Preparing Drivers for Planned Control Transitions in Automated Cars

Kai Holländer
LMU Munich
Munich, Germany
kai.hollaender@ifi.lmu.de

Bastian Pfleging
LMU Munich
Munich, Germany
bastian.pfleging@ifi.lmu.de

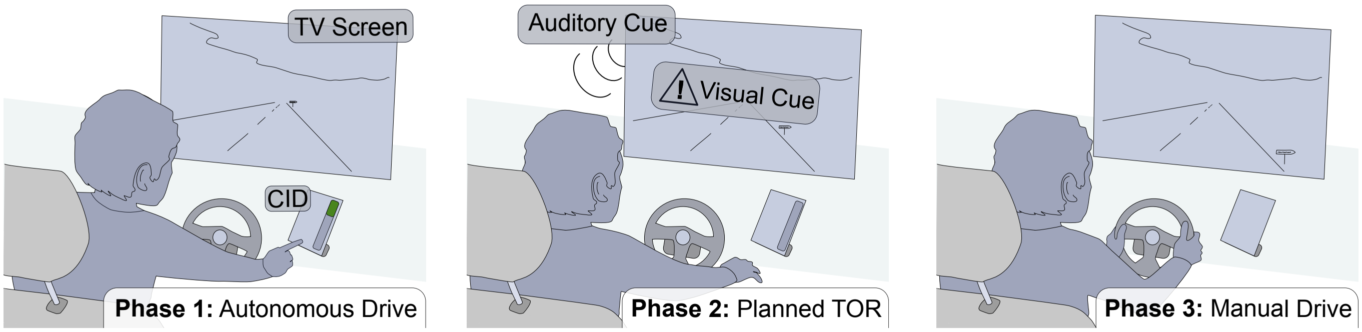

Figure 1: We investigate the influence of pre-warnings, devices, and non-driving-related activities (NDRAs) on driving performance and usability when drivers take over from the automated car (planned control transitions). Participants drive autonomously and perform an NDRA with a central information display (CID) or smartphone. While driving, either a countdown or a dynamic bar visualization indicates the upcoming take-over situation (*Phase 1*). The planned take-over request is then issued with an auditory cue and a visual warning in the head-up display (*Phase 2*). After regaining control, participants finish the drive in the static driving simulator manually (*Phase 3*).

ABSTRACT

In the near future we expect automated driving to be available for specific segments of a journey, e.g., when driving on the highway. At the end of such a route segment, a (planned) control transition from system to driver occurs. While immediate (unpredictable) take-over situations are heavily investigated, there is still a gap in understanding how to present planned take-over requests, especially while drivers might be involved in non-driving-related activities (NDRAs). We investigated the effect of three different visual representations to indicate planned take-over requests (TOR) on *usability*, *comfort*, and *driving quality*. Additionally, we explored the influences of different NDRAs and the device used for this activity. The results of our simulator study (N=24) indicate that (1) upcoming take-over requests should be displayed dynamically, (2) preferred devices depend on the performed task and (3) take-over requests should be presented with auditory, visual, and tactile cues. Based on our findings, we contribute design recommendations to support the development of safe and comfortable planned control transitions.

MUM '18, November 25–28, 2018, Cairo, Egypt

ISBN 978-1-4503-6594-9/18/11…$15.00

DOI: https://doi.org/10.1145/3282894.3282928

CCS Concepts

•**Human-centered computing** → *Interaction devices; Ubiquitous and mobile devices;*

Author Keywords

Automated Driving; Planned Control Transitions; Non-Driving-Related Activities; Take-Over Requests

INTRODUCTION

Automated driving features are expected to be available very soon for public use. This ground-braking change in mobility will alter the role of the driver: So far, drivers are responsible for operating and monitoring the vehicle and its surroundings. The more this task gets automated, the more drivers move into a passive role regarding vehicle control [6]. Automated cars for SAE (Society of Automotive Engineers) level 3 *(Conditional Automation)* and level 4 *(High Automation)* do not require operators to steer, brake, accelerate, or even monitor the vehicle permanently [28]. One advantage of such systems is that parts of a journey can be used for non-driving-related activities (NDRA). For example, writing text messages, watching videos, or browsing the Internet [23].

Take-over requests (TOR) may occur once the automated system reaches its limits [18]. Unexpected events, e.g., missing road markings, accidents, sensor or software errors, construction zones, and extreme weather conditions could lead to TORs. In this situation the driver is expected to immediately take back control of the vehicle (unplanned TOR). There has been a lot of research about certain aspects of such unexpected take-over situations like required take-over times [8, 12], the influence of traffic density [26], modalities used for warnings [25], or the general design of take-over processes [20]. Hence, in many studies critical circumstances were observed which are impossible to foresee in the wild.

The aforementioned unpredictable cases represent important but hopefully exceptional situations. In contrast, we explore how safety and comfort can be obtained in situations which are spot on predictable, i.e., expectation-conform scenarios. Since car manufacturers are developing autonomous driving systems and soon the first consumer car[1] features SAE level 3 automation, planned TORs will become a part of everyday drives.

Understanding and designing take-over situations is an essential aspect to ensure safe and comfortable (partly) automated drives. Considering take-over requests, this includes questions on how to design (multimodal) cues for the take-over request, and which influence the driver's engagement in non-driving-related tasks has.

It is likely that drivers involve in NDRAs during an autonomous drive, potentially using either in-vehicle technology or a mobile device [23]. Thus, we focus especially on cases including interaction with such technology. Young drivers tend to use their cell phones when driving already, although this behavior implies a great risk [5, 27]. Consequently, specific NDRAs could also be performed with a smartphone, e.g., writing text messages, watching video clips, or phone calls. On the other hand, modern cars are often equipped with a central information display (CID) as part of their in-vehicle infotainment system. We are interested if the choice of device includes a trade-off regarding comfort and safety during control transitions.

