This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Effects of Non-Driving Related Tasks During Self-Driving Mode

Saad Minhas, *Student Member, IEEE*, Aura Hernández-Sabaté, Shoaib Ehsan, *Senior Member, IEEE*, and Klaus D. McDonald-Maier, *Senior Member, IEEE*

***Abstract*— Perception reaction time and mental workload have proven to be crucial in manual driving. Moreover, in highly automated cars, where most of the research is focusing on Level 4 Autonomous driving, take-over performance is also a key factor when taking road safety into account. This study aims to investigate how the immersion in non-driving related tasks affects the take-over performance of drivers in given scenarios. The paper also highlights the use of virtual simulators to gather efficient data that can be crucial in easing the transition between manual and autonomous driving scenarios. The use of Computer Aided Simulations is of absolute importance in this day and age since the automotive industry is rapidly moving towards Autonomous technology. An experiment comprising of 40 subjects was performed to examine the reaction times of driver and the influence of other variables in the success of take-over performance in highly automated driving under different circumstances within a highway virtual environment. The results reflect the relationship between reaction times under different scenarios that the drivers might face under the circumstances stated above as well as the importance of variables such as velocity in the success on regaining car control after automated driving. The implications of the results acquired are important for understanding the criteria needed for designing Human Machine Interfaces specifically aimed towards automated driving conditions. Understanding the need to keep drivers in the loop during automation, whilst allowing drivers to safely engage in other non-driving related tasks is an important research area which can be aided by the proposed study.**

***Index Terms*— Mental work capacity, human–computer interaction, vehicle ergonomics, perception, virtual environment.**

I. Introduction

THE automotive industry is moving at a rapid phase towards autonomous vehicles. This involves the cars to make complex judgments depending upon the road scenarios in hand. Several big-name companies such as Waymo, Tesla, and Uber are pushing immense resources into achieving the most efficient autonomous vehicle technology on the market.

Manuscript received September 21, 2019; revised March 14, 2020 and August 1, 2020; accepted September 8, 2020. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC). The Associate Editor for this article was D. F. Wolf. *(Corresponding author: Saad Minhas.)*

Saad Minhas, Shoaib Ehsan, and Klaus D. McDonald-Maier are with the School of Computer Science and Electronic Engineering, Essex University, Colchester CO4 3SQ, U.K. (e-mail: ssminh@essex.ac.uk; sehsan@essex.ac.uk; kdm@essex.ac.uk).

Aura Hernández-Sabaté is with the Computer Science Department, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain, and also with the Computer Vision Center, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain (e-mail: aura@cvc.uab.cat).

Digital Object Identifier 10.1109/TITS.2020.3025542

The development today reflects that there is still quite a long way to go in achieving fully autonomous driving under all road environments as well as weather conditions [1]. The SAE International standard J3016 [2] defines six levels of car automation, from totally manual (level 0) until totally automated (level 5). Far from being at level 5, most current research efforts focus on level 3 and level 4 (like autonomous minibuses CabiBus [3]). These levels include situations in which the car handles all the dynamic driving aspects except the driver taking back control when required (conditionally automated) or in some driving modes. This, in turn, frees up the driver to indulge in other Non-Driving Related tasks (NDR-tasks) while the car is autonomously driving, although the driver must be able to regain the control as soon as the system requires it. At this point, it is necessary to have a comprehensive knowledge under which circumstances the system should require the driver to regain control of the car, taking into account the system boundary and the situational awareness and capabilities of the driver.

Perception Reaction Time (PRT) and mental workload have proven to be crucial in manual driving [4]–[6]. Meanwhile, in highly automated cars, Take-Over Performance (TOP) is an addition variable to take into account for road safety [7]. In these cases, the mental workload is closely related to immersing the driver in NDR-tasks (s)he is performing while the car is driving autonomously [8]. Previous studies found impairments in take-over performance while engaged in NDR-tasks, but little is known about the impact of specific task characteristics [9], [10].

This study aims to explore how the immersion in NDR-tasks affects the success of TOP of drivers in a highway critical situation and evaluate the influence of several variables, such as PRT, on this success. For that, 40 participants (mean age, 31 years old) were selected to evaluate their feet and hands reaction times under different circumstances in a simulation environment. In particular, they had to avoid a crash after taking back the control of the car that previously was autonomously driving when an alarm was triggered due to the detection of an obstacle on the road in front of the car.

The remaining article is structured as follows. Section II summarizes the state-of-the-art related with the paper. Section III is devoted the experiment, detailing the data set, the simulation environment, the procedure and the measures used for the study. Finally, section IV shows the results

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

obtained and section V discuss the results and betrays the further lines.

II. Related Work

The state of the art is described for the following three categories.

A. Perception Reaction Time

PRT of human drivers is an active research area within the manual driving performance domain because it plays a central role in different road incidents.

Several studies have been carried out to deepen the comprehension of PRT's role in crash risks [4], [11]–[14]. The methodologies range from proposing accident situations involving surprise factor to examine the reaction times of drivers and also analyzing reaction times as a factor to take into account within crash surrogate indicators.

The main results found in these studies are that the reaction time of drivers seem to be approximately a linear function of Time To Collision (TTC), and the mean reaction time and inter-individual variability progressively increases with age although some other factors such as driver gender, cognitive load, and urgency might influence in human perception-brake reaction time. However, the most influential factor is driver expectation.

