured attention	scenarios	-	16.32	15.59	14.64	18.74
	ratio to benchmark	-	84 %	79 %	90 %	83 %

Table 1. Overall descriptive measures of the TOP and attention – averages by scenario

We evaluate the attention devoted to all NDRTs combined based on a calculated attention score for each NDRT. This score represents the ratio of the correctly classified stimuli (true positive or true negative) out of all classified stimuli. To factor in that the participants could influence the speed of the test, we weigh this ratio by the amount of correctly classified tests. The scores are normalized to take the length of the different scenarios into account. To balance individual predisposition, we compare the score with the individual attention scores measured in the benchmark phase. Table 1 indicates that the ratio of scenarios to benchmark normalized by time varies in dependence on the NDRT.

To analyze the effect of attention to the NDRT on the TOP in critical scenarios, we divide the scenarios into two groups. For this, we argue that a ratio of scenarios to the benchmark of 90 % depicts that the participant devoted almost as much attention to the NDRT in the specific scenario as during the benchmark phase. Hence, we assign scenarios with a ratio to the benchmark of at least 90 % into the 'high' attention category and those with less than 90 % to 'low'. This classification leads to 54 scenarios with 'high' and 34 scenarios with 'low' attention. To assess the hypotheses, we only consider the critical scenarios as the less critical scenarios did not require the participants to react as soon as possible and might falsify the evaluation of the TOP. For the statistical analysis, we employ one-sided paired t-tests. There is a maximum of 8 critical scenario data points for each participant, two for each scenario (i.e., the three NDRTs and the control condition with no NDRT). Accordingly, if we compare no NDRT with all NDRTs, there are 2 (no NDRT) times 6 (the three NDRTs) pairs. Analogously, comparing visual multitasking to auditory multitasking only leads to a maximum of 2 times 2 pairs for each participant.

For H1, we compared each NDRT individually and all NDRTs aggregated to no NDRT. Table 2 provides the p-values of the paired one-sided t-tests and the degrees of freedom (df). It depicts that performing an NDRT significantly affects the take-over time compared to not executing an NDRT ($p < .001$). This remains significant when comparing the individual NDRTs. Consequently, we provide evidence that the performance of an NDRT increases the take-over time. Yet, this does not translate to brake application and crash incidence. Hence, we can only partly support H1. It is remarkable though, that in the case of visual multitasking the results of the t-tests are significant for all TOP dimensions.

	Comparing … to no NDRT			
TOP	visual sequential	visual multitask.	auditory multitask.	all NDRTs
-- take-over time (lower)	< .001	< .001	< .001	< .001
-- brake application	.449	.002	.721	.063
-- crash incidence	.129	.002	.615	.014
df	106	104	106	318

Table 2. P-values of paired one-sided t-tests for H1.

In terms of H2, Table 3 compares multitasking-compatible NDRTs, respectively auditory and visual multitasking, with visual tasking. It provides evidence that the take-over time is significantly lower for multitasking-compatible NDRTs. As the results for brake application and crash incidence are not significant though, we find only partial support for H2. Analyzing the NDRTs individually, the results show that for auditory multitasking the take-over time and the crash incidence are significantly lower. Yet, as the brake application is not significantly higher, we can only partly support that auditory multitasking has a better TOP than visual sequential tasking. Also, in contrast to our hypothesis, visual multitasking has a significantly lower brake application than visual sequential tasking.

	Comparing … to visual sequential tasking		
TOP	auditory multitask.	visual multitask.	multitasking-compatible NDRTs
-- take-over time	.023	.208	.024
-- brake application	.178	.997	.878
-- crash incidence	.031	.925	.345
df	107	107	215

Table 3. P-values of paired one-sided t-tests for H2.

Table 4 illustrates the results for H3. It illustrates that the non-visual NDRT leads to a significantly better TOP in all aspects when compared to visual NDRTs. Hence, the data support H3. This also partly transfers to the two visual conditions, yet the sample size might be too small to show significant effects on the take-over time of visual multitasking and brake application for visual sequential tasking.

	Comparing … to auditory multitasking		
TOP	visual sequential	visual multitask.	visual NDRTs
-- take-over time	.023	.065	.007
-- brake application	.178	<.001	<.001
-- crash incidence	.031	<.001	<.001
df	107	114	222

Table 4. P-values of paired one-sided t-tests for H3.

