Drivers' Ability to Engage in a Non-Driving Related Task While in Automated Driving Mode in Real Traffic

(article starts on next page)

Drivers' Ability to Engage in a Non-Driving Related Task While in Automated Driving Mode in Real Traffic

MARIA KLINGEGÅRD[1,2], JONAS ANDERSSON[1], AZRA HABIBOVIC[1], EMMA NILSSON[3,4], AND ANNIE RYDSTRÖM[3]

[1]RISE Research Institutes of Sweden, 417 56 Göteborg, Sweden
[2]Folksam Insurance Group, 116 67 Stockholm, Sweden
[3]Volvo Cars, 418 78 Göteborg, Sweden
[4]Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 412 96 Göteborg, Sweden

Corresponding author: Jonas Andersson (jonas.andersson@ri.se)

ABSTRACT Engaging in non-driving related tasks (NDRTs) while driving can be considered distracting and safety detrimental. However, with the introduction of highly automated driving systems that relieve drivers from driving, more NDRTs will be feasible. In fact, many car manufacturers emphasize that one of the main advantages with automated cars is that it "frees up time" for other activities while on the move. This paper investigates how well drivers are able to engage in an NDRT while in automated driving mode (i.e., SAE Level 4) in real traffic, via a Wizard of Oz platform. The NDRT was designed to be visually and cognitively demanding and require manual interaction. The results show that the drivers' attention to a great extent shifted from the road ahead towards the NDRT. Participants could perform the NDRT equally well as when in an office (e.g. correct answers, time to completion), showing that the performance did not deteriorate when in the automated vehicle. Yet, many participants indicated that they noted and reacted to environmental changes and sudden changes in vehicle motion. Participants were also surprised by their own ability to, with ease, disconnect from driving. The presented study extends previous research by identifying that drivers to a high extent are able to engage in a NDRT while in automated mode in real traffic. This is promising for future of automated cars ability to "free up time" and enable drivers to engage in non-driving related activities.

INDEX TERMS Automated driving, driver behavior, driver experience, non-driving related task, secondary task.

I. INTRODUCTION

The current automation and digitalization trends are transforming the driving experience [1], [2]. The number of driving support systems such as speed control, lane centering assistance and traffic jam assistance are increasing. The driver's role is shifting from actively controlling the vehicle via steering, breaking, and accelerating to supervising the automatic functions or even being a passenger. The SAE international classification system classifies automated functionality in vehicles into different levels [3]. SAE Level 0-2 vehicles correspond to manual vehicles in which the driver is in control with limited automation. SAE level 3 vehicles correspond to vehicles in which the driver is responsible for monitoring the situation and being ready to take over from automation within a specific timeframe, denoted in this paper as "*supervised* automated driving mode". SAE level 4 vehicles, on the other hand, correspond to vehicles that are automated to such an extent that the driver is allowed to do other activities, denoted in this paper as "*unsupervised* automated driving mode". SAE Level 5 vehicles are vehicles that can drive without any input from a human driver.

Introducing unsupervised automated driving mode functionality in vehicles opens new possibilities for the driver to perform non-driving related tasks (NDRTs); tasks that

The associate editor coordinating the review of this manuscript and approving it for publication was Pedro Neto.

normally are considered as distracting and dangerous. It is suggested that the increased level of automation in the vehicles will free up time for drivers and increase productivity during time spent in traffic jams and on motorways [2] by creating a "mobile office" [4] where they can perform NDRTs. The underlying assumption is that while being in the automated vehicle one can become more productive as one could perform tasks which are normally done, for example, in an office, extending available time and space. Several surveys have also identified positive attitudes in drivers towards NDRT engagement, as well as explored what activities people would like to perform if they had the possibility to travel in a highly automated vehicle (SAE Level 4-5) (see e.g., [5]–[7] for further information). Simulator studies to some extent confirm this by showing that with increasing levels of automation drivers are more likely to engage in NDRTs [8]–[10]. It has been shown that drivers are comfortable in adopting NDRTs and that they quickly engage in them if given the possibility, both in supervised [11] and in unsupervised automated driving mode [12].

Despite being one of the promises of automated vehicles [2], [4], there is a lack of public empirical studies that investigate to what extent people are in fact able to engage, and maintain engagement, in NDRTs while traveling in unsupervised automated mode in real traffic. That is, it has not yet been determined whether one can perform an NDRT in a vehicle equally well as in an office. In several studies, NDRTs are used as tools to investigate their effects on driving performance in for example takeover requests (see [13] for a review), with less interest in the completion of the NDRT itself (cf. [14]).

From previous studies on engagement in NDRTs in manual driving it is known that drivers adapt their driving behavior and task engagement to the traffic situation, roadway and environmental conditions (e.g., [15]–[19]). Actually, similar patterns have been seen in supervised automated driving mode when the drivers are responsible for monitoring and responding to automation. For instance, a recent simulator study with a self-paced NDRT showed that the task was rejected prior to a takeover situation, but once already engaged in the task the participants continued despite the upcoming takeover situation [9]. An on-road study [20] testing task engagement showed a strong dependency between task execution and velocity, with tasks mainly being performed during low-speed driving. It has also been shown that the level of experienced flow (i.e. level of engagement) affected the reaction time to take over requests while in supervised automation mode [21].

Several studies have witnessed that drivers become engaged and maintain engagement in NDRTs, both while in manually operated vehicles as well as in supervised vehicles (where the driver is responsible for the overall safety and driving activities). It is however yet to be determined to what extent drivers are able to engage in an NDRT while travelling in unsupervised automated driving mode, where the driver is not requested to monitor the traffic situation or supervise the automation. In this study, the assumption was that drivers are able to perform an NDRT while traveling in unsupervised mode with similar performance as in an office.

A. GOAL, OBJECTIVE, RESEARCH QUESTIONS
The goal of the presented research was to explore if automated driving can fulfil its promise of being a platform that "frees up time" for other activities while on the move. The objective of this on-road experiment was to investigate to what extent drivers can shift attention to and perform a cognitively, visually and motorically demanding NDRT in unsupervised automated driving mode in real traffic. The following research question was considered (Figure 1): How well are drivers able to engage in an NDRT while in unsupervised automated driving mode in real traffic? More specifically:

1. How well are drivers able to shift attention to the NDRT while in unsupervised automated driving mode?
2. How well are drivers able to perform the NDRT while in unsupervised automated driving mode?
3. How do drivers experience engaging in the NDRT while in unsupervised automated driving mode?

A specific NDRT was created with visual ("eyes off road"), cognitive ("mind off road") and motoric ("hands off wheel") demands to determine drivers' task engagement. The drivers' ability to engage in the NDRT while in unsupervised automated driving mode was determined by investigating their 1) ability to shift attention to the NDRT by analyzing visual behavior, 2) NDRT performance by contrasting the performance on the road and in the office, and 3) experience of engaging in the NDRT on the road by self-assessment and post-trial interviews.

II. METHOD
In this section, it is described how the study was conducted as well as the methodologies and equipment used. It should be noted that an overview of the complete experiment is provided, and as such, includes a set up beyond the scope of the research questions presented in this paper.

A. OVERALL EXPERIMENTAL SET UP
The study was conducted on a public highway in Gothenburg, Sweden. It involved 10 participants who repeatedly experienced manual driving and two modes of automated driving – supervised mode and unsupervised mode (in which an NDRT was performed).

These two automated modes differed only in the verbal explanation of the driver's responsibility, where the supervised mode meant that the driver was responsible for safe driving, while the unsupervised mode posed all responsibility for safe driving on the car. The automated driving was, to some extent, simulated using an automated vehicle test platform. Each participant performed a designated NDRT twice per test occasion, once in the office and once while traveling in the unsupervised mode (balanced by test occasion). The effects of drivers' ability to engage and shift focus were
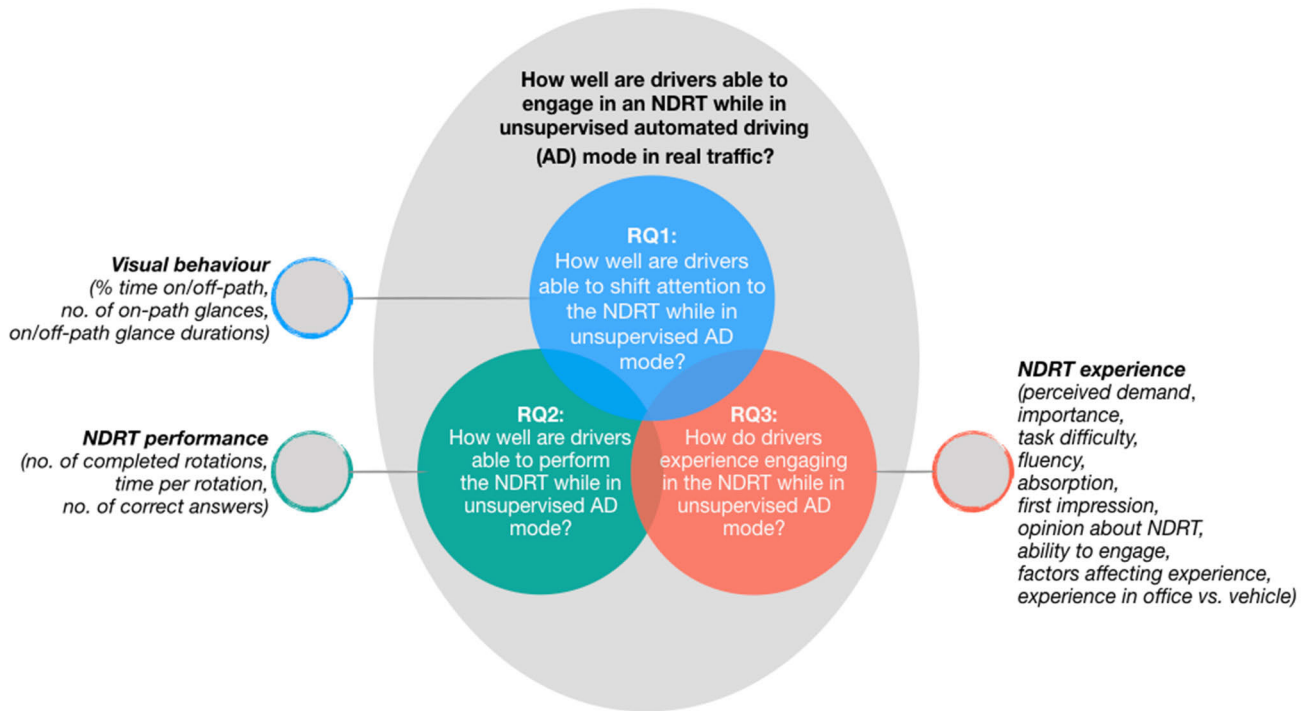
measured using various qualitative and quantitative methods. To account for habituation and learning effects, the drivers took part in the study twice, with about one week in-between the occasions.