All of the above studies were conducted under manual driving conditions, so they do not take into account the PRT when the driver is carrying NDR-tasks.

B. Mental Workload

Other studies have been centred on the influence of mental workload as a crash risk factor [15]–[19]. Mental workload is not only related to being stressed, fatigued or drowsy but performing a divided-attention task causes an increase of mental workload and task demands can exceed the driver's attention resources. These studies explore several indicators from many external sensors, such as pulse rate, skin electric potential activity or surface temperature, to better determine the physical and psychological state under different provoked circumstances. Although the general workload is not well defined psychometrically [20], all of them coincide that excessive (related to stress) or too low (related to vigilance) mental workload could derail the quality of driving [21].

Besides, human performance can either deteriorate or improve depending on the degree of automation which is introduced in that particular environment [22], so that mental workload should also be taken into account in highly automated cars. Indeed, PRT and mental workload can be closely related [23], [24], since increasing workload of the driver reduces the driver's ability to process information at different distances and thus deteriorates driving performances and increases reaction times. The first study shows that mental calculations increase the average reaction time for each age group, while the second one suggests that reaction times can increase by 40%-87% due to increased fatigue levels, giving valuable insight into how reaction times are taken into account via visual perception.

The above studies do appear to be invaluable in assessing the relationship between PRT and mental workload but assessing these variables under a controlled simulated autonomous car environment is quite crucial in exploring their effects further.

C. Control Switching

Last but not least, in highly automated cars, the process of getting the driver back into the loop is very important. In this fashion, some authors [25]–[27] explore different ways to get the driver back to the driving task in a safe manner, either focusing on signal modalities [25] or designing complete human-machine interfaces [26], [27]. But still, the automated system needs to know how far in advance and under which circumstances it has to warn the driver, depending on the NDR-tasks the driver is doing or can do, so that the analysis of drivers' take-over performance is crucial. Concerning the lead time to safely allow the driver to regain control. Eriksson *et al.* [28] review several papers exploring driver control transitions, although they not take into account secondary tasks, and carry on an experiment involving secondary tasks. On one hand, they claim that the reviewed results differ depending on the emergency the driver perceives (s)he has to cover on, and, on the other one, they find significant differences when drivers are engaged in secondary tasks.

Besides, [29] suggests that a take-over request with lead time at 10–60 s led to lower crash rate, greater minimum TTC, and lower lateral acceleration. However, both studies do not account for critical control transitions. Some other experiments [9], [30] exposed drivers to critical take-over situations and showed evidence that cognitive load on its own might not influence takeover time but have effects on the takeover quality. As well, reaction times might be in line with the driver's perception of emergency.

In case of being behind the wheel of an autonomous car such as Tesla S [31], although the drivers were also told that they were responsible for the safe operations of the vehicle regardless of it's driving mode, the recorded data demonstrated behaviour indicative of complacency and over-trust.

Still, prospectively evaluate the expected limitations caused by NDR tasks on the driver's ability to take control of an autonomous vehicle [32], more research is needed so that different aspects of NDR tasks can be translated into a modelling of a framework to predict takeover time or quality. This makes our present article more relevant as we explore in detail the different effects that the NDR tasks have on reaction times of drivers in an autonomous scenario.

III. Experiments

A. Participants

Forty participants (10 female, 30 male) between 19 and 45 years old (*mean* = 30.73, *std* = 7.086) were recruited. All of them held a valid driver's license at the time of the experiment with a seniority of at least 1 year and at most 26 years (*mean* = 9.725, *std* = 7.66). Informed consent was obtained from each participant before the trials.

B. Simulated Environment and Scenarios

All the experiments studying conditions of the drivers and involving crash accidents are dangerous to approach on real

Fig. 1. Driver simulator setup.

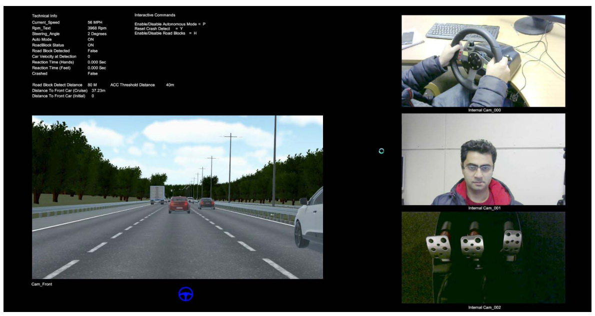


Fig. 2. Simulator recorded data.

roads. Simulators are an effective tool in recognizing and assessing problematic driving scenarios, and hence can provide vital data that can be used to design more efficient Advanced Driver Assistance Systems (ADAS), from predicting driver steering behaviours [33] to evaluating the mental workload of the driver [19], [34].

Our simulated environment is an efficient and inexpensive real looking driving environment that immerses the driver into driving on the highway [35]. The assets required are fairly standard and upgrade-able, which makes the setup quite extendable. The setup consists of one camera facing the driver straight on for capturing the important aspects of driver's facial features, a second camera looking towards the steering and a third one recording precisely the driver's feet placement. Consequently, the three cameras can generate a detailed picture of exactly what the driver is doing at any given instance in time during a trial. Figure 1 shows the general simulator setup.

The system records several variables (reaction times of hands and feet, distance to the front car at the alarm triggering and after a successful stop, and cruise distance of the car in automated mode) that are visualized on a primary screen as shown in Figure 2.