As for the previous hypotheses, we analyze H4 using paired one-sided t-tests. Contrary to our expectations, high attention to the NDRT comes with a lower average take-over time than in the case of low attention. This effect is significant (df = 68, p = .046) depicting that higher attention to the NDRT leads to a faster take-over time. Regarding the further dimension of TOP, low attention to the NDRT entails a higher brake application (df = 68, p = .150) and a lower crash incidence (df = 68, p = .186) but these effects are not significant.

5 Discussion

Diving deeper into the results, our analysis yields several interesting findings regarding the interaction between humans and vehicles in the context of automated driving. Most importantly, our findings indicate that the allowance of NDRTs during conditionally automated driving may lead to rising traffic safety risks. The rules imposed by the UN Regulation No. 157 (2021) seem to fall short to mitigate the increased safety risks associated with unplanned take-overs from the automated vehicle to its human driver. Instead, our findings indicate that lawmakers should consider more nuanced policies concerning traffic safety, for example, by restricting the execution of NDRTs that demand high attention from the driver. Also, questions of liability in the case of critical take-over scenarios should be clarified and communicated. In the following, we discuss our results more specifically.

Performing NDRTs is a trade-off between safety and productivity. Within the results, Table 2 depicts that NDRTs significantly increase the take-over time (H1: p < .001) and crash incidence (H1: p = .014); brake application (H1: p = .064) is also significant to a 10 % confidence level. Adhering to H1, these results show that the execution of NDRTs reduces the TOP. In addition, the scenario's ratios to the benchmark depict that the participants do not devote the same level of attention to the NDRT while driving as they do during the benchmark. Consequently, performing an NDRT during driving is less efficient. Hence, the opposite effect is also conceivable: the need to split the attention between traffic and NDRT might reduce the quality of the NDRT to a level that requires the driver to redo the NDRT after the ride, making a sequential execution more efficient.

Communicating the criticality of a take-over request is vital. Table 1a shows that the TOP is lower in non-critical situations compared to critical situations. While this aggregated result is not surprising as there is no urgency, it is interesting that the participants apply the brake in 41 % of the non-critical scenarios although there is no need to. Braking is not condemnable in general but there are scenarios like driving on a crowded highway in which unnecessary braking is not recommendable. It might even cause an accident if the vehicle is closely followed by another vehicle. Hence, the driver needs to be aware of the take-over request's criticality. Our study suggests that emphasis should be placed on designing human-vehicle interfaces that reliably communicate the take-over's assessed criticality to the driver to prevent accidents and contribute to increased driving safety.

Criticality and high attention lead to displacement activities. In the comparison of critical and non-critical scenarios, we observe that on average a lower take-over time occurs in critical scenarios. This finding coincides with previous studies (e.g., Lotz et al., 2020; Marberger et al., 2018) posing that the situation's criticality reduces the take-over time. In H4, we further claim that higher attention to the performed NDRT negatively affects the TOP. To test this hypothesis, we divided all scenarios into two groups, one including all scenario observations where the participants directed high attention to the NDRT and another one comprising those observations in which the drivers dedicated low attention to the NDRT. We see that the take-over time is on average lower across all NDRTs for high attention than for low attention and that this difference is significant (H4: $p = .046$). For all these scenarios, the driver either needs to process a lot of information at once (high attention on NDRT) or has little time to validate them (high criticality). Hence, we argue that such take-over requests might overstrain drivers and put them under stress. This can lead to displacement activities that manifest in quick (low take-over time) but not necessarily good reactions (higher crash incidence). Our experiment supports this reasoning as there is a higher crash incidence with high attention to the NDRT. Although this difference is not significant (H4: $p = .186$), this might be due to the rather small sample size.