We conducted a simulator study with 24 participants. Our dependent variables include visual presentations of planned TOR, interaction devices, and NDRAs. We collected qualitative and quantitative data regarding usability, mental workload, driving quality and users' preferences. Based on our results, we formulated design recommendations for planned take-over requests. These are meant to support researchers and developers when designing such control transitions.

Our results show that an abstract bar-indication of planned TOR shows higher usability than a time countdown or no visualization. Furthermore, 75% of the participants preferred a dynamic bar-visualization. According to their feedback, using a dynamic bar provides a possibility for mental preparation via peripheral vision. Regarding the device choice there are hardly any differences in perceived mental workload, usability, or driving quality. However, the performed non-driving-related activity seems crucial for the preferred interaction device.

We contribute to enhance safety and comfort in semi-automated driving by:

- Investigating user behavior during control transition.
- Exploring effects of used devices and presented warnings on driving quality, usability, and mental workload.
- Focusing on planned transitions instead of exceptional TORs.
- Providing recommendations for the design of take-over requests.

RELATED WORK

The start of a take-over procedure is usually defined by an issued take-over request [18]. According to Petermann-Stock et al. the subsequent user reactions can be categorized into four classes [22]: First, there is an orientation phase including viewing up from the NDRA and focusing on the street. This is followed by the willingness to act, which begins with a manual intervention in the vehicle control. Afterwards, the transition is performed and in the last phase, the vehicle gets stabilized. In this paper, we address the first phase. We investigate if orientating and taking vehicle control could be eased, if drivers get a chance for mental preparation while performing NDRAs.

Many projects already looked into different aspects of take-over requests. Focus is often put on determining the take-over time, i.e., the time drivers need to take back vehicle control [12, 18, 34]. Gold et al. conducted a study investigating gaze behavior, driving quality and reaction times at a dynamic driving simulator [12]. Similar to Kuehn et al., performance in manual drives were compared to semi-autonomous drives [12, 18]. According to their findings, drivers pass obstacles more narrowly and tend to look less on the speedometer, road and surroundings if the granted take-over time is shorter. However, analog to Walch et al. they only observed a single NDRA [34]. Since NDRA can be manifold [7, 23, 29] we chose a broader approach by investigating three NDRAs and having a more diverse distribution of participants age and academic background. Furthermore, resources needed for processing NDRAs might interfere with the driving task [15]. Thus, investigating multiple NDRAs might uncover how specific tasks influence driving safety and comfort.

Vogelpohl et al. conducted a study with two NDRAs, reading a newspaper an playing tetris [18]. These tasks got processed by participants on a tablet computer. They explored multiple triggers for TOR, including missing road markings, sensor or software errors, construction zones, extreme weather conditions and a (planned) highway exit. They state that drivers in an autonomous vehicle have a longer response time in all the situations mentioned above (compared to manual driving). Vogelpohl et al. recommend to inform drivers as early and obvious as possible about upcoming TOR. Langdon et al. suggest to use a time countdown as early indication for upcoming TORs [19]. Volvo introduced a concept displaying remaining

[1] https://www.audi-mediacenter.com/en/on-autopilot-into-the-future-the-audi-vision-of-autonomous-driving-9305, last access: August 2018

autonomous travel time abstractly and numerically[2]. We investigate differences in safety and comfort between a countdown and an abstract visualization.

Diederich und Colonius observed the influence of multiple stimuli on reaction times in general [10]. The authors argue that multimodal cues lead to faster respond times. Several driving simulator studies suggest that their findings are relevant in context of automated driving as well [1, 21, 24, 30]. Acoustic hints prior to handovers can enhance safety [32] and language based warnings might perform worse than abstract ones [25]. Hence, we use a take-over request with abstract sound cues and a visual warning in the HUD, see Figure 1.

USER STUDY

To investigate our research questions we chose an autonomous drive which gets interrupted by a take-over request. This drive took place on a highway, as highway assistants appear to be the first ready-to-market technology allowing highly automated driving.

We answer the following research questions in context of planned transitions from automated to manual driving:

RQ1 Should upcoming take-over requests be indicated with a time-countdown or an abstract dynamic bar? Which visualization provides lower mental workload, higher driving quality and better usability?

RQ2 How do the performed non-driving-related activity and the associated device influence driving quality, usability, and mental workload during planned control transitions?

We envision a notification system which already informs the driver in advance. By providing details such as the remaining time or distance until automation ends, we expect the driver to be able to prepare on time for the take-over situation. Hence, we implemented a prototype providing drivers with a pre-warning regarding upcoming TOR while performing NDRAs. Figure 2 illustrates behavior of the pre-warnings and Figure 3 shows screenshots of the application. The pre-warning designs are deduced from an expert interview at the BMW research and innovation center as well as literature reviews [18, 19, 25]. We implemented three different NDRAs inspired by the work of Pfleging et al. [23].

Prior research suggests that take-over requests should employ multiple modalities [1, 10, 16, 21, 24]. Thus, we presented TORs visually at the head-up display simultaneously with an additional sound cue. We used a fixed central console and respectively a smartphone to display pre-warnings for upcoming TORs and NDRAs. The devices show the same content adapted to the screen format and resolution by accessing the same responsive website. Although we could have made it possible for individuals to use their own mobile phone, we decided to hand out the same device for everyone to guarantee consistency. Figure 1 illustrates an exemplary study sequence with a CID and a dynamic bar pre-warning. Figure 4 shows pictures of participants during the study.