The software side consists of Unity3d, which is used for all the programming of the driver simulator, and other third-party softwares, like Autodesk 3dsmax and Adobe Photoshop, are used to generate the required 3d and 2d assets. This includes virtual cars and road surface assets.

The road surface was modelled to reflect as close as possible the London Orbital Motorway. To target a modular approach towards the road environments only two road sections were produced, the first one being a straight section and the second one being a 25-degree curved road section. Figure 3 shows

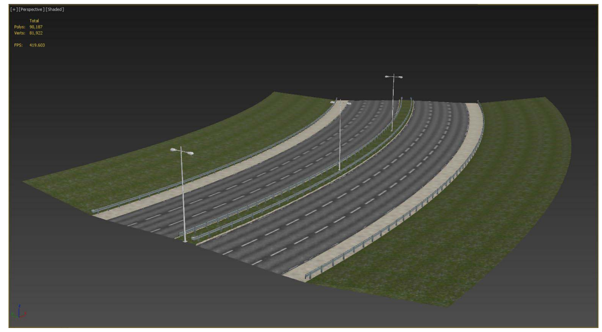


Fig. 3. Road surface model.

Fig. 4. Traffic cars.

Fig. 5. Simulator interior model.

the curved road section. These two sections were then used in such a pattern that would result in a continuous looping road environment. By doing a looping road environment we achieved a never-ending motorway section that could be used for as long as an experiment is required. The road section environment is roughly 4 miles long and proves to be a base for experimenting in the virtual world. Besides the road, basic tree models were then placed to simulate nature.

The traffic cars as shown in Figure 4 are populated via a third-party plug-in called ITS (Intelligent Traffic System). This enables the cars to spawn at a random location every time a new trial is executed. The cars can overtake other cars and their speeds can be controlled depending on their placement on the road i.e. the cars on the fast lane can have a top speed ranging from 70-80 mph, cars in the middle lane can range from 60-70 mph and cars in the third lane can range from 50-60 mph. Moreover, the traffic cars can keep a safe distance from the cars ahead, and when the front car stops due to an incident the other traffic cars respond accordingly.

Figure 5 shows the virtual interior of the driver car. All the useful elements of the main virtual car were rigged, ranging

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

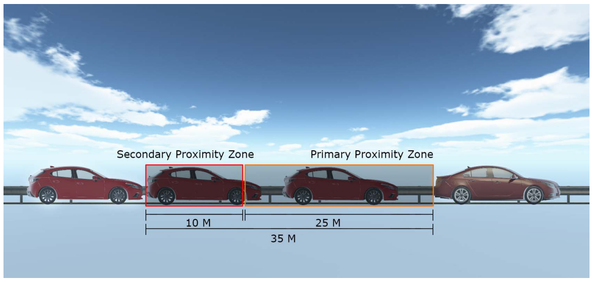


Fig. 6. Simulator proximity model.

from the RPM needle in the dashboard to the brake and accelerator pedals underneath. The setup consists of both a hardware steering wheel as well as a virtual steering wheel. The visibility of the virtual steering wheel behaves like a bridge between the virtual and the real-world counterpart. Apart from providing visual feedback, the virtual steering also acts as a somewhat intrusive figure further simulating the real-world steering position. The input hardware was rigged accordingly to respond to the slightest movement from the driver. The rear-view mirrors were also rigged up to properly simulate the mirror functionality. Besides, a fully functional adaptive cruise control was rigged, which allowed the driver car to keep its lane on the motorway and also allows the car to maintain a specific distance from the car ahead. The Adaptive cruise control is split up into two proximity zones, namely Primary Proximity zone and Secondary Proximity Zone as shown in the figure 6. The Primary Proximity Zone forces the autonomous car to maintain an average speed cruise control resulting in approx 70MPH speed whereas the Secondary Proximity Zone forces the autonomous car to brake and increase the distance to the car in front.

The autonomous driving model uses a basic path follow algorithm with an obstacle avoidance system simulating the Lidar component. As the study was based mostly on the driver's perception there was little to no need for a complex autonomous model for this experiment.

C. Simulated Situation

The simulated situation consisted of an infinite three-lane motorway of 4 miles loop, as explained in the subsection above. The car was driving in autonomous mode and, suddenly, the vehicle detected an invisible obstacle, an alarm was triggered and all the cars in front of it stopped. The detection time for obstacles randomly varies from trial to trial but usually, it happens between 2-5 minutes during a trial. At that moment, all the cars in front stopped and the driver had to take over the control of the car in order to avoid a collision. The average distance and the standard deviation of the main car to the car in front at the time the alarm was triggered was 40.64 ± 13.34 meters and the velocity of the car was 45.2 ± 10.97 mph. The traffic can be turned ON on both sides of the road, but this was not a requirement for the current experimental setup. Moreover, the drivers had the instruction of only braking or dodging if needed, depending on the situation.