B. PARTICIPANTS

The study involved 10 participants (4 male, 6 female) with an age range between 18 and 51 years (3 in the age range 18-30, 3 in the age range 31-40, and 4 in the age range 41-51). All participants had a driving license and drove a passenger car regularly. However, none of them had previous experience of automated driving. Eight drivers stated that they were novice or had no previous knowledge in automated driving, while two drivers stated that they were knowledgeable or expert. Six of them stated that they have some experience of using the driving assistance function Volvo Pilot Assist (provides lateral and longitudinal control assistance). On a technology readiness scale, four drivers classified themselves as early adopters, one as early majority, four as majority, and one as late majority. The drivers were employed at Volvo Cars and RISE and were recruited through internal databases at these organizations.

C. DATA COLLECTED

The data collected in the study include both quantitative and qualitative data (Figure 1). The quantitative data consisted of:

Video: From which participants' overall behavior (in vehicle and in the office) and visual behavior was analyzed. Only visual behavior is presented here.

Self-assessments: Background questionnaire, repeated questionnaire (not elaborated upon here), NDRT questionnaire and post questionnaire (not elaborated upon here).

NDRT performance: Response time, correct or incorrect response.

Vehicle data: Speed, acceleration, deceleration, brake onset/offset (which are not elaborated upon here).

Physiological data: Electroencephalography (EEG), electrocardiography (ECG), electrooculography (EOG) and respiratory inductance plethysmography (RIP) (which are not elaborated upon here).

The qualitative data consisted of interviews with the participants. The interviews focused on capturing the overall driver experience of using the automation and of conducting the NDRT in the vehicle and in the office.

D. MATERIAL AND EQUIPMENT

Data were collected using various materials and equipment, including an automated vehicle test platform and a data collection system. These are described in more detail in the following sections.

1) WIZARD OF OZ PLATFORM

The study was conducted using an automated vehicle test platform – a Wizard of Oz (WOz) vehicle, which is used to simulate automated driving (cf. [12], [22], [23]). The platform is based on a Volvo XC90 and is equipped with additional sensors and vehicle controls in the backseat. These controls were concealed for the participant who was seated in the front

behind the steering wheel. A "driving wizard" was seated in the backseat and was responsible for the driving when the automated driving mode was activated. The car is approved for public roads by the local Swedish authorities and the study was ethically approved by the Swedish Ethical Review Authority (2019-01827).

The interface was controlled by the "interface wizard" from the back seat using a small tablet. An icon showing that automation is available was presented in the instrument cluster display and was supplemented by a sound. To activate (or deactivate) the function the participant had to simultaneously press two buttons on each side of the steering wheel for 2 s. When the participant was to deactivate the function on request from the automated function, a take-over request was issued by the interface wizard, consisting of visual information as well as an audio signal. The participant had 800 m (about 35 s) to resume control of the car. If the participant would fail taking over the control within 800 m, the car was supposed to continue travelling in the automated mode. However, all participants managed to take over accordingly and this was not applied in practice.

The vehicle simulated supervised mode and unsupervised mode, which differed only in the information on responsibility that was orally given by the test leader to the participants. In the supervised mode, the participants were told: "You are the driver and you are responsible for the overall safety". In the unsupervised mode, they were told: "Volvo Cars are responsible for the safety onboard after you have activated the function".

2) CAR DATA LOGGER AND CAMERAS
The car was equipped with a Dewesoft S-Box measurement system, which was logging data from the controller area network of the car as well as video data from four webcam cameras of the type Logitech C90E. Video data was recorded from the following views; facing driver from the front, facing driver from the front right, facing driver and displays from the back, and facing the road ahead. In addition, a fifth camera (GoPro) was used to record the behavior of the participant while conducting the NDRT at the office.

3) THE NON-DRIVING RELATED TASK
The NDRT was based on mental rotation [24], [25] and was designed to shift the driver's attention in such a way that the NDRT became the "primary task". It was designed to be cognitively demanding as well as visually and motorically intensive, corresponding to a driver state of "mind off", "eyes off" and "hands off". It was visually presented on a tablet and demanded physical interaction (button presses). Time criticality was introduced (a set time limit and a timer counting down) to further enhance the shift in driver attention (Figure 2).

The mental rotation task consisted of 86 polygons and was implemented using the PsyToolkit web service for psychological experiments [26], [27]. In each trial, a set of three polygons (reference, correct, mirrored) was displayed in a
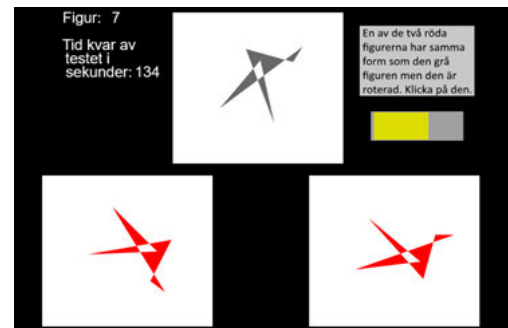

FIGURE 2. The mental rotation task with the reference (top), correct (bottom left) and mirrored (bottom right) polygons.

randomized order, with a condition that all sets must be shown once before any set is repeated (Figure 2). The reference polygon incorporated four to nine edges. The correct polygon had the same shape as the reference, while the mirrored polygon was a mirrored version of the reference figure. The correct and the mirrored polygons were rotated compared to the reference one and to each other. They were displayed side-by-side in a randomized order, while the reference polygon was always displayed on the top.

The task of the participant was to determine which of the lower polygons is the correct one by tapping on the selected polygon (Figure 2). The maximum response time per set of polygons was 7 s (indicated by a progress bar). If the participant did not answer within this time limit, a message was displayed for one second to answer faster. Directly after answering, it was displayed during 1 s whether the answer was correct. The total length of the experiment was three minutes. Prior to the experiment, the participants took part in a training session consisting of additional 5 sets of polygons.

4) SELF-ASSESSMENT QUESTIONNAIRES
A total of four questionnaires were used:

Background Questionnaire consisting of multi-option questions on demographics, experience of driver assistance systems, knowledge of automated driving, as well as technical readiness (completed on paper).

Repeated Questionnaire assessing how much participants trust their own driving and/or automated driving, perceived safety [23], as well as how good they find the driving capability (own/automated) [28]. This was asked verbally after each road segment at approximately the same location. The answers were on a 10-grade scale and were noted on paper. (This data is not reported as it was not relevant for the research questions reported in this paper).

NDRT Questionnaire assessing the experience of conducting the NDRT. It was completed on the iPad directly after completing the NDRT (i.e. in the office or while still in unsupervised mode). It was based on flow theory [29] and mental workload theory [30] as in line with [21]. It consisted of items from the flow short scale (FKS): fluency, absorption, perceived demand, perceived importance and task difficulty.

To assess workload the simplified SWAT technique was used with the items mental effort load, time load and psychological stress load [30]. All answers were on a 10-grade scale, except for SWAT that uses a scale 1-3.

Post Questionnaire assessing participant's experience of the test equipment. (This data is not reported in this paper as it was not relevant for the research questions.)

5) INTERVIEWS

The qualitative data consisted of interviews that were audio recorded using a digital audio recorder. The interviews were conducted in Swedish or English, depending on the participant's preference, and were led by the test leader. After being recorded, they were transcribed by one of the authors, and translated into English. The interviews were semi-structured and focused on contrasting the experience of performing the NDRT in the vehicle and in the office from a user experience perspective. Events that contributed to the NDRT experience were identified and contextual factors contributing to that experience were elaborated upon, in line with the critical incident method [31]. The test leader had interview guidelines at his/her disposal.