Visual multitasking is not (yet?) better than visual sequential tasking in terms of TOP. To test H2, we compare the TOP indicators for visual sequential tasking with the visual and auditory multitasking-compatible conditions. H2 investigates if performing a multitasking-compatible NDRT has a lower negative effect on the TOP than performing a sequential NDRT. Table 3 indicates that multitasking-compatible NDRTs go along with a significant lower take-over time (H2: $p = .024$) than sequential tasks, yet the quicker reaction does not lead to a significantly better reaction in terms of brake application (H2: $p = .878$) or crash incidence (H2: $p = .345$). Especially when comparing visual sequential and multitasking, it is noticeable that multitasking has a better take-over time ($p = .208$) but a higher crash incidence ($p = .925$) and significantly worse brake application ($p = .997$). This observation accords with studies that did not find a difference between visually distracting and visually not distracting tasks (Kun et al., 2017; Radlmayr et al., 2019) or between hand-held and hands-free mobile phone use (World Health Organization, 2011). Additionally, it corresponds to the previously outlined observation on displacement activities. Indeed, this might be an explanation for our observations because for the majority of the participants wearing a HoloLens was a new experience that might have strained them. Since the adoption of mixed reality devices is currently growing and they become a frequent element of human-machine interfaces, we are interested to see follow-up comparisons in the future.

Should we reconsider allowing visual NDRTs? In H3, we derive that performing a non-visual NDRT has a lower negative effect on the TOP than performing a visual NDRT. Our results clearly support this hypothesis for take-over time (H3: $p = .007$), brake application (H3: $p < .001$), and crash incidence (H3: $p < .001$). This indicates that visual NDRTs as newly permitted by the UN Regulation No. 157 (2021) reduce the TOP and, hence, impair the driver's safety in comparison to previously permitted NDRTs. As it is comprehensible that drivers want to take full advantage of the automation zone and behave more freely, permitting such NDRTs seems unreasonable when they correlate with lower TOP and, thus, reduced safety for all road users. The same counts, however, for auditory NDRTs which reportedly increase the risk of car crashes by a factor of four (World Health Organization, 2011).

5.1 Contribution & Implications

Our research contributes to research by enabling a deeper understanding of the safety risks associated with conditionally automated driving as an emerging and successful application of artificial intelligence. This knowledge advances information systems research by indicating potential ways for increasing the acceptance and adoption of automated vehicles.

As a highly societally relevant hazard responsible for thousands of deaths across the globe every year, traffic safety requires rules, forbearance, and care to work out to everybody's benefit. Although driving automation has the advantage to eliminate the notable risk of human error, the current transitional state in conditionally automated driving adds new risks depending on the interaction between vehicles and drivers. Potential users of automated driving apparently feel threatened by these risks and uncertainties, currently limiting the acceptance and adoption of conditionally automated driving (Bornholt and Heidt, 2019). A better understanding of what influences the success of vehicle-human take-overs could help build users' trust. Our research contributes to this by delivering insights into the effects of NDRTs on the TOP. Our research expands on the findings of fellow researchers presenting contradicting results related to the effect of NDRTs (Zhang et al., 2019). It abstracts from individual tasks by analyzing common factors for the effects or rather non-effects of certain NDRTs. For this, we investigate how the NDRT's potential for multitasking and addressed sensory channels – visual and auditory – affect the TOP. Hence, our study has the potential to lay an important cornerstone to better understand how NDRTs influence TOP and, thus, driving safety in conditionally automated driving.

Besides the theoretical contribution, our research has several practical implications. Aiming to establish a common theoretical ground for universal regulations, our research contributes to discussions regarding recent laws balancing the trade-off between driver safety and driver's freedom during automated driving level 3. Lawmakers can take our research as a starting point for deriving the hazard potential associated with different NDRTs and for defining what NDRTs should be permitted in level 3 automated driving. Besides its implications for lawmakers, our research underpins the need for careful design of the take-over request. Car manufacturers may learn from our study to what extent the performance of NDRTs during automated driving might impair the driver's ability to effectively take over control of the vehicle. Further, they can take this as an inspiration to redesign the vehicle's interfaces to both humans and their devices with an emphasis on driving safety in the case of take-overs. This might be achievable by combining NDRT and vehicle controls in a smart steering wheel. Lastly, car and device manufacturers should jointly work on interfaces and standards that facilitate the communication between the vehicle and the devices used for NDRTs. Thereby, early warnings of possible take-overs could be presented at the device used for the NDRT and allow the driver to catch up with the current traffic situation.