To carry out the experiment, a series of plausible scenarios were needed that a driver would find behind the wheel of an autonomous car. Following a previous experiment performed by Eriksson *et al.* [36], in which he used a newspaper reading scenario while the car was in autonomous mode, the challenge was to enable the driver to lose focus on the road ahead while engaging in secondary tasks inside the car. Secondary tasks were selected in such a way that they resembled real-life situations as closely as possible. Hence keeping that in mind, three different scenarios for the driver were posed while the car was in autonomous driving mode. The first scenario, henceforth Default, is termed as the base scenario in which the driver was aware to the road with the hands and feet ready to react (hands on the wheel but without touching them and feet on a marked place very close to the pedals). The second scenario, henceforth Social, deals with the fact that the driver was not paying attention down the road, but was freely immersed in social media activities on her/his smartphone. Since immersion to social media could be total and the system is automatically driving, to suppress the variable off-road glance time, the driver was not allowed to look at the road. In the third scenario, namely Immersive Question and Answers (IMQA), the driver was answering a list of questions via a smartphone. These were relatively basic questions consisting of the driver's name, age, hobbies, etc and more cognitive loaded ones such as some basic mathematical addition and subtraction questions. Strict guidelines were provided to the drivers so that they won't even peek at the road ahead while the car was driving automatically and the driver is indulging in secondary tasks (second and third scenarios), the penalty for which was to start the trial again if the driver looked at the road during specific scenarios.

Each driver performed the experiment four times in each scenario under different circumstances, that is a combination of enabling to dodge or not and a threshold on the initial distance to the car in front during the autonomous driving (the approximate distance that the car would keep from the car in front in autonomous mode). Thus, there are a total of 12 trials that were carried out on each subject, although we group the four trials per driver and scenario as 4 different samples. There are a number of variables that are being recorded: Hands/feet PRT, cruise distance threshold to the front car, distance to the front car at which the alarm is triggered, final distance to the front car after the car comes to a stop, the speed at which the alarm is triggered, any traffic cars to the left and right side of the driver's car, crashed or not, and if the driver tried to dodge and save himself. The experiment also includes other basic variables like the age and gender of the driver and the seniority of his/her license.

D. Experimental Design

After a brief introduction to the driver simulator, the participants were given a trial run with the autonomous mode disabled just so that they could get the feel and sensitivity of the steering wheel as well as the pedals. Afterwards, the autonomous mode was turned on and the participants were told to take back control as soon as the audio/visual alarm was triggered by the simulator. Moreover, the participants were also instructed to place their hands and feet in a neutral position during this time. This gave them the initial confidence in tackling the simulated event and the baseline scenario (Default). For the next scenario, Social Media, the participants were briefed to indulge themselves in social media activities on their smartphones while behind the wheel of the virtual autonomous car. For the Immersive Question and Answers scenario, the participants were only given a short briefing regarding the type of questions that would be asked for this part of the experiment. All the data recorded during the trials was used in an anonymous form.

E. Objective Measures

During the experiment, we recorded several objective variables to be explored under different scenarios.

1) PRT: We understand the PRT as the time that elapses from the instant that the driver recognizes the existence of a hazard on the road to the instant that the driver takes appropriate action. The hazard can result from traffic cars in front of the driver's car suddenly stopping and queuing up, which would alert the driver of a possible hazard. In this case, the hazard recognition is perceived thanks to an acoustic alarm. Since we do not have a tool to measure the time from when the alarm is triggered until the driver perceives it and the time elapsed from when the driver perceives it until (s)he acts, we can only measure the elapsed time from the moment the alarm is triggered until the moment when the driver reacts.
   a) PRT of hands (PRTH): The appropriate action for hands is steering the wheel. The system only records the reaction times of the steering once it receives at least 1 degree of input in either direction from the driver. Notice that the alarm is only triggered during straight road sections, so that normal steering inputs on a curve are absent.
   b) PRT of feet (PRTF): The appropriate action for feet is braking, so that, the system records the reaction time as soon as it detects pressure on the brake pedal.
2) Velocity of the car at the time the alarm is triggered.
3) Distance of the car to the one in front at the time the alarm is triggered.
4) Success of TOP: We consider a success when the car does not crash.

To answer how the immersion in NDR-tasks affects the TOP of drivers we analyze the PRTH and PRTF as soon as the alarm is triggered across different scenarios. As well, we explore the relationship between the velocity and distance at the moment the alarm is triggered.

F. Statistical Analysis

According to the objective variables explained in the above subsection notice that success of TOP can be considered as a dependent variable, while the remaining ones are independent, so that we analyze how such independent variables can influence on the success of TOP. Also, since the scenario can influence on the performance of the driver, PRTH and PRTF can be analyzed across the scenarios to explore if they are affected by the current scenario.

To decide if PRTH and PRT are affected by the scenario, a one-way ANOVA should be computed for each variable. This test is usually used to detect significant differences between the distributions of more than two factors (in this case the different scenarios). That is, its hypothesis test associated considers as null hypothesis H_0 meaning the factor has no effect, and as an alternative that it does. In terms of parameters, the ANOVA test can be written as follows:

$$\begin{cases} H_0: & \mu_1 = \mu_2 = \mu_3 \\ H_1: & \exists\ \mu_i \text{ s.t. } \mu_i \neq \mu_j \text{ for some } j = 1, \ldots, 3 \end{cases}$$

where $\mu_i, i = 1, 2, 3$ corresponds to the mean of the objective variable for each scenario.

A requirement for applying an ANOVA is that data is normally distributed, which can be contrasted by means of a Kolmogorov–Smirnov test. In particular, the Lilliefors test is a normality test based on the Kolmogorov–Smirnov one that compares the empirical distribution of the data with a normal distribution without any expected value and variance of the distribution [37]. In case the data does not follow a normal distribution we can use a non-parametric statistical test, instead of a parametric one, which analyzes differences among group medians instead of means. In particular, since each subject repeats the test for all scenarios in our experimental design, we can consider a repeated measure one way ANOVA, so that we use a Friedman test [38].