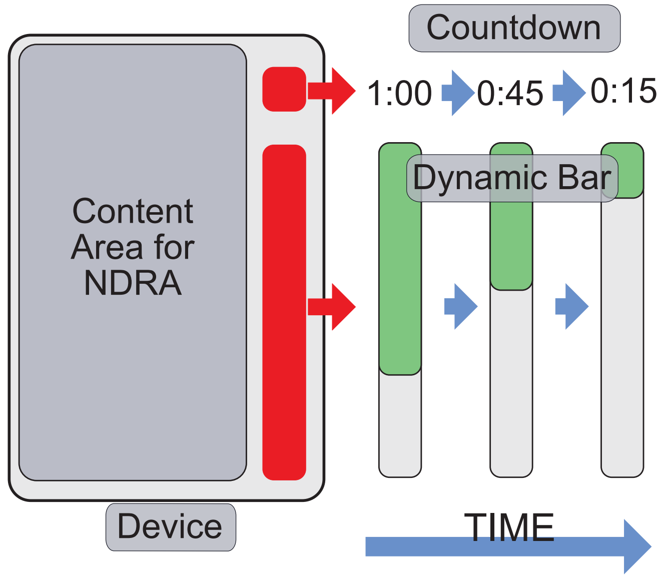

Figure 2: Either a countdown (top) or a dynamic bar (bottom) filled the corresponding red colored area. Pre-warnings indicate remaining autonomous driving time (countdown) or distance (bar) until a take-over request is issued.

The simulated scenario includes a highway without traffic, because other road users can impair driving quality [13]. A TOR appeared at a speed of 80 km/h. Participants were supposed to take-over manual control, drive on a highway exit and stop at a parking lot.

To investigate our research questions, we conducted a 3 *(pre-warnings)* x 3 *(NDRA)* x 2 *(device)* within-subject experiment. "Pre-warning" is used here as a synonym for the visual indication of an upcoming TOR. The term is supposed to express that this visualization is present prior to any warning or TOR. In future, such pre-warnings could be displayed as soon as the system accesses information about the route including locations or paths which cannot be driven autonomously. We explored a total of three NDRAs which are derived from a literature review. The used device was either a CID or a smartphone.

24 participants were introduced to each pre-warning and NDRA condition once. Hence, every subject completed three drives which resulted in a sum of 72 samples. Indicators, devices and NDRAs were evenly mixed (counterbalanced) amongst participants. Each input device was used 36 times. Every possible condition-set made of a pre-warning, NDRA and a device was tested 4 times. Each driver either used a smartphone once and the CID twice or vice-versa. The amount of participants using smartphone or CID twice was equally distributed. In order to avoid sequence effects, the distribution of conditions alternated for each participant. The collected quantitative data set was parsed with multivariate analysis of variances to find correlations between dependent and independent variables. Insights received from qualitative data got

[2] https://www.volvocars.com/en-kw/about/our-innovation-brands/intellisafe/intellisafe-autopilot/c26, last access August 2018

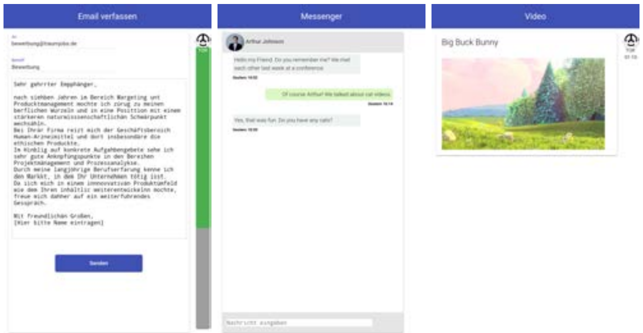

Figure 3: Screenshots of NDRAs and indicators. From left to right: e-mail correction with bar indicator; text messenger without pre-warning; video task with time countdown.

transferred together with findings from statistical analysis into design recommendations.

Independent Variables

The following list contains three independent variables explored in this user study:

Pre-warnings For this variable, we had three levels: In *Bar* condition, the interfaces shows a dynamic bar that decreases from bottom to top and reflects the vehicle moving direction. The bar continuously indicates remaining distance until the upcoming TOR. In *Countdown* condition the interface displays a digital countdown indicating time left until take-over as numerical value in seconds. In addition, we implemented a *Base* condition with no graphical representation at all. Figure 2 shows bar and countdown behavior.

NDRAs We implemented three different non-driving-related tasks. During a (long) drive it is likely that more than one NDRA gets executed, hence we reflected on that. We identified NDRAs which are expected to be performed by future users of semi-automated cars via literature review. Each implemented task fits in a different context. We implemented a *text messenger* application for chatting (social and communication), an *e-mail* correction task (typical daily work routine) and watching a *video* (relaxation & multimedia), as shown in Figure 3.

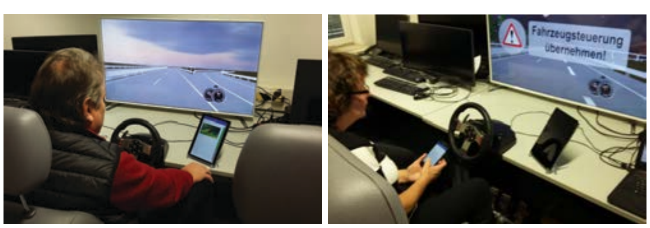

Figure 4: Participant performing video task with dynamic bar pre-warning for upcoming TOR on CID while driving autonomously (left). Participant looking up from smartphone because of a take-over request (right).

Devices Pre-warnings and non-driving-related activities were either shown on a smartphone or on a central information display as both devices are often present in modern vehicles. On th one hand, there are many studies suggesting that people want to use their phone while driving autonomously. On the other hand, car manufacturers provide sophisticated in-vehicle infotainment systems. We investigated how differences regarding the used device are perceived by participants, how they influence take-over time and driving quality.

Dependent Variables

We evaluated mental workload with the Driver Activity Load Index (DALI) questionnaire [31]. Additionally, participants filled out the System Usability Scale (SUS), which is an established measurement tool for usability [2, 3]. Both questionnaires produce metric outcome scores. DALI-values are computed with the same method as NASA TLX-scores [14].

To achieve an understanding of users' experiences during the experiment we collected qualitative insights with the help of a semi-structured interview. Participants were asked about their preferred device and pre-warning. We encouraged them to think aloud about alternative pre-warnings and TOR-procedures. Additionally, we asked how take-overs of manual control were processed and why. There was a total of six questions which got answered in an open conversation. These interviews lasted from about 10 min to 30 min conversations.