5.2 Limitations & Further Research

Although we performed the behavioral study with the necessary care, some limitations reduce the validity of our results. First, the limited scope of the experiment cannot represent the vast number of drivers and possible traffic situations but only consider a small subset. Although the obtained sample (14) is slightly larger than the samples of similar experimental studies (e.g., 9 for Borojeni et al., 2017 and Ekman et al., 2018, 10 at Kun et al., 2017) and the minimum sample size suggested by our power analysis, we primarily focused on specific critical traffic situations that might occur in automated driving. This reduction in complexity though offers better comparability between the tested scenarios. However, a field experiment cannot guarantee the repeatability of always the same scenarios. For this reason and because a field experiment would have violated existing laws in our country, we used a driving simulator and employed a laboratory experiment. To prevent learning effects across the scenarios, the participants did not receive any feedback on their TOP. Instead, the driving scenarios stopped as soon as the participant intervened. While this increases the comparability of the results in different scenarios, the participants' motivation to actively influence the scenario could decrease and, thus, worsen the measurement results concerning a quick and effective intervention. The operationalization of the visual multitasking condition through a HoloLens requires drivers to switch between a far and a near focal point. This might have contributed to the low difference in TOP between the two visual conditions. Furthermore, even though our analysis assesses all proposed hypotheses, model-based analysis such as ANCOVA might add additional value by providing a greater overview.

In the investigation of multitasking-compatible NDRTs during automated driving level 3, we still see substantial potential for further research. A greater amount of different traffic scenarios and road sceneries, for instance, might allow more detailed insights into the effect of NDRTs on the TOP. Similarly, the use of a more realistic simulator that directly simulates the driving reactions and the resulting traffic situations would lead to better transferability into practice. First, the participants would get clear feedback on their reactions, since they would see the effects of their reactions. Second, it would be easier to validate the reactions of the participants, since it would be easier to evaluate which resulting traffic situation a participant's reaction led to. Additionally, the experiment can also be extended to other groups of people, for example, novice drivers, to check whether the findings obtained in this paper also apply to these groups of people. In addition, the adoption of technical devices rises with their use. For most of the participants, it was the first time using a HoloLens. Since this may have influenced our results for visual multitasking, recruiting more participants who are familiar with the HoloLens may result in a better TOP for performing an NDRT with a HoloLens. Furthermore, in our experiment, we noticed that in absence of any take-over request some participants took over the driving process. Consequently, it is arguable that the reason for this interference is trust-related and that the participants did not trust autonomous driving enough. This underpins ongoing trust-related research in autonomous driving (e.g., Ekman et al., 2018) and provides evidence for individual design for every maturity level (0 to 5) to fully automated driving.

6 Conclusion

Once the dawn of conditionally automated driving has come, drivers may use otherwise passive driving time by performing more and more NDRTs while driving. Despite the tremendous potential for additional productivity arising from the use of mixed reality devices (here, the Microsoft HoloLens) to facilitate multitasking, drivers need to be prepared to take control of the vehicle at all times. Our study investigated the effects of different NDRTs (auditory task, visual task on regular display, and visual task in HoloLens) on the TOP in critical traffic situations as well as the hazard potential for driving safety associated with the drivers' behavioral adaptation. We employed a behavioral experiment in which participants operate a simulated automated vehicle and evaluate their reaction when the vehicle requests them to take over control at some point in time. During the automated ride, the different NDRTs distract the driver. We find that NDRTs with a visual component impair the TOP more than auditory NDRTs, yet UN Regulation No. 157 (2021) suggests that visual NDRTs shall be allowed for conditionally automated driving. Visual NDRTs might enable greater productivity during driving but also come with greater risk. This trade-off must be taken into account with great care to prevent severe consequences.

7 References

Atwood, J. R., F. Guo and M. Blanco (2019). "Evaluate driver response to active warning system in level-2 automated vehicles" *Accident Analysis & Prevention* 128, 132–138.

Biondi, F., I. Alvarez and K. Jeong (2019). "Human–Vehicle Cooperation in Automated Driving: A Multidisciplinary Review and Appraisal" *International Journal of Human–Computer Interaction* 35 (11), 932–946.

Blanco, M., J. Atwood, H. M. Vasquez, T. E. Trimble, V. L. Fitchett, J. Radlbeck, G. M. Fitch, S. M. Russell, C. A. Green, B. Cullinane and others (2015). *Human factors evaluation of level 2 and level 3 automated driving concepts.* Report No. DOT HS 812 182. National Highway Traffic Safety Administration.