To measure the strength of agreement between subjects (effect size) we also compute Kendall's W, defined as $W = \chi^2/N(k-1)$, where χ^2 is the test statistic, N the number of samples (160) and k the number of scenarios (3). The results can be categorized as small, medium and large, which in our case will be $[0, 0.10)$, $[0.10, 0.30)$ and $[0.30, 1]$, respectively.

To analyze the influence of an objective variable on the TOP success in each scenario, we also need to compute a non-parametric test in case the data does not follow a normal distribution. In this case the Wilcoxon rank-sum test [39] is an alternative to the Student's t-test for independent (unpaired) samples and the effect size is computed as $r = \|Z/N^{0.5}\|$, where Z is the Z-statistic, and N is the number of participants. In this case, the results are categorized as small effect = $[0.10, 0.30)$, medium effect = $[0.30, 0.50)$ and large effect = $[0.50, 1]$.

IV. Results

For each variable recorded there is a total of 480 samples (40 participants x 4 trials x 3 scenarios). In all tests, the significance level is 0.05. As well, none of the variables follow a normal distribution because the null hypothesis of the Lilliefors test is rejected with a p-value less than 0.05.

To have a visual idea of the distributions of PRT on each scenario, figure 7 shows the boxplots of the 3 global scenarios (Default, Social and IMQA) for PRT of hands and feet. The first observation is that the hands reaction times are higher as compared to their feet counterpart in all scenarios. Outliers, in this case are reflecting the non-normality of the data. One plausible explanation about this non-normality is that there

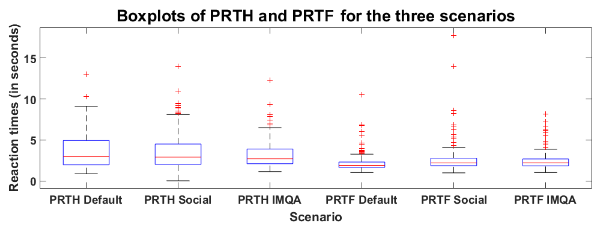


Fig. 7. BoxPlots of PRTH and PRTF among the three scenarios.

TABLE I

MEAN ± STD, MEDIAN (IN SECONDS) AND p-VALUES FOR PRTH AND PRTF

		PRTH		PRTF		p-value
		$\mu \pm \sigma$	median	$\mu \pm \sigma$	median	
Scenario	Default	3.645 ± 2.212	2.99	2.211 ± 1.158	1.90	$7.88e-14$
	Social	3.632 ± 2.326	2.9	2.723 ± 1.967	2.2	$1.81e-11$
	IMQA	3.247 ± 1.716	2.7	2.485 ± 1.139	2.21	$6.75e-11$

were some instances where the drivers failed to input any motion within the steering or the feet resulting in a crash, the reaction times were recorded only after the user performed any input, this caused the system to record higher than normal reaction times, hence the outliers.

The corresponding measures of central tendency are reported in Table I. For each scenario we report the ranges (mean ± std), medians and p-values of the Wilcoxon Signed-rank test, which is the same as Wilcoxon rank-sum test, but for paired data.

The results of the test verify the above visual observation with a global p-value of 6.2954e-33, ensuring the significant difference between PRT for hands and feet. This is due to the fact that drivers, when suddenly encountering an obstacle, tend to prioritize to use the brake pedal before putting any input into the steering wheel hence resulting in the above higher reaction times for hands.

Still, the global average reaction time was 3.51 seconds for hands and 2.47 seconds for feet which coincides with the minimum amount of time described in [40] in which drivers can take over the control of vehicle safely and comfortably in this situation.

The results of Friedman test for PRTH show that there are no significant effects among the 3 scenarios, $\chi^2(2, N = 160) = 3.6904$, $p = 0.1580$, $W = 0.0115$, although the strength of agreement among drivers is very small. That means that drivers take more or less the same time to steer the wheel in both scenarios. However, in the case of feet there are significant differences between at least one of the scenarios, with a medium strength of agreement: $\chi^2(2, N = 160) = 44.3974$, $p = 2.2868e-10$, $W = 0.1387$. To know what scenario is different from the rest, a multiple comparison test has been computed using the output of the Friedman test and shown in Figure 8. Two means are significantly different if their intervals are disjoint, and are not significantly different if their intervals overlap, so that the significant difference in PRTF is on the default scenario against the other two.

These results make sense from the point of view that drivers have to avoid a sudden hazard on the road, so probably the first instinctive action would be braking, taking into account

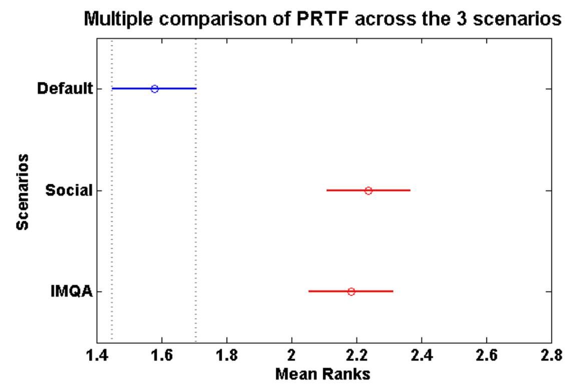


Fig. 8. Multiple comparison test of PRTF among the three scenarios.