Furthermore, we investigated the time between each TOR and the driver's first intervention with vehicle controls, e.g. steering, braking or accelerating (in seconds). The standard deviation from a fixed ideal driving-trajectory (in meters) and driving speed at the exit (in kilometers per hour).

Appropriate responses to a TOR are strongly situation-dependent. For the evaluation we considered driving quality to be ideal if the following four conditions are met:

Standard deviation of the ideal line is less than 0.8 m. If this value is exceeded, it is very likely that road markings have been crossed. Hence, a value over 0.8 m is an indicator for poor lane keeping.

The response time after a TOR until the first control intervention is less than 10 seconds. If this value is exceeded it is impossible to fulfill the given driving task.

The first reaction is either steering or braking, since accelerating is not constructive in the scenario presented here.

Reach of a moderate speed at exit (less than 80 km/h). Driving faster than 80 km/h would not comply with local traffic regulations.

Participants

A total of 24 individuals aged between 21 and 64 years took part in our laboratory study *(M = 29.58 years, SD = 11.87 years; 11 women and 13 men)* (No data set was excluded). Students provided a share of 50%, the others employment were encountered each only once and spread over diverse professions including a pensioner. Participants covered various levels of education.

We requested a valid driver's license to guarantee knowledge of local traffic laws. Participants reported to drive once a month or less (37.5%), once a week (29.17%), several times a week (20.83%) or daily (12.5%).
We recruited participants via internal university e-mail lists. Each participant received an expense allowance of five euro.

Apparatus

We implemented a static driving simulator as demonstrated in Figure 1 and Figure 4. To imitate the CID we built a laser-cut mount and inserted a Huawei Media Pad M3 Lite tablet, display size: 10.1 " (25,6 cm). This was placed in portrait mode to the right of the steering wheel, similar to central consoles of current models from different car manufacturers such as Volvo XC60[3], Renault Mégane[4], or Tesla Model S[5]. Participants used a Motorola MotoG (2nd generation) smartphone with a 5 " (12,7cm) display (720 x 1.280 Pixel).

The simulation ran on a HP Envy computer with an Intel Core i7-8700 processor, 16 GB RAM and an NVIDIA GeForce GTX 1060 GPU with 50 fps. A 48 " (121cm) LED TV, as well as a Logitech G27 steering wheel with corresponding pedals complete the simulator setup.

Three driving scenarios were implemented with the OpenDS Version 4.5[6] simulation environment. Three NDRAs were presented in the form of a responsive website, which was invoked on the tablet or the mobile phone. Both devices showed the same content relative to their display size. To set up communication from the PC which ran OpenDS with the devices, we implemented a website and two servers with HTML, Javascript (Node.js), XAMPP and Websockets. In order to exchange data wireless with external devices (CID and smartphone) we employed an LinksysWRT 54G router.

Implemented Take-Over Scenario

The arrangement of each TOR is static and independent of the study conditions. We presented handover requests with the sequence demonstrated in Figure 1. TORs occurred with an acoustic warning signal and a visual request in the head-up display. Hence, we followed suggestions stated by Walch et al. by implementing visual and auditory cues [34]. The auditory notification consisted of 16 single "beep" sounds which appeared within a time span of ten seconds. The acoustic warnings became slightly louder for each cue. A transparent message in the HUD occurred simultaneously, this moment is captured in the right photo of Figure 4. Drivers could take back control of the vehicle by either steering, braking or accelerating *(immediate handover)*. If test persons intervened in vehicle control both warnings (HUD and auditory cues) disappeared at once. As a result of the control intervention, drivers transferred into manual mode. If there was no takeover of control within ten seconds the autonomously moving vehicle passed the exit and continued to drive on the highway. In that case the take-over request would have disappeared, too. Hence, there was no risky situation implied at any time. However, all participants took back control within 10 seconds.

Procedure

We welcomed participants in the laboratory and briefed them about the experiments' course. Subsequently, they signed a consent form and filled out an initial questionnaire on job, age and driving frequency. Afterwards, subjects completed a manual test drive to become familiar with the simulator. The first autonomous trip followed. Each drive took place on a highway. At the beginning the experimenter prepared the environment (selection of NDRA, pre-warning type, and device) according to the current condition. In addition, he explained the NDRA which participants were supposed to perform during the current condition. Shortly after the start of each drive, participants began performing the assigned NDRA. The car was driving highly automated for about 90 seconds, which allowed subjects to focus on the selected task. Tasks were designed to exceed overall duration of each condition. A TOR interrupted the process, prompting participants to switch to manual driving. After taking over control of the vehicle, they drove a distance of about 450 meters manually.

Each participant completed three runs, which took approximately 45 minutes in total. After each run attendees completed the DALI and SUS questionnaires. After the last drive, participants were invited to a semi-structured interview. Our study design guaranteed that each participant experienced all three NDRAs, pre-warning types, and both devices.

In order to prevent participants from forming a habit of executing one drive multiple times, we introduced three different highway scenarios. Each run took place on another track. Track one and three only differ visually. Objects at the roadside (parking cars, buildings, plants and signs) got replaced, as well as the daytime (morning and evening) and the background (Mountains and a flat, rural area). Track two has a different start an end point than tracks one and three. Additionally, the environment got adjusted. Daytime was mid-day and the background showed a cloudy sky. Our data analysis shows that the track did not influence the dependent variables.

Limitations

Although we used a within-subjects design the sample size of 24 participants (72 samples) is still rather small. Especially, since we employed a *3 x 3 x 2* design. Therefore, we particularly considered qualitative insights.

In order to investigate planned TOR while spending a reasonable amount of time as well as money and to prevent harming our participants we conducted a simulator study. Hence, a general validity in context of highly automated driving is not necessarily to accept. This study rather represents an initial investigation regarding specific parts of an in-car interface.