Bornholt, J. and M. Heidt (2019). "To Drive or not to Drive - A Critical Review regarding the Acceptance of Autonomous Vehicles" *ICIS 2019 Proceedings*.

Borojeni, S. S., T. Wallbaum, W. Heuten and S. Boll (2017). "Comparing Shape-Changing and Vibro-Tactile Steering Wheels for Take-Over Requests in Highly Automated Driving". In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*: ACM, pp. 221–225.

Brechtel, S., T. Gindele and R. Dillmann (2014). "Probabilistic decision-making under uncertainty for autonomous driving using continuous POMDPs". In: *Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems*: IEEE, pp. 392–399.

Buser, T. and N. Peter (2012). "Multitasking" *Experimental Economics* 15 (4), 641–655.

Casner, S. M., E. L. Hutchins and D. Norman (2016). "The Challenges of Partially Automated Driving" *Communications of the ACM* 59 (5), 70–77.

Charlton, S. G. (2009). "Driving while conversing: Cell phones that distract and passengers who react" *Accident Analysis & Prevention* 41 (1), 160–173.

Dogan, E., V. Honnêt, S. Masfrand and A. Guillaume (2019). "Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving" *Transportation Research Part F: Traffic Psychology and Behaviour* 62, 494–504.

Dosovitskiy, A., G. Ros, F. Codevilla, A. Lopez and V. Koltun (2017). "CARLA: An open urban driving simulator" *arXiv preprint arXiv:1711.03938*.

Duncan, J. (1984). "Selective attention and the organization of visual information" *Journal of experimental psychology: General* 113 (4), 501–517.

Ekman, F., M. Johansson and J. Sochor (2018). "Creating Appropriate Trust in Automated Vehicle Systems: A Framework for HMI Design" *IEEE Transactions on Human-Machine Systems* 48 (1), 95–101.

Endsley, M. R. (1995). "Toward a Theory of Situation Awareness in Dynamic Systems" *Human factors* 37 (1), 32–64.

Ernst, C.-P. and P. Reinelt (2017). "Autonomous Car Acceptance: Safety vs. Personal Driving Enjoyment" *AMCIS 2017 Proceedings*.

Gold, C., R. Happee and K. Bengler (2018a). "Modeling take-over performance in level 3 conditionally automated vehicles" *Accident Analysis & Prevention* 116, 3–13.

Gold, C., F. Naujoks, J. Radlmayr, H. Bellem and O. Jarosch (2018b). "Testing Scenarios for Human Factors Research in Level 3 Automated Vehicles". In N. A. Stanton (ed.) *Advances in Human Aspects of Transportation*, pp. 551–559. Cham, Switzerland: Springer International Publishing.

Goldstein, E. B. (2015). *Cognitive psychology. Connecting mind, research and everyday experience.* 4th edition. Stamford, CT: Cengage Learning.

Hein, D., P. Rauschnabel, J. He, L. Richter and B. Ivens (2018). "What Drives the Adoption of Autonomous Cars?" *ICIS 2018 Proceedings*.

Hergeth, S., L. Lorenz, R. Vilimek and J. F. Krems (2016). "Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving" *Human factors* 58 (3), 509–519.

Horberry, T., J. Anderson, M. A. Regan, T. J. Triggs and J. Brown (2006). "Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance" *Accident Analysis & Prevention* 38 (1), 185–191.

Hutchins, N. and L. Hook (2017). "Technology acceptance model for safety critical autonomous transportation systems". In: *36th DASC. Digital Avionics Systems Conference*. Piscataway, NJ: IEEE, pp. 1–5.

Keller, K., C. Zimmermann, J. Zibuschka and O. Hinz (2021). *Hawaii International Conference on System Sciences 2021*. Honolulu, HI: University of Hawai'i at Manoa Hamilton Library.

Kolodko, J. and L. Vlacic (2003). "Cooperative autonomous driving at the intelligent control systems laboratory" *IEEE Intelligent Systems* 18 (4), 8–11.

Kuehn, M., T. Vogelpohl and M. Vollrath (2017). "Takeover Times in Highly Automated Driving (Level 3)". In: *Proceedings of the 25th International Technical Conference on the Enhanced Safety of Vehicles*.