TABLE II

PERCENTAGE OF SUCCESSES OF TOP AMONG THE 3 GLOBAL SCENARIOS

	Default	Social	IMQA	global
Failure	18.75	26.25	19.38	21.46
Success	81.25	73.75	80.63	78.54

TABLE III

PERCENTAGE OF SUCCESSES OF TOP BY GENDER AMONG THE 3 GLOBAL SCENARIOS

	Default	Social	IMQA
Female	85	67.5	85
Male	80	75.84	79.17

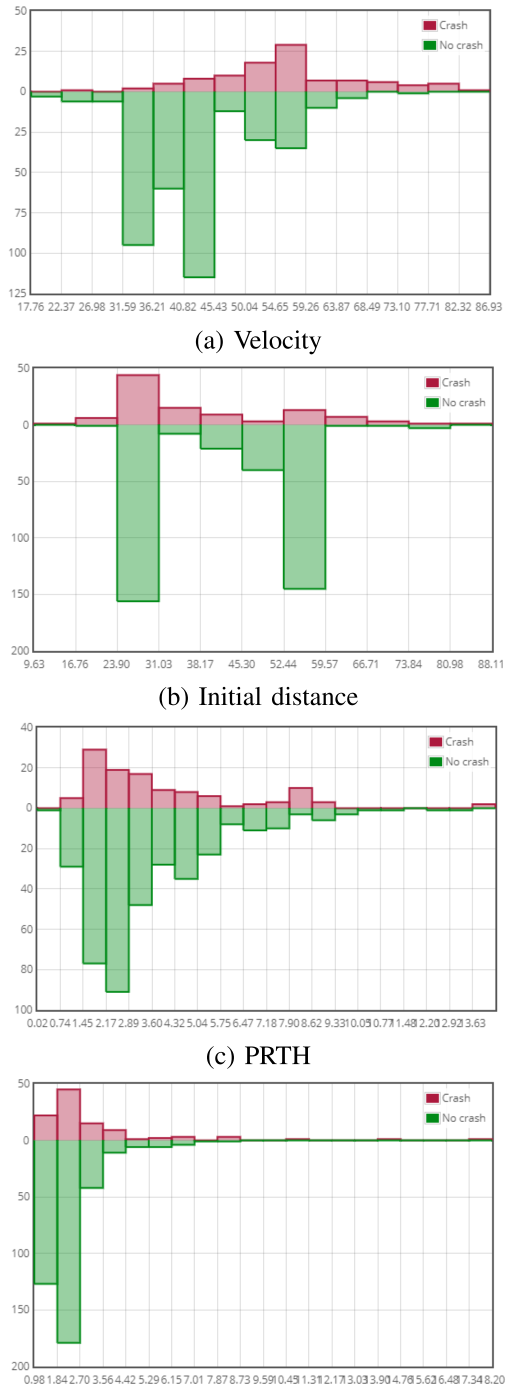
that the alarm is triggered at the same time the cars in the front stop. In this way, as we pointed out before, the reaction time of hands is not relevant in any scenario, but drivers' reaction time of feet might be slower in NDR-tasks scenarios due to their cognitive processing [41], and the driver's lack of attention to the road.

Still, another interesting point is the success of the TOP action by itself, which is reflected in table II. We can appreciate a slight peak of failures in the Social scenario, although it is not significant (Pearson's chi-squared test [42]: $\chi^2 = 3.2881$, $p - value = 0.193$ at significant level of 0.05). There are no significant differences in PRT among Social and IMQA scenarios either, but probably social activities provokes a deepen immersion than questions and answers, so that there are more crashes. If we separate by gender, we can observe that the peak in social scenario remains, as table III suggest. As well, we can notice that females have less crashes than male, although these data could be biased since the number of females is $\frac{1}{3}$ than males.

As well, we can assess the relationship between variables such as velocity, distance or PRTH and PRTF and the success of TOP in each scenario. Global descriptive statistics from figure 9 show that the clearest variable that has significant differences between crashing or not is the velocity. Other variables like Occlusion and aggressive traffic cars had little part to play because the experiment was targeted towards a controlled study of driver's perceptions in a given scenario. That is, the bihistogram of velocity is the most asymmetric one, having most of the no-crash samples lower velocity than most of the crash ones.

This fact is proved by means of the Wilcoxon rank-sum test, summarized in table IV. In all the scenarios we can reject the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) Velocity

(b) Initial distance

(c) PRTH

(d) PRTF

Fig. 9. Bihistograms of the distributions of the different variables about their success of TOP.

null hypothesis so that we have evidence that the medians of the velocity when crashing or not differ.

Table V for the variable of the initial distance of the car to the one in front at the time the alarm is triggered shows that, depending on the scenario, their medians when crashing or not significantly differ, although p-value is very close to the significant level. If we do not take into account the scenario, the p-value obtained rejects the null hypothesis, but we can appreciate that the effect size is small, unlike in the case of velocity.