E. TEST ENVIRONMENT

The study was performed in September/October 2019 on a public motorway in Gothenburg, Sweden. A road segment of approximately 9 km between the motorway interchanges Fiskebäcksmotet and Åbromotet was chosen (Figure 3). The participants travelled four times in each direction, that is, in a loop between the interchanges. The road segment involves merging lanes at a few instances, but it does not include any controlled intersections and has no major known traffic congestions, and thus allows for a continuous automated driving experience. The posted speed limit varied between 70 and 90 km/h. The WOz vehicle has a maximum allowed speed of 80 km/h. There was ongoing roadworks immediately after the on-ramp at Fiskebäcksmotet, and at this location the vehicle was always operated in manual driving mode (i.e. the automated function was made available after the roadworks). The study was conducted under daylight conditions and weather conditions were varying, from no rain to rain. The tests were scheduled to avoid morning and afternoon rush hours. The road segment included variations in traffic flow and density, and the situations drivers experienced.

F. PROCEDURE

The participants took part in two succeeding occasions with about one week in-between (Figure 4). Upon their arrival at the office, they were informed by the test leader about the study and were asked to complete an informed consent and the Background Questionnaire. They were then given information in verbal and written form about unsupervised and supervised automation modes. Next, the equipment for physiological measurements was fitted by an assistant. During the first occasion, the participants were then asked to complete the NDRT in the office (for the second occasion
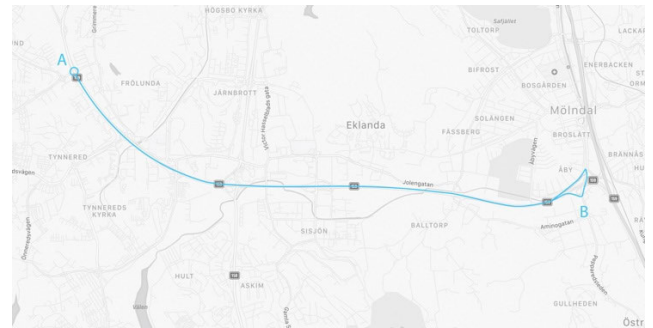
FIGURE 3. A map of the road segment used in the study. One drive from A to B corresponds to one lap.

this was done after the driving session). The participants were informed that there is a training session prior to the NDRT experiment, including a set of instructions. In essence, they were instructed to determine which of the lower polygons is the correct one by tapping on the selected polygon as fast as possible. At last, the participants were asked to be seated in the WOz vehicle. After a brief introduction to the vehicle, they got a demonstration of the interface for automated driving and practiced a few minutes on using it while in standstill (the interface for supervised and unsupervised driving was the same). Next, the participants were instructed to drive to the selected road segment and complete a total of 8 laps (cf. Lap 1-8 in Figure 4), they traveled four times in the direction A to B, and four times in the direction B to A (Figure 3). They were asked to stay in the middle lane and avoid changing lanes, unless it was needed for safety reasons. In Lap 1, the participants practiced how to follow the instructions and how to activate and deactivate the automated driving function. The interface wizard (who also was the test leader) prompted the function as available, and requested to disengage, several times for practice reason. Prior to Lap 2-8, the test leader told the participant what his/her task would be; manual driving, supervised automated driving ("You are the driver and you are responsible for the overall safety"), or unsupervised automated driving with/without NDRT ("Volvo Cars are responsible for the safety onboard after you have activated the function"). Once it became available, it was up to each participant to decide when to activate and deactivate the automated driving function. As the end of the road segment was reached (cf. A and B in Figure 3), the participant was prompted by the system to disengage the automation and then instructed by the test leader to exit the highway. While exiting the highway, the test leader asked the participant the questions in the Repeated Questionnaire and noted the answers. The automated driving function was made available at the same locations for all participants, and the order of laps (Figure 4) was the same for all participants on both occasions. The NDRT task was conducted during Lap 7 while in the unsupervised driving mode. The test leader handed the tablet to the participant as soon as the unsupervised mode was activated. The participant was encouraged to complete the NDRT using the tablet as good and as fast as possible.
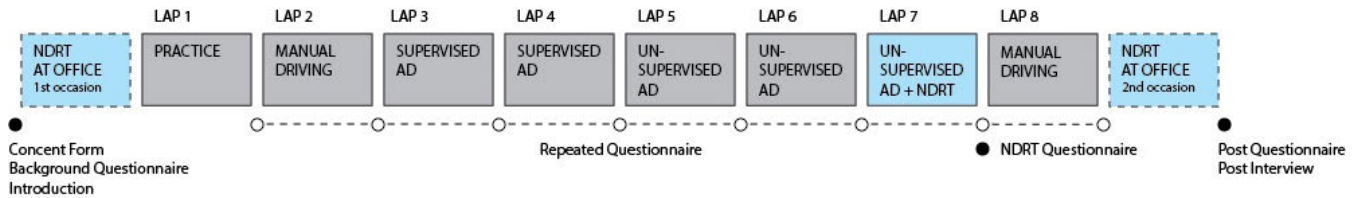
FIGURE 4. Overall procedure of the test. For occasion 1, the NDRT was performed at the start. The following week, at occasion 2, the NDRT was performed at the end.

However, it was up to the participant to decide how much he/she engaged in the task and at what pace. In total, the study took about 3 hours to complete (out of which 90 minutes were spent on the driving session). After the driving session, the participants took part in the Interview and competed the Post Questionnaire.

G. DATA ANALYSIS
This section describes how the data was analyzed to answer the research questions presented in Figure 1.

1) VISUAL BEHAVIOR
To facilitate an analysis of visual behavior, the video data from the lap with NDRT was manually annotated using the camera facing the driver from the front. The annotations were made in Dewesoft and the data were then imported into MATLAB for processing and analysis. The analysis included two glance directions: *on path* (eyes directed out the forward windshield towards the direction of the vehicle's travel) and *off path* (eyes directed to other areas than the direction of the vehicle's travel). A glance started when the eye movement towards the new glance direction started and ended when the eye movement towards the next area started [32]. Eye blinks shorter than 0.3 s were included in the current glance direction. There were no eye blinks longer than 0.3 s in the data sample.

2) NDRT TASK PERFORMANCE
The task performance was analyzed in terms of number of completed rotations, time per rotation (ms) and number of correct answers. Learning effects were assessed by comparing the performance between the first and second occasion in the office. The habituation effects of experiencing automated driving for the first time were examined by comparing the performance between the first and second occasion in the vehicle. A performance comparison between the vehicle and office assessed the ability to engage in a NDRT, where a poor(er) performance in the car would indicate inability to fully engage in the given task.

3) SUBJECTIVE EXPERIENCE
The subjective experience was assessed from ratings of the task considering perceived demand, perceived importance, task difficulty, fluency, absorption, time load, mental demand and stress. Learning effects were assessed in similar way as to task performance.

The interviews were analyzed by one of the authors based on the principles of grounded theory [33]. All interviews (in total 20 interviews) were transcribed and analyzed. Grounded theory consists of a series of activities, however, due to the nature of the research question only the first stage of the analysis, "open coding" was employed. It analyses qualitative data by extracting concepts from the data in which similarities and differences are identified across the participants. A combination of coding of text passages line-by-line with conceptual codes and a guided search was performed to characterize task and driving experience. Each segment was first coded by (1) occasion (first/ second) and (2) location (office/car) together with participant ID. Within the segments, general themes and trends emerged and were identified as a category (Table 5 and Table 6). Each category was provided with a thematic description and the corresponding citations.

4) STATISTICAL ANALYSIS
The statistical analyses were chosen and performed in accordance with Laerd statistical guides (https://statistics.laerd.com). IBM SPSS version 26 software package was used for the statistical testing. Table 1 to 5 presents the results from the statistical analyses.

III. RESULTS
This on-road experiment aimed to investigate the effect of, and possibility for, drivers to engage in an NDRT while driving in unsupervised automated driving mode. The presented data includes glance behavior, task performance, NDRT questionnaires, and interviews (Figure 1). For specification of test occasions see Figure 3 and 4.

A. HOW WELL ARE DRIVERS ABLE TO SHIFT ATTENTION TO THE NDRT?
The glance analysis showed that all participants looked away from the forward path during the vast majority of the task execution (Figure 5). All drives had a gaze direction that was on-path for less than 20% of the total time of NDRT execution. In 50% of the drives, the gaze direction was on-path for less than 2% of the time.

The number of times that the test participants looked up at the road during the three minutes of task execution differed between zero and 44 times (Figure 6). Cumulative glance duration distributions for on- respectively off-path glances are plotted in Figure 7. Among the on-path glances, approximately 95% have a duration of less than one second and
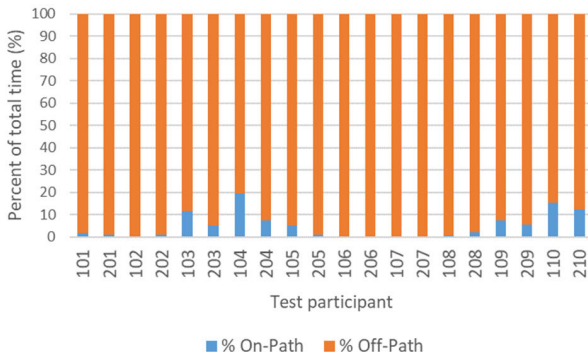
FIGURE 5. The percent of total time that the participants had a glance-direction on-path respectively off-path for each occasion. The first number in the three-digit test participant sequence represents the test occasion (1 or 2) and the following two numbers represents the test participant (01, 02, 03, etc.).