Kujala, T., K. Kircher and C. Ahlström (2021). "A Review of Occlusion as a Tool to Assess Attentional Demand in Driving" *Human factors*, 187208211010953.

Kun, A. L., H. van der Meulen and C. P. Janssen (2017). "Calling While Driving: An Initial Experiment with HoloLens". In: *Proceedings of the 9th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design: driving assessment 2017*: University of Iowa, pp. 200–206.

Lackes, R., M. Siepermann and G. Vetter (2020). "Where can I take You? – The Drivers of Autonomous Driving Adoption" *ECIS 2020 Research Papers*.

Lee, J. and J. H. Yang (2020). "Analysis of Driver's EEG Given Take-Over Alarm in SAE Level 3 Automated Driving in a Simulated Environment" *International Journal of Automotive Technology* 21 (3), 719–728.

Lee, S. C., S. H. Yoon and Y. G. Ji (2021). "Effects of Non-Driving-Related Task Attributes on Takeover Quality in Automated Vehicles" *International Journal of Human–Computer Interaction* 37 (3), 211–219.

Lotz, A., N. Russwinkel and E. Wohlfarth (2020). "Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks" *Cognition, Technology & Work* 22 (4), 733–744.

Marberger, C., H. Mielenz, F. Naujoks, J. Radlmayr, K. Bengler and B. Wandtner (2018). "Understanding and Applying the Concept of "Driver Availability" in Automated Driving". In N. A. Stanton (ed.) *Advances in Human Aspects of Transportation*, pp. 595–605. Cham, Switzerland: Springer International Publishing.

Marinik, A., R. Bishop, V. Fitchett, J. F. Morgan, T. E. Trimble and M. Blanco (2014). *Human factors evaluation of level 2 and level 3 automated driving concepts: Concepts of operation.* Report No. DOT HS 812 044. National Highway Traffic Safety Administration.

Miller, J. and M. Durst (2014). ""Just do it when you get a chance": the effects of a background task on primary task performance" *Attention, perception & psychophysics* 76 (8), 2560–2574.

NHTSA (2013). "National Highway Traffic Safety Administration. Preliminary Statement of Policy Concerning Automated Vehicles".

Petersen, L., L. Robert, J. Yang and D. Tilbury (2019). "Situational Awareness, Driver's Trust in Automated Driving Systems and Secondary Task Performance" *SSRN Electronic Journal* Forthcoming.

Pettitt, M., G. Burnett and A. Stevens (2009). "Defining Driver Distraction". In: *World Congress on Intelligent Transport Systems*. URL: https://www.researchgate.net/publication/242449092_Defining_Driver_Distraction.

Philip, P., P. Sagaspe, N. Moore, J. Taillard, A. Charles, C. Guilleminault and B. Bioulac (2005). "Fatigue, sleep restriction and driving performance" *Accident; analysis and prevention* 37 (3), 473–478.

Radlmayr, J., F. M. Fischer and K. Bengler (2019). "The Influence of Non-driving Related Tasks on Driver Availability in the Context of Conditionally Automated Driving". In: *Proceedings of the 20th Congress of the International Ergonomics Association*: Springer International Publishing, pp. 295–304.

Roenker, D. L., G. M. Cissell, K. K. Ball, V. G. Wadley and J. D. Edwards (2003). "Speed-of-processing and driving simulator training result in improved driving performance" *Human factors* 45 (2), 218–233.

SAE (2016). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. J3016.* Warrendale, United States of America: SAE International.

Schmidt-Atzert, L. (2004). “Kapitel 4 Test d2: Aufmerksamkeits-Belastungs-Test”. In G. Büttner and L. Schmidt-Atzert (eds.) *Diagnostik von Konzentration und Aufmerksamkeit*, pp. 87–102. Göttingen, Germany: Hogrefe Verlag.

Seppelt, B. D. and T. W. Victor (2016). “Potential Solutions to Human Factors Challenges in Road Vehicle Automation”. In G. Meyer and S. Beiker (eds.) *Road Vehicle Automation 3*, pp. 131–148. Cham, Switzerland: Springer International Publishing.