The three investigated non-driving-related tasks call upon different driver resources. Drivers are required to use visual and manual channels for the text messenger and e-mail, but visual

[3] https://www.volvocars.com/us/cars/new-models/xc60, last access August 2018
[4] https://www.renault.co.uk/vehicles/new-vehicles/megane.html, last access August 2018
[5] https://www.tesla.com/models, last access August 2018
[6] https://www.opends.eu/, last access May 2018

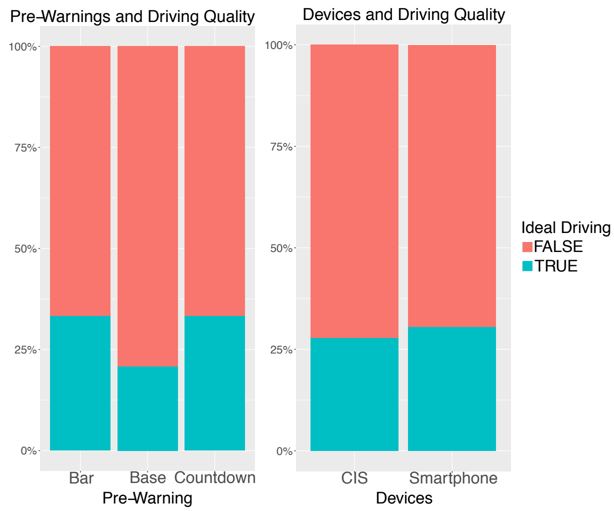

Figure 5: Presenting a pre-warning shows higher frequency of ideal driving compared to no visual indication (Base) of planned TOR (left). Measured driving quality balances for both investigated devices (right).

and auditory resources for the video task. Based on the Multiple Resource Theory [36], this could affect the take-over response. In essence, there is a greater resource conflict with the take-over response expected for the visual-manual NDRAs (e-mail and messenger) than the visual-auditory NDRA (video). However, we only found slight hints for such a conflict in our data. There was no statistically significant evidence that tasks and mental workload would interfere differently within the investigated scenario.

RESULTS

Since subjective preferences, attitudes and feelings are difficult to quantify we especially profited from results of a semi-structured interview. Not limiting participants to fixed answers resulted in manifold suggestions to improve our prototype. The following section shows quantitative results first, followed by qualitative outcomes. Collected data directly refers to our research questions presented in Section 3.

Quantitative Results

Kolmogorov-Smirnov fitting tests validated normal distribution of the data [9]. Mauchly tests proofed sphericity to avoid type 1 (alpha) errors.

Driving Quality

Chi-Square homogeneity tests resulted in *X-squared = 0.33, p-value = 0.75* for the distribution of ideal driving quality on the used devices and *X-squared = 1.5, p-value = 0.61* for the distribution of ideal driving quality on NDRAs. Fisher's exact test did not show significance on ideal driving behavior in relation to pre-warnings or NDRAs *(p-value = 1)*.

Not indicating upcoming TOR leads to a decreased driving performance. Additionally, our results suggest that either device influences driving quality equally, see Figure 5.

	Reaction Time		
	Estimate	*Conf. Int.*	*p-value*
(Intercept)	2.78	2.43 – 3.14	<.001
NDRA - Messenger	-0.09	-0.59 – 0.41	.737
NDRA - Video	0.46	-0.03 – 0.96	.072
Device - CID	0.15	-0.21 – 0.51	.415
Pre-Warning - Baseline	-0.24	-0.75 – 0.26	.345
Pre-Warning - Time	-0.02	-0.52 – 0.48	.945
Observations		72	

Figure 6: Linear fixed effects model with estimated influence of dependent variables on response time.

Predictors	*Dependent Variables*					
		DALI			SUS	
	Estimate	*Conf. Int.*	*p-value*	*Estimate*	*Conf. Int.*	*p-value*
Fixed Parts						
(Intercept)	45.32	40.92 – 49.71	<.001	81.22	76.72 – 85.71	<.001
NDRA - Messenger	7.39	3.27 – 11.51	.001	-3.35	-7.56 – 0.86	.126
NDRA - Video	-3.55	-7.66 – 0.56	.098	-0.90	-5.11 – 3.30	.676
Device - CID	1.18	-1.86 – 4.22	.450	-0.63	-3.74 – 2.48	.692
Pre-Warning - Baseline	0.57	-3.57 – 4.71	.789	4.93	0.69 – 9.17	.028
Pre-Warning - Time	-2.03	-6.14 – 2.08	.339	-2.26	-6.46 – 1.95	.299
Random Parts						
σ^2		158.440			165.769	
$\tau_{00, Participant}$		67.810			71.036	
$N_{Participant}$		24			24	
$ICC_{Participant}$		0.300			0.300	
Observations		72			72	

Figure 7: Linear fixed effects model with estimated influence of dependent variables on mental workload (DALI) and usability (SUS) scores.

Reaction Times

A linear fixed effects model regarding response times shows that our independent variables do not have a significant influence on measured reaction times, see Figure 6. Collected response times are: *Median = 2.59 s, Mean = 2.78 s, Min = 0.53 s, Max = 8.89 s*.

Mental Workload and Usability

The higher a SUS-score, the higher perceived usability. The lower a DALI-score, the lower perceived mental workload. A regression model predicts the influence of NDRAs, devices and pre-warnings on our criterion variables mental workload (DALI) and usability (SUS).

In contrast to correcting an e-mail, expected DALI scores decrease significantly for the messenger task (*NDRA - Messenger: Estimate = 7.93; p* = 0.001).