FIGURE 6. The number of on-path glances during task execution for each occasion. The first number in the three-digit test participant sequence represents the test occasion (1 or 2) and the following two numbers represents the test participant (01, 02, 03, etc.).
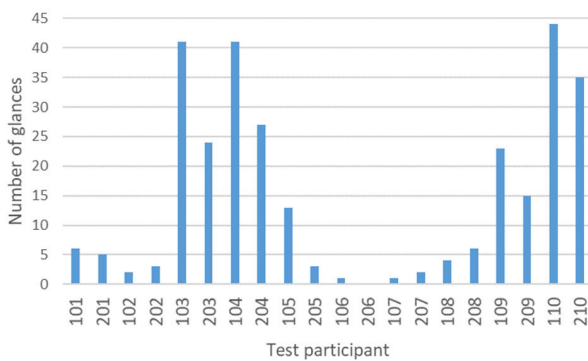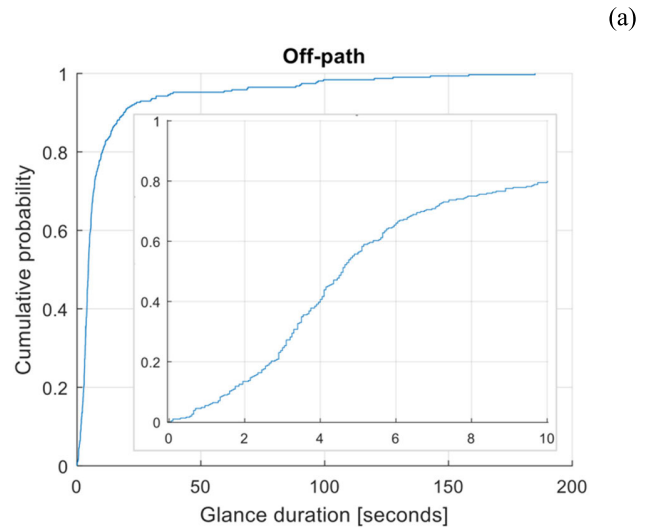
(a)

(b)

FIGURE 7. Cumulative glance duration distributions for (a) off-path glances, and (b) on-path glances.

TABLE 1. Analysis of glance data using Wilcoxon signed rank test.

Measure		Occasion 1	Occasion 2	Statistic
% on-path	m	6.22	3.60	$z=-1.89$,
	M	3.44	1.72	$p=.059$
	StD	6.99	3.93	
# on-path	m	17.60	12.00	$z=-1.80$,
glances	M	9.50	5.50	$p=.073$
	StD	18.12	12.46	
On-path	m	0.50	0.46	$z=-.578$,
glance	M	0.53	0.50	$p=.563$
durations	StD	0.11	0.19	

m: mean, M: median

50% have a duration of less than half a second. Among the off-path glances, more than 85% have a duration above two seconds and 60% have a duration above four seconds. There were no statistically significant differences between the first and second test occasion in the proportion of on-path glances, the number of on-path glances, nor in the on-path glance durations (Table 1).

It is however noteworthy that the five participants that clearly stands out with the highest levels in both proportion on-path glance time and number of on-path glances (participants ×03, ×04, ×05, ×09 and ×10) all show distinct decreases in both measures between the first and second test occasion.

B. HOW WELL CAN DRIVERS PERFORM THE NDRT?

Looking at all test occasions, the results show that the participants were able to complete 38.9 ± 5.6 (mean $\pm$ SD) rotations, with 33.9 ± 8.3 correct answers, needing 3880 ± 964 ms per rotation in the vehicle. Likewise, the participants were able to complete 36.2 ± 6.9 rotations in the office
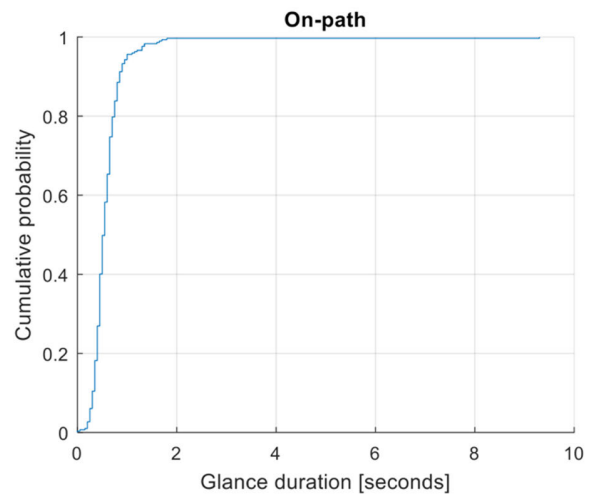
with 31.3 ± 8.6 number of correct answers, needing 4157 ± 1001 ms per rotation. Paired sample t-tests were used to examine differences in task performance in the car and in the office. For number of mental rotations, the results show

a difference in the first and last test occasion (both in the office), where the second occasion in the office yielded an improved performance ($m = 37.8$, $SD = 5.9$) compared to the first occasion ($m = 33.2$, $SD = 6.2$), $t(8)=-2.559$, $p = 0.034$, $d = .853$ (Figure 8), indicating a learning effect. The other measures followed the same pattern where time per rotation decreased and correct answers increased the second time in the office. No outliers were identified in the data and the assumption of normality was not violated in any of the measures, as assessed by the Shapiro-Wilk's test of normality ($p > .05$ for all measures).
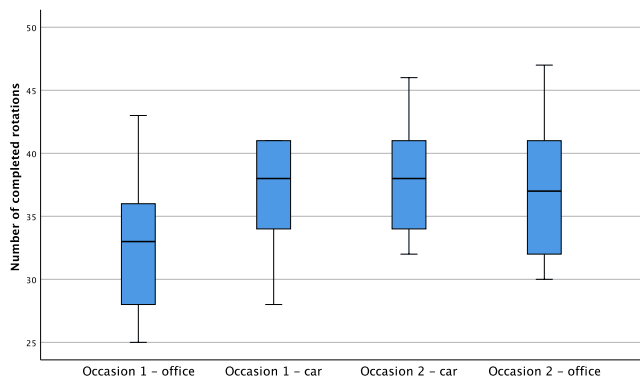

FIGURE 8. Number of completed mental rotations in the NDRT. Please observe that y-axis scale starts at 25 rotations.

The results also show no statistically significant differences for either number of completed mental rotation images, time per image or number of correct answers between the first and second occasion in the car. The results thus indicate that there was no habituation effect from riding in an automated vehicle on NDRT performance.

To account for the learning effect shown between first- and second-time performance of the NDRT, the second in-vehicle occasion were compared to the second occasion in the office (when the drivers were more trained in the NDRT). The results show no statistically significant differences for the number of completed mental rotation images, time per image or number of correct answers (Table 2). Thus, it can be concluded that the drivers were able to perform the NDRT equally well in the car as in the office environment.

C. HOW DO PEOPLE EXPERIENCE ENGAGING IN THE NDRT?

1) NDRT QUESTIONNAIRE: EXPERIENCED LEVEL OF TASK ENGAGEMENT

Table 3 presents subjective ratings of experienced task demand, importance and task difficulty along with a comparison between office and car. To avoid potential learning effects from the first occasions, the second occasion in the car was compared to the second occasion in the office. The results show no statistically significant differences for perceived demand, perceived importance and task difficulty. Although not statistically significant, it can be noted that the perceived demand ratings are leaning towards higher

TABLE 2. Comparison of NDRT performance using a paired sample t-test.

Item		Office	Car	Statistic
# mental rotations	m	38.9	39.5	t (9) = 1.068, p=.313
	M	38.0	39.0	
	StD	6.6	5.7	
Time per image (ms)	m	3757	3963	t (9) = .590, p=.570
	M	3702	3808	
	StD	834	1166	
# correct answers	m	33.9	34.6	t (9) = -.685, p=.511
	M	34.5	35.0	
	StD	8.1	7.6	

m: mean, M: median

TABLE 3. Comparison of task experience using a wilcoxon signed rank test.

Item		Office	Car	Statistic
Perceived demand *"I felt somewhat challenged by the task"*	m	2.90	3.40	z=1.890, p=.059
	M	3.0	3.5	6 rated higher in car
	StD	1.29	1.71	1 rated lower in car
				3 rated same
Perceived importance *"The task was of great importance to me"*	m	2.70	3.10	z=1.265, p=.206
	M	1.5	2.5	6 rated higher in car
	StD	2.45	2.28	1 rated lower in car
				3 rated same
Task difficulty *"How do you judge the task's difficulty?"*	m	4.44	4.89	z=.962, p=.336
	M	4.0	5.0	4 rated higher in car
	StD	1.33	2.15	1 rated lower in car
				4 rated same

m: mean, M: median, Completely agree - Completely disagree (1-10), Easy - Difficult (1-10)

ratings in the car. Moreover, the task engagement ratings (in absolute numbers) show that the participants experienced fluency and became highly absorbed by the task (Table 4). No statistically significant differences were found comparing the task engagement ratings between the second occasions in the car and in the office. Again, it should be noted that results are leaning towards higher ratings in the car (i.e. were less engaged compared to office) while ratings are still on the low end of the scale (i.e. overall high level of engagement in the NDRT).

Additionally, the NDRT was assessed using SWAT. Similarly, the second occasion in the office was compared with the second occasion in the car. The SWAT ratings show that the participants experienced moderate levels of mental workload while performing the NDRT (Table 5). There were no significant differences found between the time load, mental demand and stress experienced in the office and on the road.