Stephanidis, C., G. Salvendy, M. Antona, J. Y. C. Chen, J. Dong, V. G. Duffy, X. Fang, C. Fidopiastis, G. Fragomeni, L. P. Fu, Y. Guo, D. Harris, A. Ioannou, K. Jeong, S. Konomi, H. Krömker, M. Kurosu, J. R. Lewis, A. Marcus, G. Meiselwitz, A. Moallem, H. Mori, F. Fui-Hoon Nah, S. Ntoa, P.-L. P. Rau, D. Schmorrow, K. Siau, N. Streitz, W. Wang, S. Yamamoto, P. Zaphiris and J. Zhou (2019). “Seven HCI Grand Challenges” *International Journal of Human–Computer Interaction* 35 (14), 1229–1269.

Swadesh, M. (1936). “Phonemic Contrasts” *American Speech* 11 (4), 298–301.

Tellis, L., G. Engelman, A. Christensen, A. Cunningham, R. Debouk, K. Egawa, C. Green, C. Kawashima, G. Nicols, D. Prokhorov, B. Wendling, S. Yako and S. Kiger (2016). *Automated Vehicle Research for Enhanced Safety*. National Highway Traffic Safety Administration. URL: http://www.campllc.org/avr/.

Trimble, T. E., R. Bishop, J. F. Morgan, M. Blanco and others (2014). *Human factors evaluation of level 2 and level 3 automated driving concepts: Past research, state of automation technology, and emerging system concepts.* Report No. DOT HS 812 043. National Highway Traffic Safety Administration.

Tsang, P. S. and M. A. Vidulich (2006). “Mental workload and situation awareness”. In *Handbook of human factors and ergonomics, 3rd ed*, pp. 243–268. Hoboken, NJ, US: John Wiley & Sons, Inc.

Underwood, G., A. Ngai and J. Underwood (2013). “Driving experience and situation awareness in hazard detection” *Safety Science* 56, 29–35.

United Nations (2021). *Addendum 156 – UN Regulation No. 157*. URL: https://unece.org/transport/documents/2021/03/standards/un-regulation-no-157-automated-lane-keeping-systems-alks (visited on 11/15/2021).

Villalobos-Zúñiga, G., T. Kujala and A. Oulasvirta (2016). “T9+HUD: Physical Keypad and HUD Can Improve Driving Performance While Typing and Driving”. In: *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Ed. by P. Green, S. Boll, J. Gabbard, S. Osswald, G. Burnett, S. S. Borojeni, A. Löcken, A. Pradhan. New York, NY, USA: ACM, pp. 177–184.

Vogelpohl, T., M. Vollrath, M. Kühn, T. Hummel and T. Gehlert (2016). *Übergabe von hochautomatisiertem Fahren zu manueller Steuerung. Teil 1. Review der Literatur und Studie zu Übernahmezeiten.* Berlin: Unfallforschung der Versicherer, GDV.

Wickens, C. D. (2008). “Multiple resources and mental workload” *Human factors* 50 (3), 449–455.

Wolfe, B., L. Fridman, A. Kosovicheva, B. Seppelt, B. Mehler, B. Reimer and R. Rosenholtz (2019). “Predicting road scenes from brief views of driving video” *Journal of vision* 19 (5), 8.

Wood, J. M. (2002). “Age and visual impairment decrease driving performance as measured on a closed-road circuit” *Human factors* 44 (3), 482–494.

World Health Organization (2011). *Mobile phone use. A growing problem of driver distraction.* Geneva, Switzerland. URL: https://www.who.int/publications/i/item/mobile-phone-use-a-growing-problem-of-driver-distraction (visited on 11/15/2021).

Wu, Y., K. Kihara, K. Hasegawa, Y. Takeda, T. Sato, M. Akamatsu and S. Kitazaki (2020). "Age-related differences in effects of non-driving related tasks on takeover performance in automated driving" *Journal of safety research* 72, 231–238.

Zeeb, K., A. Buchner and M. Schrauf (2015). "What determines the take-over time? An integrated model approach of driver take-over after automated driving" *Accident; analysis and prevention* 78, 212–221.

Zhang, B., J. de Winter, S. Varotto, R. Happee and M. Martens (2019). "Determinants of take-over time from automated driving: A meta-analysis of 129 studies" *Transportation Research Part F: Traffic Psychology and Behaviour* 64, 285–307.