Welch's two sample t-test to find differences between smartphone and CID usage on DALI *(t = 1.17, df = 66.22, p-value = 0.24)* and SUS *(t = -0.57, df = 67.12, p-value = 0.56)* mean scores did not reveal correlations with a p-value smaller

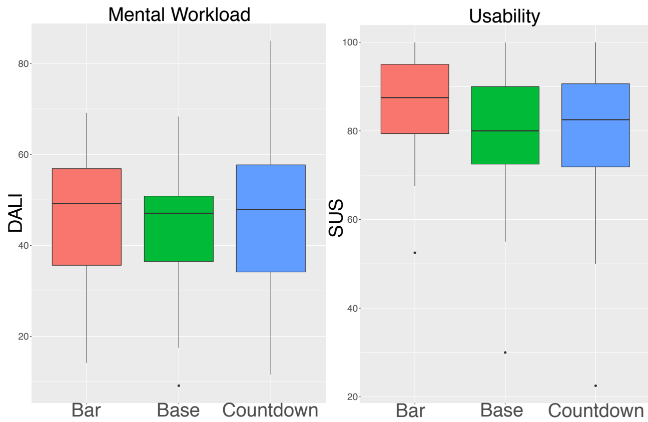

Figure 8: Boxplots with even distribution of means for DALI (workload) and significantly higher SUS (usability) scores in case of a dynamic bar pre-warning.

5%. Mental workload is evenly distributed over pre-warning conditions, see Figure 8.

According to the results, a dynamic bar increases expected usability compared to no pre-warning (Baseline). Hence, not displaying planned TOR leads to significantly lower SUS scores than a visual feedback with a dynamic bar (*Pre-Warning - Baseline: Estimate = 4.93;* $p = 0.028$), see Figure 7 for estimates and p-values. Comparing the usability ratings of the visualizations, the dynamic bar shows significantly higher scores than any other condition, Figure 8 shows a corresponding box plot.

Qualitative Findings

For our qualitative evaluation, we conducted a semi-structured interview, inspected the answers manually, and prepared them for a coding process. We split participants answers into 159 single statements and derived three main categories based on frequency patterns: *modalities for TOR, TOR process and presentation of pre-warnings*. All interview questions were open-ended. The procedure was similar to the work of Eiband et al. [11].

Alternative Pre-Warnings and Procedure

As a first question, we asked the participants if they could think of an alternative pre-warning or TOR procedure. Most stated improvements in context of the categories are:

1. *Modalities for TOR:* Include additional tactile feedback. (41.6 %)

2. *TOR process:* Show TOR on the device which is currently used for the NDRA. (41.6 %)

3. *Presentation of pre-warnings:* Show the remaining time as a numeric value in addition to the continuously moving dynamic bar. (29.2 %)

We did not investigate how single NDRAs were perceived, since they are selected via literature based on profound user surveys and further related work [23, 7, 29].

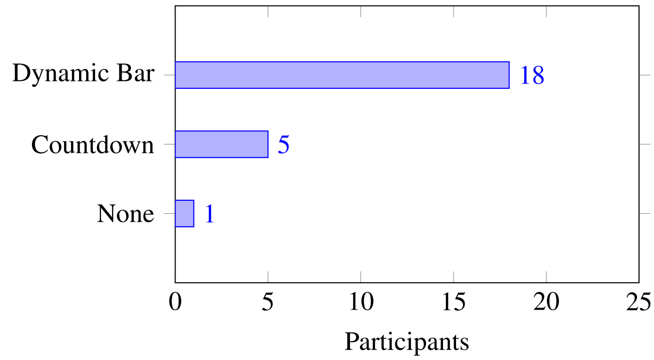

Figure 9: Participants' favorite visual indicators (pre-warnings) of upcoming TOR.

Pre-Warnings

We asked all 24 participants about their preferred type of pre-warning. 18 participants (75 %) favored a dynamic bar indicator for early warnings, see Figure 9. Four of them mentioned that the bar indicator can be perceived peripherally and thereby serves as a helpful tool for mental preparation. Five participants (20.83 %) favored the time countdown as indicator. However, 2 (8.33 %) described it as "pressurizing" and 4 subjects (16.67 %) stated that they ignored the countdown completely, as it was outside their focus of attention. One participant (4.17 %) preferred no graphical pre-warning at all. In his opinion, auditory and visual cues of the TOR were sufficient for a safe take-over.

Device Preferences for Performing NDRAs

In addition to the presentation of pre-warnings, we asked the participants about their preferred device for performing NDRAs. The smartphone was favored by 13 participants (54.17 %). A vast majority of 21 subjects stated that they were used to a smartphone more than using a touch-enabled CID. Furthermore, five participants (20.83 %) stated to prefer a device which can be moved freely. The CID was favored by 11 participants (45.83 %), due to a larger display size. Nine people (37.5 %) stated to prefer the CID for watching videos and the smartphone for messenger applications, see Figure 10. For example, participant 15 stated:

> "I would like to chose the device dependent on the task. For messaging I want to have a smartphone and for watching videos the screen should be as big as possible."

DISCUSSION

According to our results, providing a possibility for users to prepare mentally prior to planned TOR improves driving performance after a control transition, see Figure 5. The majority of Participants preferred a pre-warning visualized with a dynamic bar. Additionally, this graphical representation significantly increased reported usability. Participants stated that a dynamic bar is intuitive and that a mental preparation towards the upcoming TOR gets best supported. Furthermore, it can be understood via peripheral observation, without moving the focus away from the NDRA at hand. Peripheral information processing can benefit drivers in complex driving situations [17, 35]. Many participants suggested a combined display with an abstract dynamic bar including the remaining time or distance as a precise number. Based on our data analysis and the participants statements **we recommend to show pre-warnings as a dynamic bar in combination with a numeric value.**

Regarding the examined devices our results show little influence on the dependent variables. There was an almost equal share for the preferred interaction device (54% smartphone to 46% CID). In addition, several participants stated that their device choice depends on the intended activity. Many prefer to watch videos on the CID due to the larger display size. For the messenger application a smartphone is often favored because people are used to processing messenger applications on a mobile device. Hence, we suggest to offer both devices redundantly in such a way that users can choose which device is better suited for the specific task. Furthermore, copying functionality between a fixed in-car and mobile device includes the benefit for drivers to continue their NDRA without interruptions after leaving the car. For these reasons, **we recommend implementing in-car experiences where users can choose between devices to work on NDRAs.** This includes a presentation of pre-warnings at least on the currently used device.