D. POST INTERVIEW: CHARACTERISING THE NATURE OF PERFORMING NDRT

The interview started by identifying the participants' spontaneous reaction to what they experienced as they did not have

TABLE 4. Comparison of task engagement using a wilcoxon signed rank test.

Item		Office	Car	Statistic
Fluency	m	1.90	2.60	z=1.382, *p*=.167
"I could concentrate	M	1.0	3.0	5 rated higher in car
on the task the	StD	1.37	1.35	1 rated lower in car
whole time."				4 rated same
Absorption	m	2.20	2.90	z=1.403, *p*=.161
"I was totally	M	2.0	3.0	5 rated higher in car
absorbed by doing	StD	1.03	1.44	2 rated lower in car
the task"				3 rated same

m: mean, M: median, Completely agree - Completely disagree (Scale 1-10)

TABLE 5. Comparison of task demand from swat using a wilcoxon signed rank test.

Iteam		Office	Car	Statistics
Time load	m	2.00	2.13	z=0.577, *p*=.564
How much time	M	2.0	2.0	2 rated higher in car
pressure did you	StD	0.77	0.64	1 rated lower in car
experience?				5 rated same
Mental demand	m	2.13	2.00	z=.577, *p*=.564
How demanding	M	2.0	2.0	1 rated higher in car
was it to perform	StD	0.64	0.54	2 rated lower in car
the task?				5 rated same
Stress	m	1.57	1.86	z=1.00, *p*=.317
How pressed	M	2.0	2.0	3 rated higher in car
(stressed) did you	StD	0.54	0.69	1 rated lower in car
feel while doing				3 rated same
the task?				

m: mean, M: median, Little - Much (scale 1-3)

any previous experience of automated driving and the NDRT. The participants' first impression of the vehicle consisted of expressions such as: "cool" [TP102], "nice to be able to do other stuff" [TP103], "fun, I want to buy" [TP105], "impressive" [TP106], and "exiting" [TP107, TP108, TP109]. Three of the participants were less enthusiastic and spontaneous highlighted that limitations in the vehicle's driving style affected their experience [TP101, TP104, TP110].

The thematic analysis from the first interview showed that the participants could easily engage in the NDRT [TP101, TP202, TP105, TP107, TP108]. They also enjoyed engaging in the task [TP101, TP102, TP103, TP105, TP108, TP110], or at least thought the experience was rather good [TP106, TP104, TP107, TP109]. As some of them explained: "it feels good" [TP203], "I could enjoy the ride" [TP201], "it was good - you forced me to get disconnected" [TP107].

One contributing factor for enjoying the task was that it made them "happily unaware of the situation" as they did not notice things that they may otherwise wonder about [TP101, TP102, TP107, TP108, TP109, TP207, TP208, TP209]. As stated by some participants: "I was in a bubble, I did not

look at what the car did" [TP101], "I thought I would experience it [performing the task] as more unsafe, but I experienced it as more safe as I did not notice situations I otherwise would wonder about" [TP109], "It is rather nice to do a task while traveling, I embraced the task, and can enjoy that the car is driving" [TP201].

However, a few participants expressed that they felt "forced" to do the task, indicating that it may not come completely natural for all participants to fully rely on the vehicle [TP101, TP102, TP103, TP107]. Indeed, two participants [TP110/210, TP104/204] found this as most challenging, although they expressed that it was somewhat easier the second time. As one participant reflected: "at the beginning, I had an idea of just pressing random buttons and focusing on driving, but then I thought it would be cheating so I tried to do the task" [TP110].

In line with this, several participants were surprised by their own ability to rely on the vehicle and engage in the NDRT [TP101, TP102/202, TP103/203, TP104, TP106, TP107, TP109]. To exemplify: "I was pleasantly surprised by my ability to let go, I could actually complete the task " [TP203], "Last time one was more surprised that one let go of driving" [TP202], "What I reacted to was that you really let go of the driving when you were provided with a task" [TP102], "all of a sudden, I did not care about the vehicle and I completely trusted it" [TP101], "It is easy to become embraced with the task and one does not think of the surroundings so much" [TP202], "I could focus on the task and I could zoom out from everything else" [TP207].

From the interviews, it is possible to distinguish four levels of engagement ranging from fully engaged to minimally engaged (Table 6), reflecting different level of disconnection from the driving (role). It seems that participants that experienced "fully engaged" or "minimally engaged" did not move between categories between occasion one and two (Table 6). However, the participants that experienced "partly engaged", changed category on the second occasion, indicating that they could become more engaged in the task the second time. This was also confirmed by their own self reflections [TP203, TP204, TP206, TP205, TP208]. For instance: "Overall, I would say that it is a better experience than the previous time" [TP204], "Last time, there was one occasion when the car slowed down and then I looked up to make sure everything worked, this time I did not need to do this" [TP206], "Last time, I noticed a small bridge along the road, but I didn't notice it today at all" [TP205]. This indicates that participants belonging to the more extreme categories had a more stable experience over time. It is possible that participants belonging to the "minimally engaged" category can adapt but require more time to do so.

It was noted that sudden changes in vehicle speed and in the surrounding environment affected the task engagement [TP101, TP102, TP103, TP105, TP109]: "It is just those physical movements generating a feeling in your body that disconnect you from the iPad, you look up, and twitch"

TABLE 6. Participants' ability to engage in the NDRT. while in unsupervised automated driving mode.

Level of engagement	Description	Citations
Fully engaged	One can mentally disconnect from driving with ease, and one is not affected by external (inside/outside vehicle) and/or internal (mental) factors	Occasion 1: "I could completely disconnect" [TP101]; I could let go of driving" [TP102]; "I could let go of driving completely" [TP105]; "It is a bit strange to do other stuff while driving, but it works " [TP107] Occasion 2: "I get fulfilled by the task and enjoy that the vehicle can drive by itself" [TP201]; "It felt good, you get quite into the task" [TP202]; "I could trust the car to the level that I could do the task" [TP203]; "There is a small bridge that I noticed last time and today I did not notice it" [TP205]; "The need to drive was not there anymore, I was fully focused on the task" [TP206]"; "I could focus on the task and I could zoom out from everything else" [TP207]
Almost fully engaged	One can mentally disconnect from driving, but express a sensitivity to changes in the surrounding traffic environment, i.e. one is easily interrupted by external factors	Occasion 1: "When a car passes by fast, you want to look, and you disconnect from the task" [TP109] Occasion 2: "I felt confident and was much more focused on the computer and did not look at the road" [TP204]; "I think I looked up two or three times, but it worked well" [TP208]; "I reacted to the surrounding environment, but I did not feel stressed" [TP209]
Partly engaged	One cannot fully mentally disconnect from driving as one is ready to act, i.e. one actively direct attention to the driving task as well as the NDRT	Occasion 1: "You need to look up sometimes" [TP103]; "I tried to answer, but at the same time be ready to take over" [TP104]; "I felt that I could both look at the road and to the test" [TP106]; " I looked in the rear-view window"[TP108]
Minimally engaged	One cannot mentally disconnect from driving, and one observes the surrounding traffic environment continuously	Occasion 1: "At the beginning, I had an idea of just press random buttons and focus on driving, but then I thought it would be cheating so I tried to do the task, while monitoring the traffic, and it was difficult" [TP110] Occasion 2: "I wanted to do it as quickly as possible to be able to look at the road" [TP210]

The first number in the three-digit sequence within brackets represents the test occasion (1 or 2) and the following two numbers represents the test participant.

[TP201]. It seems like a reflexive reaction beyond human self-control: "It is a reflex" [TP103]. These interruptions were experienced negatively by some participants as they thought it took time from the task [TP201, TP109].

Also, the actual location affected the participant's task experience. They indicated that they felt more stressed while doing the task in the vehicle as compared to the office: "When I needed to do the task, I felt more stressed than regularly" [TP104], "It is a more stressful situation in the vehicle" [TP203], "You feel more safe here [office]" [TP202]. The participants felt that it was extra demanding to perform the task in the vehicle: "I was feeling ready to act" [TP204], "It demanded a bit more" [TP109], "You do not want to drive off the road, and you do not have that feeling in the office" [TP103]. In general, participants experienced the office was a calmer environment, e.g. "the environment was more calm and more quiet, tranquil" [TP105], "It is quieter in the office" [TP205], "When I did the test in the office, I could fully focus on it and I have no insecure feeling of making any accident, stress level is much lower" [TP204]. One participant also expressed a preference for the car: "It is a bit more performance anxiety here [office], it is more a competition" [TP206].

As the participants reflected upon how they executed the task a set of themes/categories emerged that can indicate how they dealt with the fact that the NDRT was performed while in the unsupervised automated driving mode (Table 7). These (behavior) strategies ranged from being ready to act, to consciously looking up every now and then, to looking up on demand when something changed in the surroundings, for example when the vehicle slowed down.

Some participants purposely divided their attention between driving and the task (dual activity), while others moved into a mindset of being a passenger, for some, a forced activity (Table 6). At the end, participants reflected upon the ability to perform the NDRT in the vehicle. They commonly expressed that they saw a usage for such a function during motorway driving [TP109, TP103, TP106, TP107], for long distance travels [TP109, TP03, TP106, TP107], on straight roads [TP103, TP107], and when there is enough time available to activate it: "if it is available for at least 5 minutes" [TP105]. As one participant explained: "I would use it on the road sections where I today use the ACC" [TP108]. Another participant reflected: "I could do stuff you would do on a train, for example read a book, look at a movie, rest, eat" [TP207].