TABLE IV

WILCOXON RANK-SUM TEST FOR THE RELATIONSHIP BETWEEN VELOCITY AND CRASHING

	Crash	No crash	z-score	r	p-value
Default	55.55	41.01	6.103	0.482	< 0.001
Social	54.88	42.09	6.121	0.484	< 0.001
IMQA	55.16	40.95	5.744	0.454	< 0.001
global	55.34	41.09	10.436	0.476	< 0.001

TABLE V

WILCOXON RANK-SUM TEST FOR THE RELATIONSHIP BETWEEN INITIAL DISTANCE AND CRASHING

	Crash	No crash	z-score	r	p-value
Default	31.5	45.99	−0.560	−0.044	0.576
Social	31.44	48.11	−1.871	−0.148	0.061
IMQA	28.34	47.53	−2.046	−0.162	0.041
global	31.15	47.55	−2.553	−0.117	0.011

TABLE VI

WILCOXON RANK-SUM TEST FOR THE RELATIONSHIP BETWEEN PRTH AND CRASHING

	Crash	No crash	z-score	r	p-value
Default	2.870	2.990	−0.981	−0.078	0.326
Social	3.100	2.820	1.256	0.099	0.209
IMQA	2.660	2.740	−0.052	−0.004	0.959
global	3.100	2.760	1.1017	0.046	0.309

TABLE VII

WILCOXON RANK-SUM TEST FOR THE RELATIONSHIP BETWEEN PRTF AND CRASHING

	Crash	No crash	z-score	r	p-value
Default	2.290	1.900	2.018	0.159	0.044
Social	2.380	2.14	1.943	0.154	0.052
IMQA	2.620	2.100	3.104	0.245	0.002
global	2.420	2.060	4.403	0.201	< 0.001

If we focus on the relationship between PRT and success of TOP we can observe that PRTH maintain the little relevance they already had across the scenarios. Table VI shows that p-values are much greater than the significance level. It makes sense in the light of the foregoing. On the contrary, PRTF seems to impact on the success of TOP, since there are significant differences between crashing or not in most of the scenarios, with a medium effect size. In the case of Social scenario, the null hypothesis can not be rejected. but the p-value is very close to the significance level.

V. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we explore the influence of several variables related to driver-vehicle interactions in an automated driving mode under different scenarios in a simulated environment. In particular, we assess the reaction times of drivers after an unforeseen hazard appears in front while the car is in autonomous driving mode during which the driver is engaged in NDR-tasks. This hazard is alerted by an acoustic alarm and the driver has to regain car control to avoid a crash with the cars stopped in front. This experiment has proven to be a step up from the last version of the driver simulator [35]. Significant modifications were made to get the most amount of usable data from the 40 subjects. As well, the recorded data can then serve in designing complex Advanced Driver Assistance Systems for future Autonomous Vehicles.

One interesting aspect of the recorded results indicate that PRTF of subjects that were involved in NDR-tasks has significantly longer reaction times as compared to subjects who were paying attention to the road. The experiment also seems to show that not all secondary tasks result in higher PRTH. In accordance to that, the results reflect the fact that the reaction times of Social Media usage is higher than the reaction times for IMQA, although these results are not significant. The results also show that the reaction times for subjects are within the 5-second mark, which was noted in other previous research papers [34].

Experiments also shed a light on the fact that the increase in reaction times due to secondary tasks can affect the quality of driving of subjects in any given scenario [43], as well as other variables that can play an important role. In this way, our results sustain the hypothesis that the velocity is one of the most influential variables in the success of TOP as we previously claimed [35].

From the point of view of safety for passengers, we can observe that the most influential variable for safe driving is speed, whether the car drives automatically or the driver does. Besides, drivers have a shorter reaction time with their feet than with their hands when trying to regain control after performing NDR-tasks, so PRTF could be the second most influential variable for safe driving under that conditions. At this point, gender should be another factor to consider, having balanced data to infer.

Since another important variable, which is out of the scope of this article, is the lead time [29], further lines will be focused on the relationship between lead time, velocity and NDR-tasks. Indeed, the implications of these results are important for understanding the criteria needed for the appropriate design of Human Machine Interfaces in automated driving conditions, to ensure that messages regarding the transfer of control are given in a timely and appropriate manner, which is another important area in need of further research [10]. Understanding how to keep drivers in the loop during such automation will allow drivers to safely engage in other NDR-tasks.

Finally, the experiment also forwards the notion that using simulators is a vital part of pushing forward the development of new systems that can provide the drivers of future with necessary assistance in keeping themselves and other road users safe [19]. The resulting data is quite beneficial and it will serve as a dataset for future research like convolutional neural networks development and analysis to test new ADAS algorithms. As well, we have recorded data from 3 webcams monitoring the head, hands and the feet, that can be used to further analyze driver performance with regards to driver psychology and mental workload under autonomous driving conditions. This is particularly important as described by Gaku Iizuka [19], who claims that changes in driver behaviour are observed in all parameters such as blood oxygen levels and gaze perception. Monitoring the cognitive state of the driver could replace several parameters from other convoluted sensors.

This study provides significant insight into how driver's would react when they are behind the wheel of an increasingly autonomous car in a real life scenario. It also reflects the implications of poor driving behaviors which can result in fatalities for the future drivers. Better acoustic alarms, cognitive screening of driver's interactions and better lead time calculations can all result in a more effective and safe autonomous driving systems.

VI. Key Points

- Experiment shows that not all secondary tasks result in higher Perception Reaction Times.
- Experiment shows that the Perception Reaction Times have a global average of 3.51 seconds for hands and 2.47 seconds for feet.
- The study shows that any secondary tasks while driving result in deterioration of quality of driving.
- The implications of the results are important in understanding the criteria needed for designing Human Machine Interfaces for autonomous driving vehicles. This can include entities such as Driver's awareness to his/her surroundings which can be monitored by ADAS resulting in a more enhanced driving experience.
- The study also highlights the fact that the true self-driving cars can be hazardous if they lack the proper systems that can observe the driver's behavior in a timely manner.