Vitense et. al state that bimodal feedback positively influences user performance [33]. However, sensory channels of an NDRA could interfere with stimuli of warnings and thus reduce their effectiveness. For example, listening to loud music in the car may mask an audible warning. In this case, tactile and visual signals would be more appropriate. A driver who already focuses the windshield of a car could benefit from a visual reference in a head-up display more than from a warning on the CID or mobile device. If drivers are immersed in a video, tactile impulses could be more effective than visual cues. Such an impulse would neither interfere with audio nor video channels of the film. Hence, the type of NDRA can determine which modality is best suited for a TOR. Therefore, it could be useful for drivers to adapt take-over modalities dependent on the executed task. Alternatively, multiple modalities could be combined for every TOR. The aim is to ensure that drivers are able to understand the request timely and take over control safely. Therefore, **we suggest to present predictable take-over requests using auditory, visual, and tactile modalities.**

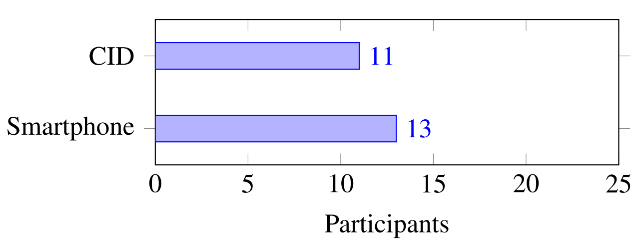

Figure 10: Participants' favorite devices for processing NDRA.

DESIGN RECOMMENDATIONS

As main contribution of this work, we transfer our findings into three recommendations for the design of plannable take-over requests and non-driving-related activities:

1. *Pre-warnings* should be displayed with a *dynamic bar* in combination *with numerical information*.
2. *Processing* of *NDRA* should be possible redundantly with the *CID and a smartphone*.
3. *Predictable take-over requests* should include *auditory, visual, and tactile* cues.

CONCLUSION & FUTURE WORK

We identified recommendations on how to achieve safe and comfortable planned take-over scenarios. To that end, we examined three different representations of pre-warnings for indicating upcoming TOR. Additionally, we took three NDRAs into account and explored which interaction device (smartphone or CID) is more suitable in case of a handover scenario. Therefore, we set up a driving simulator study and explored the influence of dependent variables on usability, mental workload and driving quality. Based on gained results we composed three design recommendations.

In addition to visual representations, future research could investigate other approaches for planned TOR e.g., using speech or tactile cues. Additionally, a generally valid definition of take-over quality remains an open challenge.

The investigated NDRA are based on behavioral observations of passengers on public transport and questionnaires [23]. We took a first step investigating these theoretical insights in an internally valid and reproducible laboratory study [4]. Future work could validated whether drivers actually cover these NDRAs in a naturalistic environment.

REFERENCES

1. Steven M. Belz, Gary S. Robinson, and John G. Casali. 1999. A New Class of Auditory Warning Signals for Complex Systems: Auditory Icons. *Human Factors* 41, 4 (1999), 608–618. DOI: http://dx.doi.org/10.1518/001872099779656734 PMID: 10774131.
2. John Brooke. 2013. SUS: A Retrospective. *J. Usability Studies* 8, 2 (February 2013), 29–40. http://dl.acm.org/citation.cfm?id=2817912.2817913
3. John Brooke and others. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7. http://dag.idi.ntnu.no/IT3402_2009/sus_background.pdf
4. G. E. Burnett. 2011. On-the-Move and in Your Car: An Overview of HCI Issues for In-Car Computing. In *Human-Computer Interaction and Innovation in Handheld, Mobile and Wearable Technologies*. IGI Global, 60–79. DOI: http://dx.doi.org/10.4018/jmhci.2009010104
5. Francesca Cazzulino, Rita V. Burke, Valerie Muller, Helen Arbogast, and Jeffrey S. Upperman. 2014. Cell

Phones and Young Drivers: A Systematic Review Regarding the Association Between Psychological Factors and Prevention. *Traffic Injury Prevention* 15, 3 (2014), 234–242. DOI: http://dx.doi.org/10.1080/15389588.2013.822075