However, the participants also identified an additional area of usage, namely when there is a (sudden) need to hand over the control for a short amount on time when focus is elsewhere, for example "when making a call in a city environment" [TP101], or "when you look for something such as your child's pacifier" [TP102]. During the second occasion

TABLE 7. Participants' strategies when performing the NDRT while in unsupervised automated driving mode.

Category	Description	Citations
Individual behavioral		
Looking up	Either (1) Occasionally looking up to increase awareness of the surrounding environment and vehicle behavior (2) When needed shift focus from task, trust senses to alert	(1) "I looked up and then back to the task" [TP102]; "I checked and looked up once in a while, I wanted some control" [TP103]; "I looked up to be ready to take action" [TP104]; "I needed to look up" [TP106]; "I looked up" [TP107]; "I look up … it is nice to be in control" [TP108]; "It is easy to check the mirrors then you are prepared" [TP108] (2) "As soon as I heard something, a approaching car, the vehicle breaking, I would have reacted to see what happens … but I miss everything that you do not physically feel with gas and break," [TP109]; "When something changes you look up" [TP102, TP106, TP108]
Body state	Consciously being in a state to be ready to act if needed	"I was thinking how I can quickly grab the steering wheel" [TP104]; "I was always ready to react to changes in the environment" [TP109]
Individual mental state		
Dual activity	Purposely conscious division of focus between driving and NDRT	"While doing the task I felt that I, at the same time, was analyzing what the car was doing" [TP101]; "be ready to take action … I tried to answer and at the same time be ready to take over I always had the task and car in my head [TP103]; "It was a shared focus [TP104]; "I could do two things … the road and do the task" [TP106]; "I looked up frequently to check … always ready to react to changes in the environment" [TP109]; "I tried to do the task, and I tried to focus on the traffic" [TP110];
Becoming a passenger	Active choice of distancing oneself from the driving task, simply letting go of the driving	"By doing a task I become a passenger… I embraced the task, and enjoyed the drive" [TP201]; "I think it was my mindset ... it was more that I accepted it" [TP101]; "… I just let go when I got the task" [TP102]; "I immersed so much in the task that I just let go of driving" [TP103]; "I could let go of driving completely" [TP105]; "Now I shall trust the vehicle and shall do the task" [TP203]; "I could focus and I felt I trusted the car, I did not need to look up and check" [TP207]; "The driving was completely gone, I focused completely on the task" [TP206]; "During the test I was able to disconnect [TP107]; "I was not responsible for the safety" [TP108]; "It is all about determining to trust the vehicle" [TP109]; "It is easy to become embraced with the task and not think of the surroundings so much" [TP202]; "I could focus on the task and zoom out of everything else" [TP207]
Forced engagement	Conscious effort to force oneself into the task demanding mental effort	"I tried to let go … and fully trust the vehicle" [TP102]; "I needed to make an effort "[TP103]; "I should be able to handle it" [TP106]; "After half of the task I realized that I need to complete the task as the time is running out" [TP103]; "You forced me to do the task" [TP107]; "If you force yourself to perform a task and focus on it, you cannot be afraid at the same time - you trick your brain" [TP108]
Environment		
Placement of artefact	Maximize possibilities by adjusting positions	"I placed the iPad there, and had it quite high up" [TP103]; "It is a bit more difficult, the steering wheel is in the way … change the position of the seat" [TP205]
Task		
Quality management	Being not as meticulous in the vehicle as in the office (reading task instruction, and own performance)	"I was not as careful in the vehicle" [TP101]; "I did not dare to be that careful while doing the task … and to look so closely [at object]" [TP109]; "I think I made more mistakes in the car than in the office" [TP110]; "At the beginning I had an idea of just pressing random buttons and focusing on driving but then I thought, that would be cheating" [TP110]

The first number in the three-digit sequence within brackets represents the test occasion (1 or 2) and the following two numbers represents the test participant.

some concern was raised regarding the physical constraints and the restrictions that may be caused by being seated in a moving vehicle: "I like to work with the computer and the steering wheel will be in the way" [TP205], "you have noise, sun glare and similar, so it is not an optimal working environment" [TP205].

IV. DISCUSSION

The focus of this study was to examine to what extent drivers could engage in an NDRT that was cognitively, visually and motorically demanding. To assess the validity of the NDRT as a cognitively demanding task (i.e. a task that would require a change in focus from driving to the NDRT, enabling the driver to change primary task), participants were asked to rate the NDRT in terms of perceived demand, perceived importance, and task difficulty, as well as according to the SWAT (mental demand, time, physiological stress) as in line with [21]. The task was judged to be of high importance, demanded much effort, and the task was somewhat difficult to perform (Table 3). The participants could concentrate and become absorbed by the task (Table 4). The task also received a moderate score on the mental workload scale (Table 5). In addition, the results show that the task was visually engaging as all participants looked away from the forward path during the vast majority of the task execution (Figure 5). Participants had extremely few, or even no glances on the road while doing the NDRT. The results from the task performance (e.g. number of errors, time to completion, number of completed frames) show that all participants could complete the task above chance (i.e. they made conscious choices focused

on the NDRT) (Figure 8). The number of completed frames also indicates the requirement of a high motoric demand to complete the task. All of this indicates the validity of the task design, fulfilling the goal of being a task that represents a "mind off/eyes off/hands off" situation. The following section discusses the result in more detail as according to the research questions in this paper (Figure 1).

A. HOW WELL ARE DRIVERS ABLE TO SHIFT ATTENTION TO THE NDRT?

The glance data clearly shows that the drivers were able to let go of the task of driving and engage in the NDRT. Most people had extremely few, or even no, glances at the road while doing the NDRT. No one looked at the road more than 20% of the time, which can be compared to a typical on-path glance proportion (commonly referred to as Percent Road Center, PRC) of 70-90% during both manual driving and driving with partial automation [15], [34], [35] and down to 30% when doing visually demanding secondary tasks during manual driving [15], [36]. Most on-path glances were also very short, with 95% of the glances having a duration of less than one second and 50% of the glances having a duration of less than half a second. It appears highly unlikely that so few and short on-path glances could provide even close to a sufficient scene awareness for safe manual driving [37], strongly indicating that the drivers disengaged from the task of driving.

Yet another strong indication that the drivers disengaged from driving was the length of the off-path glances. More than 85% of the off-path glances had a duration above two seconds, and 60% had a duration above four seconds. During manual driving, drivers rarely look away from the road for more than two seconds. In a reference model based on naturalistic manual driving [34], only 3.1% of the off-path glances have a duration above two seconds, and as little as 0.1% have a duration above four seconds [38]. [35] found an increase in off-path glance durations when drivers of Tesla Model S had the Enhanced Autopilot function (containing Traffic Aware Cruise Control, Autosteer, and Auto Lane Change) active compared to when not, but it was still very rare with single off-path glances of a duration above four seconds.

B. HOW WELL ARE DRIVERS ABLE TO PERFORM THE NDRT?

The NDRT performance data (e.g. number of correct answers) show that the drivers were able to complete the NDRT and to perform it equally well in the car as in the office environment. No major significant differences were identified. Although, a slight learning effect was noted during the first occasion in the office indicating that the task practice time should have been a bit longer before deployment. The quick adaptation and ability to perform NDRTs are in line with previous studies such as [12], although studies are still few. More specifically, [39] found that drivers were relaxed after about 10 minutes of riding in an automated vehicle in traffic and that nearly all participants occasionally engaged

in a voluntary NDRT. In a test track study [23] it was found that drivers engaged in an NDRT after about seven minutes, with a shorter time to engagement at the second occasion. It has also been found that the choice to engage in NDRTs varied between 0 and 93% of the driving time, pointing out large differences between individuals [40].

C. HOW DO DRIVERS EXPERIENCE ENGAGING IN THE NDRT?

The task engagement ratings indicate that the participants could concentrate and become absorbed in the task, even though somewhat easier in the office than in the car. However, the differences were small (one increment on the 10-grade scale) and not statistically significant, leading to the overall conclusion that drivers can become roughly equally engaged in an NDRT while in unsupervised automated driving mode as in an office environment. Interviews confirm this by identifying that most participants enjoyed performing the task and could see a value for such a function.

When analyzing the data over time (i.e. between occasion 1 and 2), there was a notable change indicating that it could be easier for drivers to concentrate on an NDRT with time, as they become habituated to automated driving. However, the difference is small. Interestingly, previous studies have indicated that the change is larger within the first occasion compared to between the first and second occasion (cf. [23]).

Considering the interview data, there is a wide range of levels of NDRT engagement, ranging from minimally engaged to fully engaged. It may reflect the selection of participants in this study, ranging from novice to expert with a variety of technical readiness. This is also in line with other studies that have pointed out that the variety of user experience while traveling in supervised automated driving mode most probably contributed to the results, e.g., [40]. In future studies it would be interesting to see if personality correlates with level of engagement to anticipate any future usage.

What is also notable is that, as in line with other research on user experience and sociotechnical systems, the mental state of the participants (ease, stress etc.) and their behavior was shaped by the environment of which the NDRT took place. Highlighted factors in the interviews include: the change in environment from a stationary to moving environment, from a spacious to a physically restricted environment, and from a non-critical situation to a safety critical situation. Also, both environments can be subject to sudden unpredictable interruptions (but of different character (passing vehicle vs. passing human in activity-based office), as noted by the participants in the study.

D. OVERALL: HOW WELL DO DRIVERS ENGAGE IN AN NDRT?

So far, NDRTs have typically been studied in relation to safety issues, for example how NDRTs affect take over requests, with limited studies of the performance of NDRTs as the primary investigation. Hence, the need for a purposely designed task for this study. With this study, we show with our

task that participants are indeed able to engage in an NDRT while in unsupervised automated driving mode. Although, interviews point out that (some) participants experienced that they were sensitive to environmental changes and that it was less calm in the car. Even though people said that they could engage in the NDRT (according to their rating on fluency and absorption) and that they performed it equally well in the office and the car, they still reported that they reacted to changes in the vehicle/traffic situation. Most people said that it was difficult *not* to react to changes in vehicle movement (e.g. speed) or other unexpected events in the surrounding environment (e.g., passing of an ambulance). Similar findings that drivers adapt their behavior according to the current traffic environment (demand) has been seen both for manual driving (e.g., [16]–[19] as well as for levels of automation below SAE Level 4 in which the driver still is responsible for driving (e.g., [9], [20]). With this study, similar effects are shown in situations of unsupervised automated driving.

However, it should be noted that this experienced unintended reaction did not (apparently) affect their performance (e.g. completed frames and number of correct answers), as there was no significant difference between task performance in the office and vehicle. The ability to ''just look up'' was in a way considered as a positive aspect, providing comfort, and safety. Some participants experienced a difference in doing the task in the office compared to in the car though; the office being calmer and more tranquil, while an extra demand was experienced in the car.

What implications this has for the changing role of the driver is up for debate. It is yet to determine the distribution between self-paced looking up (i.e. when having a pause in the task) and environmental induced disturbances and its effects. As one participant reflected ''you react if you are a passenger in a normal vehicle too'' [TP209]. Indeed, only two participants experienced that ''looking up'' influenced their overall performance and their ability to complete the task. In fact, the participants developed different approaches which could be viewed as mitigation strategies to maintain task engagement and become comfortable in performing the NDRT while in unsupervised automated driving mode (Table 6). Indeed, those strategies could help explain that all participants reported that they could engage, and was mostly comfortable doing it, many even with surprise and ease.

E. THE VALUE OF AUTOMATED DRIVING: A PLATFORM FOR ACTIVITIES?

For automated vehicles to fulfil its promises of being a platform that can ''free up time'' and increase productivity [2], [4], [41] there is an assumption that the driver will be able to use the time in automated driving for other activities than mere travelling. The goal of this research was therefore to explore whether automated driving can be a platform in which drivers can engage in visually, cognitively and manually demanding activities. The results indicate that people were able to shift their attention to the NDRT and perform it equally well as in an office environment, and that they

experienced that they could engage in the NDRT. There are thus indications that there is a possibility for automated vehicles to fulfil its promises. Indeed, most participants reported enjoying performing the NDRT in the vehicle, empirically confirming previous attitude surveys on adaptability of automated vehicles [5]–[7]. Participants experienced similar flow and workload, irrespective of location - car or office. Interestingly, participants enjoyed the feeling of being ''happily unaware'', as several participants expressed that it was good that they performed a task that limited their ability to pay attention to the vehicle and surrounding traffic environment. The task itself aided their ability to move away from the driving role, allowing them to become a passenger to a greater extent than when they were given no task at all. Future studies should explore how NDRTs mediate the role of the driver and how NDRTs could be encouraged and explored as a mean to aid drivers to let go of driving.

There is also an ongoing discussion about in what situations and under what conditions unsupervised automated driving mode would be used. The results from the interviews confirm the ongoing technical development, emphasizing the usage for, for example, motorway and long-distance driving. However, this study also highlights that there is another usage: sudden, immediate support due to changes in the social/physical environment in the vehicle (phone call, fetching pacifiers, etc.), indicating a wider usage scope besides longer motorway passages. This is a finding less emphasized in current surveys and attitude studies on the role of automated driving. Indeed, unsupervised automated driving mode is not typically portrayed as a driver support system for such short time intervals.

There is also a question of what type of NDRTs are most suitable to perform while in an automated driving mode. This study identifies that some participants experienced a reflex to look up due to changes in the environment (e.g. an ambulance passes), as in line with studies on secondary task performance [18] and lower level of automation [9], [20], [21]. This indicates that some types of activities may be more suitable to perform in vehicles than others, namely activities that can manage small unpredictable interruptions (cf. [42]). An aspect that needs to be considered in the future design and marketing of automated driving.

F. METHODOLOGICAL IMPLICATIONS

Wizard of Oz (WOz) studies in real life have been proven successful in its abilities to study technology of the future such as automated vehicles and can be a great tool to learn about driver behavior and subjective experience [12], [22], [23], [43]. However, it is not without its limitations as it is difficult to mimic automation/position of the car repeatedly for a human being, even when driving assistance is used to a great extent. In this study, approaches such as formal training for the WOz drivers, the use of landmarks and limiting the number of drivers involved were used to minimize variations within the study. Measures were also made to keep the number of people in the vehicle to a minimum and letting the test

participant be alone in the front seat. The fact that participants were monitored, both by test leaders in the vehicle, as well as by video and physiological measurement recordings, can still have affected their behavior as well as their experiences though.

When performing studies of future technology there is a risk of participants having a first encounter effect which affects their subjective experience. In this study, participants experienced the automated vehicle twice, with a week in between, to reduce this effect. While the drivers' task performance did not change between occasions, the subjectively reported experience and for some participants the glance behavior, changed, by becoming more relaxed. The results also showed a learning effect for the NDRT, where the first occasion performance (which was in the office environment) was worse than all other occasions. Optimally, the drivers should have gotten more training on the NDRT, both in the office environment and in the moving vehicle, to reduce or eliminate these effects.

It should be noted that the results of the study do not account for other levels of automation (e.g. with less/more driver responsibilities), other traffic environments (e.g. city traffic), or other types of tasks (e.g. self-paced or more difficult tasks) as it might yield different behaviors and experiences.

There is not yet a standard approach on how to evaluate and measure driver engagement in terms of ability to shift focus between primary and secondary/tertiary tasks during automated driving as evident by the reviewed existing NDRTs. The difference in results between the different data collection methods of this study highlights the complexity of the subject and the importance of using different data collection methods to capture level of engagement. Data should be viewed as complementary to indicate level of engagement (cf. [44]). The presented research has taken inspiration from previous studies that explored mental workload and flow theory [21], while eyes on/off road has been used as an indication of disconnection from driving. The presented research thus contributes to the ongoing development of evaluation procedures for automated vehicles in terms overall experimental set up and choice of NDRT.

Future work includes mixed method analyses of the data collected in this experiment to explore driver attention, experiences and responses (physiological and behavioral) to the different levels of automation. For example, visual attention will be explored through gaze direction, eye blinks and eye fixation related potentials (EEG).

V. CONCLUSION

This study empirically explored the potential of automated vehicles being a platform for performing cognitively, visually and motorically demanding activities. It was done by examining to what extent drivers can engage in an NDRT, and was assessed in terms of visual behavior, task performance and subjective experience. The visual behavior analysis showed that the participants' attention was directed

towards the NDRT, the task performance analysis showed equally good performance in the vehicle as in the office, and the participants reported feeling engaged in the NDRT to different degrees. Although it should be noted that a majority of the participants experienced that they ''looked up'' occasionally or on demand due to environmental/vehicle changes. However, this did not affect their task performance, and was by some, experienced as a positive. The result of the study thereby shows that automated vehicles have the potential of being a platform for NDRTs. Since this study is limited in terms of having only 10 people participating at two occasions, future studies are needed to confirm the results.

REFERENCES

[1] G. Meixner, C. Häcker, B. Decker, S. Gerlach, A. Hess, K. Holl, A. Klaus, D. Lüddecke, D. Mauser, M. Orfgen, M. Poguntke, N. Walter, and R. Zhang, "Retrospective and future automotive infotainment systems—100 years of user interface evolution," in *Automotive User Interfaces*. Cham, Switzerland: Springer, 2017, pp. 3–53.

[2] S. M. Casner, E. L. Hutchins, and D. Norman, "The challenges of partially automated driving," *Commun. ACM*, vol. 59, no. 5, pp. 70–77, Apr. 2016.

[3] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles-SAE International*, Standard J3016B. Accessed: Jul. 2, 2020. [Online]. Available: https://www.sae.org/standards/content/j3016_201806/

[4] L. L. Chuang, S. F. Donker, A. L. Kun, and C. P. Janssen, "Workshop on the mobile office," in *Proc. 10th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, 2018, pp. 10–16.

[5] T. Hecht, E. Darlagiannis, and K. Bengler, "Non-driving related activities in automated driving–an online survey investigating user needs," in *Proc. Int. Conf. Hum. Syst. Eng. Design, Future Trends Appl.*, 2019, pp. 182–188.

[6] B. Pfleging, M. Rang, and N. Broy, "Investigating user needs for non-driving-related activities during automated driving," in *Proc. 15th Int. Conf. Mobile Ubiquitous Multimedia*, 2016, pp. 91–99.

[7] Capgemini. (2019). *The Autonomous Car: A Consumer Perspective*. [Online]. Available: https://www.capgemini.com/wp-content/uploads/2019/05/30min-%E2%80%93-Report.pdf

[8] J. C. F. de Winter, R. Happee, M. H. Martens, and N. A. Stanton, "Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 27, pp. 196–217, Nov. 2014.