6. Ben Clark, Graham Parkhurst, and Miriam Ricci. 2016. Understanding the socioeconomic adoption scenarios for autonomous vehicles: A literature review. *University of the West of England, Bristol, UK* University of the West of England, Bristol, UK. (2016), 37. http://eprints.uwe.ac.uk/29134
7. Rita Cyganski, Eva Fraedrich, and Barbara Lenz. 2015. Travel-time valuation for automated driving: A use-case-driven study, In 94th Annual Meeting of the Transportation Research Board. *Proceedings of the 94th Annual Meeting of the TRB* (January 2015). http://elib.dlr.de/95260/
8. Daniel Damböck, Mehdi Farid, Lars Tönert, and Klaus Bengler. 2012. Übernahmezeiten beim hochautomatisierten Fahren. *Schwerpunkt Vernetzung, 5. Tagung Fahrerassistenz, München, DE* 15 (2012), 16.
9. Michael J. De Smith. 2013. *Statistical analysis handbook: a comprehensive handbook of statistical concepts, techniques and software tools*. Dr Michael J de Smith. https://www.statsref.com/StatsRefSample.pdf
10. Adele Diederich and Hans Colonius. 2004. Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics* 66, 8 (01 Nov 2004), 1388–1404. DOI: http://dx.doi.org/10.3758/BF03195006
11. Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding Shoulder Surfing in the Wild: Stories from Users and Observers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4254–4265. DOI: http://dx.doi.org/10.1145/3025453.3025636
12. C Gold, D Damböck, L Lorenz, and K Bengler. 2013. "Take over!" How long does it take to get the driver back into the loop?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 57. SAGE Publications, Sage CA: Los Angeles, CA, 1938–1942. http://journals.sagepub.com/doi/abs/10.1177/1541931213571433
13. Christian Gold, Moritz Körber, David Lechner, and Klaus Bengler. 2016. Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density. *Human Factors* 58, 4 (2016), 642–652. DOI: http://dx.doi.org/10.1177/0018720816634226 PMID: 26984515.
14. S. G. Hart and L. E. Stavenland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati (Eds.). Elsevier, Chapter 7, 139–183. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000004342_1999205624.pdf
15. Renate Häuslschmid, Bastian Pfleging, and Andreas Butz. 2017. The Influence of Non-driving-Related Activities on the Driver's Resources and Performance. In *Automotive User Interfaces: Creating Interactive Experiences in the Car*, Gerrit Meixner and Christian Müller (Eds.). Springer International Publishing, Cham, 215–247. DOI: http://dx.doi.org/10.1007/978-3-319-49448-7_8
16. William J Horrey and Christopher D Wickens. 2004. Driving and side task performance: The effects of display clutter, separation, and modality. *Human factors* 46, 4 (2004), 611–624.
17. Am Isomura, D Kamiya, and K Hamatani. 1993. Driver's cognition in peripheral field of view (SAE Paper 931876). *Warrendale, PA: Society of Automotive Engineers* (1993), 8. DOI:http://dx.doi.org/10.4271/931876 International Pacific Conference On Automotive Engineering.
18. Matthias Kuehn, Tobias Vogelpohl, and Mark Vollrath. 2017. Takeover Times in Highly Automated Driving (Level 3). In *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*. Unfallvorschung der Versicherer (UDV), Wilhelmstraße 43 / 43G; 10117 Berlin, 11. http://indexsmart.mirasmart.com/25esv/PDFfiles/25ESV-000027.pdf
19. Patrick Langdon, Ioannis Politis, Mike Bradley, Lee Skrypchuk, Alex Mouzakitis, and John Clarkson. 2018. Obtaining Design Requirements from the Public Understanding of Driverless Technology. In *Advances in Human Aspects of Transportation*, Neville A Stanton (Ed.). Springer International Publishing, Cham, 749–759.
20. Vivien Melcher, Stefan Rauh, Frederik Diederichs, Harald Widlroither, and Wilhelm Bauer. 2015. Take-over requests for automated driving. *Procedia Manufacturing* 3 (2015), 2867–2873.
21. Frederik Naujoks, Christoph Mai, and Alexandra Neukum. 2014. The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Advances in Human Aspects of Transportation* 7, Part I (07 2014), 431.
22. I Petermann-Stock, L Hackenberg, T Muhr, J Josten, and L Eckstein. 2015. Bitte übernehmen Sie das Fahren!": Ein multimodaler Vergleich von Übernahmestrategien. *AAET. Google Scholar* (2015).
23. Bastian Pfleging, Maurice Rang, and Nora Broy. 2016. Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th international conference on mobile and ubiquitous multimedia*. ACM, NY, USA, 91–99. DOI: http://dx.doi.org/10.1145/3012709.3012735 ISBN: 978-1-4503-4860-7.
24. Ioannis Politis, S Brewster, and F Pollick. 2014. The Effects of Modality, Urgency and Message Content on Responses to Multimodal Driver Displays. *AutomotiveUI EA 2014, ACM* AutoUI'14 (2014), 5.

25. Ioannis Politis, Stephen Brewster, and Frank Pollick. 2017. Using Multimodal Displays to Signify Critical Handovers of Control to Distracted Autonomous Car Drivers. *Int. J. Mob. Hum. Comput. Interact.* 9, 3 (July 2017), 1–16. DOI: http://dx.doi.org/10.4018/ijmhci.2017070101

26. Jonas Radlmayr, Christian Gold, Lutz Lorenz, Mehdi Farid, and Klaus Bengler. 2014. How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58. Sage Publications Sage CA: Los Angeles, CA, SAGE journals, sagepub, 2063–2067. DOI: http://dx.doi.org/10.1177/1541931214581434

27. Thomas A Ranney, Elizabeth Mazzae, Riley Garrott, and Michael J Goodman. 2000. NHTSA driver distraction research: Past, present, and future. In *17th International Technical Conference on the Enhanced Safety of Vehicles*, Vol. 2000. NHTSA, https://www.nhtsa.gov/, 11.

28. SAE. 2014. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. *Standard J3016* 1, 1 (2014), 30. DOI: http://dx.doi.org/10.4271/J3016_201609

29. Brandon Schoettle and Michael Sivak. 2014. A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia. (2014).

30. Bobbie Seppelt and Christopher Wickens. 2003. In-Vehicle Tasks: Effects of Modality, Driving Relevance, and Redundancy. (08 2003).

31. University of Iowa 2007. *Driver's Behavior and Workload Assessment for New In-vehicle Technologies Design*. University of Iowa, IA: Public Policy Center, Iowa City, IA. DOI: http://dx.doi.org/10.17077/drivingassessment.1276

32. Remo van der Heiden, Shamsi T Iqbal, and Christian P Janssen. 2017. Priming Drivers before Handover in Semi-Autonomous Cars. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, NY, USA, 392–404. DOI: http://dx.doi.org/10.1145/3025453.3025507

33. H. Vitense, J.Jacko, and V.Emery. 2003. Multimodal feedback: an assessment of performance and mental workload. *Ergonomics* 46, 1-3 (2003), 68–87. DOI: http://dx.doi.org/10.1080/00140130303534 PMID: 12554399.

34. Marcel Walch, Kristin Lange, Martin Baumann, and Michael Weber. 2015. Autonomous Driving: Investigating the Feasibility of Car-driver Handover Assistance. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '15)*. ACM, New York, NY, USA, 11–18. DOI: http://dx.doi.org/10.1145/2799250.2799268

35. Julia Werneke and Mark Vollrath. 2011. Signal evaluation environment: a new method for the design of peripheral in-vehicle warning signals. *Behavior Research Methods* 43, 2 (01 Jun 2011), 537–547. DOI: http://dx.doi.org/10.3758/s13428-010-0054-8

36. Christopher D. Wickens. 2002. Multiple resources and performance prediction. *Theoretical issues in ergonomics science* 3, 2 (2002), 159–177.