[9] B. Wandtner, N. Schömig, and G. Schmidt, "Secondary task engagement and disengagement in the context of highly automated driving," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 58, pp. 253–263, Oct. 2018, doi: 10.1016/j.trf.2018.06.001.

[10] O. Carsten, F. C. Lai, Y. Barnard, A. H. Jamson, and N. Merat, "Control task substitution in semiautomated driving: Does it matter what aspects are automated?" *Hum. Factors*, vol. 54, no. 5, pp. 747–761, Jan. 2012.

[11] D. R. Large, G. Burnett, D. Salanitri, A. Lawson, and E. Box, "A longitudinal simulator study to explore drivers' behaviour in level 3 automated vehicles," in *Proc. 11th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, Utrecht, The Netherlands, Sep. 2019, pp. 222–232, doi: 10.1145/3342197.3344519.

[12] R. Broström, A. Rydström, and C. Kopp, "Drivers quickly trust autonomous cars," in *Advances in Usability, User Experience and Assistive Technology*. Cham, Switzerland: Springer, 2018, pp. 699–705.

[13] F. Naujoks, D. Befelein, K. Wiedemann, and A. Neukum, "A review of non-driving-related tasks used in studies on automated driving," in *Proc. Int. Conf. Appl. Hum. Factors Ergonom.*, 2017, pp. 525–537.

[14] Y. Forster, A.-K. Frison, P. Wintersberger, V. Geisel, S. Hergeth, and A. Riener, "Where we come from and where we are going: A review of automated driving studies," in *Proc. 11th Int. Conf. Automot. User Interfaces Interact. Veh. Appl., Adjunct Proc.*, 2019, pp. 140–145.

[15] T. W. Victor, J. L. Harbluk, and J. A. Engström, "Sensitivity of eye-movement measures to in-vehicle task difficulty," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 8, no. 2, pp. 167–190, Mar. 2005, doi: 10.1016/j.trf.2005.04.014.

[16] E. Tivesten and M. Dozza, "Driving context influences drivers' decision to engage in visual–manual phone tasks: Evidence from a naturalistic driving study," *J. Saf. Res.*, vol. 53, pp. 87–96, Jun. 2015.

[17] B. Wandtner, M. Schumacher, and E. A. Schmidt, "The role of self-regulation in the context of driver distraction: A simulator study," *Traffic Injury Prevention*, vol. 17, no. 5, pp. 472–479, Jul. 2016, doi: 10.1080/15389588.2015.1102231.

[18] N. Schömig, B. Metz, and H.-P. Krüger, "Anticipatory and control processes in the interaction with secondary tasks while driving," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 14, no. 6, pp. 525–538, Nov. 2011, doi: 10.1016/j.trf.2011.06.006.

[19] R. Ismaeel, D. Hibberd, O. Carsten, and S. Jamson, "Do drivers self-regulate their engagement in secondary tasks at intersections? An examination based on naturalistic driving data," *Accident Anal. Prevention*, vol. 137, Mar. 2020, Art. no. 105464, doi: 10.1016/j.aap.2020.105464.

[20] F. Naujoks, C. Purucker, and A. Neukum, "Secondary task engagement and vehicle automation–comparing the effects of different automation levels in an on-road experiment," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 38, pp. 67–82, Apr. 2016.

[21] S. M. Ko and Y. G. Ji, "How we can measure the non-driving-task engagement in automated driving: Comparing flow experience and workload," *Appl. Ergonom.*, vol. 67, pp. 237–245, Feb. 2018.

[22] A. Habibovic, J. Andersson, M. Nilsson, V. M. Lundgren, and J. Nilsson, "Evaluating interactions with non-existing automated vehicles: Three wizard of Oz approaches," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 32–37.

[23] J. Andersson, E. Asker, A. Habibovic, M. Klingegard, D. Lindström, and V. M. Lundgren, "Method development for studies at test tracks of long-term behavioral adaption in automated vehicles," SAFER Open Res. AstaZero, Göteborg, Sweden, Tech. Rep. A-0018, 2018. [Online]. Available: https://www.saferresearch.com/sites/safer.cloud.chalmers.se/files/2019-11/FinalReport_A-0018.pdf

[24] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, no. 3972, pp. 701–703, Feb. 1971.

[25] L. A. Cooper, "Mental rotation of random two-dimensional shapes," *Cognit. Psychol.*, vol. 7, no. 1, pp. 20–43, Jan. 1975.

[26] G. Stoet, "PsyToolkit: A software package for programming psychological experiments using linux," *Behav. Res. Methods*, vol. 42, no. 4, pp. 1096–1104, Nov. 2010.

[27] G. Stoet, "PsyToolkit: A novel Web-based method for running online questionnaires and reaction-time experiments," *Teach. Psychol.*, vol. 44, no. 1, pp. 24–31, Jan. 2017.

[28] J. Östlund, L. Nilsson, O. Carsten, N. Merat, H. Jamson, S. Jamson, S. Mouta, J. Carvalhais, J Santos, V. Anttila, H. Sandberg, J. Luoma, D. de Waard, K. Brookhuis, E. Johansson, J. Engström, T. Victor, J. Harbluk, W. Janssen, and R. Brouwer. (2004). *HASTE Human Machine Interface and the Safety of Traffic in Europe—Deliverable 2: HMI and Safety-Related Driver Performance*. [Online]. Available: http://www.its.leeds.ac.uk/projects/haste/Haste%20D2%20v1-3%20small.pdf

[29] S. Engeser and F. Rheinberg, "Flow, performance and moderators of challenge-skill balance," *Motivat. Emotion*, vol. 32, no. 3, pp. 158–172, Sep. 2008.

[30] A. Luximon and R. S. Goonetilleke, "Simplified subjective workload assessment technique," *Ergonomics*, vol. 44, no. 3, pp. 229–243, Feb. 2001.

[31] J. C. Flanagan, "The critical incident technique," *Psychol. Bull.*, vol. 51, no. 4, p. 327, 1954.

[32] *Road Vehicles—Measurement of Driver Visual Behaviour With Respect to Transport Information and Control Systems—Part 1: Definitions and Parameters*, Standard ISO 15007-1:2014. Accessed: Jul. 1, 2020. [Online]. Available: https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/66/56621.html

[33] A. Strauss and J. Corbin, *Basics of Qualitative Research Techniques: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA, USA: Sage publications, 1998.

[34] A. Morando, T. Victor, and M. Dozza, "A reference model for driver attention in automation: Glance behavior changes during lateral and longitudinal assistance," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 8, pp. 2999–3009, Aug. 2019.

[35] J. Gaspar and C. Carney, "The effect of partial automation on driver attention: A naturalistic driving study," *Hum. Factors*, vol. 61, no. 8, pp. 1261–1276, Dec. 2019, doi: 10.1177/0018720819836310.

[36] T. McWilliams, J. Lee, and B. Mehler, "Revisiting radio tuning: A secondary analysis comparing glance behavior across five vehicles," in *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, 2017, pp. 1924–1928.

[37] K. Kircher and C. Ahlström, "Issues related to the driver distraction detection algorithm AttenD," in *Proc. 1st Int. Conf. Driv. Distraction Inattention*, 2009, pp. 1–15.

[38] T. W. Victor, E. Tivesten, P. Gustavsson, J. Johansson, F. Sangberg, and M. Ljung Aust, "Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel," *Hum. Factors*, vol. 60, no. 8, pp. 1095–1116, Dec. 2018, doi: 10.1177/0018720818788164.

[39] K. Omozik, Y. Yang, I. Kuntermann, S. Hergeth, and K. Bengler, "How long does it take to relax? Observation of driver behavior during real-world conditionally automated driving," in *Proc. 10th Int. Driving Symp. Human Factors Driver Assessment, Training Vehicle Des.* Iowa City, IA, USA: Public Policy Center of Iowa, Jun. 2019, pp. 245–251.

[40] C. Marberger, D. Manstetten, and C. Klöppel, "Highly automated driving in the real world—A wizard-of-Oz study on user experience and behavior," *VDI-Berichte*, vol. 2360, no. 2360, pp. 109–119, 2019. [Online]. Available: https://trid.trb.org/view/1680596

[41] J. D. Rupp and A. G. King, "Autonomous driving—A practical roadmap," SAE Tech. Paper 2010-01-2335, 2010.

[42] C. P. Janssen, S. T. Iqbal, A. L. Kun, and S. F. Donker, "Interrupted by my car? Implications of interruption and interleaving research for automated vehicles," *Int. J. Hum.-Comput. Stud.*, vol. 130, pp. 221–233, Oct. 2019.

[43] K. Osz, A. Rydström, V. Fors, S. Pink, and R. Brostrom, "Building collaborative test practices: Design ethnography and WOz in autonomous driving research," *IxD&A*, vol. 37, pp. 12–20, 2018. [Online]. Available: https://www.semanticscholar.org/paper/Building-Collaborative-Test-Practices%3A-Design-and-Osz-Rydström/1711362d7bb3a22b04f23b1fe265912411f5b41e?p2df

[44] P. A. Hancock and G. Matthews, "Workload and performance: Associations, insensitivities, and dissociations," *Hum. Factors*, vol. 61, no. 3, pp. 374–392, May 2019.

• • •