References

[1] L. Kelion. (2016) *Tesla Autopilot Involved in a Crash*. [Online]. Available: http://www.bbc.com/news/technology-36783345

[2] SAE, J3016. (2019). *Sae Levels of Driving Automation*. [Online]. Available: https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic

[3] (2019). *Cabibus, the Vehicle*. [Online]. Available: https://www.cabibus.com/cabibus_vehicle.html

[4] Y. Kuang, X. Qu, J. Weng, and A. Etemad-Shahidi, "How does the driver's perception reaction time affect the performances of crash surrogate measures?" *PLoS ONE*, vol. 10, no. 9, Sep. 2015, Art. no. e0138617.

[5] N. Lyu, L. Xie, C. Wu, Q. Fu, and C. Deng, "Driver's cognitive workload and driving performance under traffic sign information exposure in complex environments: A case study of the highways in China," *Int. J. Environ. Res. Public Health*, vol. 14, no. 2, p. 203, Feb. 2017.

[6] C. Dijksterhuis, K. A. Brookhuis, and D. De Waard, "Effects of steering demand on lane keeping behaviour, self-reports, and physiology. A simulator study," *Accident Anal. Prevention*, vol. 43, no. 3, pp. 1074–1081, May 2011.

[7] C. Gold, D. Damböck, L. Lorenz, and K. Bengler, "'Take over!' How long does it take to get the driver back into the loop," in *Proc. Hum. Factors Ergonom. Soc. Annual Meeting*, vol. 57, no. 1. Los Angeles, CA, USA: SAGE, 2013, pp. 1938–1942.

[8] B. Wandtner, N. Schömig, and G. Schmidt, "Effects of non-driving related task modalities on takeover performance in highly automated driving," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 60, no. 6, pp. 870–881, Sep. 2018.

[9] B. Wandtner, G. Schmidt, N. Schoemig, and W. Kunde, "Non-driving related tasks in highly automated driving-effects of task modalities and cognitive workload on take-over performance," in *Proc. AmE Automot. Meets Electron., 9th GMM-Symp.*, 2018, pp. 1–6.

[10] N. Merat, A. H. Jamson, F. C. H. Lai, M. Daly, and O. M. J. Carsten, "Transition to manual: Driver behaviour when resuming control from a highly automated vehicle," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 27, pp. 274–282, Nov. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1369847814001284

[11] R. S. Jurecki and T. L. Stańczyk, "Driver reaction time to lateral entering pedestrian in a simulated crash traffic situation," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 27, pp. 22–36, Nov. 2014.

[12] R. S. Jurecki, T. L. Stańczyk, and M. J. Jaśkiewicz, "Driver's reaction time in a simulated, complex road incident," *Transport*, vol. 32, no. 1, pp. 44–54, 2014.

[13] M. Svetina, "The reaction times of drivers aged 20 to 80 during a divided attention driving," *Traffic Injury Prevention*, vol. 17, no. 8, pp. 810–814, Nov. 2016.

[14] M. Green, "'How long does it take to stop?' methodological analysis of driver perception-brake times," *Transp. Hum. Factors*, vol. 2, no. 3, pp. 195–216, 2000.

[15] J. Paxion, E. Galy, and C. Berthelon, "Mental workload and driving," *Frontiers Psychol.*, vol. 5, p. 1344, Dec. 2014.

[16] F. P. D. Silva, "Mental workload, task demand and driving performance: What relation?" *Proc. Social Behav. Sci.*, vol. 162, pp. 310–319, Dec. 2014.

[17] M. M. Moniri, D. Merkel, M. Feld, and C. Muller, "Incorporating the Driver's focus of attention into automotive applications in real traffic and in simulator setups," in *Proc. 12th Int. Conf. Intell. Environ. (IE)*, Sep. 2016, pp. 198–201.

[18] G. J. Andersen, R. Ni, Z. Bian, and J. Kang, "Limits of spatial attention in three-dimensional space and dual-task driving performance," *Accident Anal. Prevention*, vol. 43, no. 1, pp. 381–390, Jan. 2011.

[19] G. Iizuka, Y. Saito, and M. Yamada, "Measurement of a Driver's mental state using a 4K driving simulator," in *Proc. Int. Conf. Comput. Appl. Technol.*, Aug. 2015, pp. 66–71.

[20] G. Matthews, L. E. Reinerman-Jones, D. J. Barber, and J. Abich, "The psychometrics of mental workload: Multiple measures are sensitive but divergent," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 57, no. 1, pp. 125–143, Feb. 2015.

[21] K. A. Brookhuis and D. de Waard, "Monitoring drivers' mental workload in driving simulators using physiological measures," *Accident Anal. Prevention*, vol. 42, no. 3, pp. 898–903, May 2010.

[22] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs," *J. Cognit. Eng. Decis. Making*, vol. 2, no. 2, pp. 140–160, Jun. 2008, doi: 10.1518/155534308X284417.

[23] H. Makishita and K. Matsunaga, "Differences of drivers' reaction times according to age and mental workload," *Accident Anal. Prevention*, vol. 40, no. 2, pp. 567–575, Mar. 2008.

[24] R. Abbasi-Kesbi, H. Memarzadeh-Tehran, and M. J. Deen, "Technique to estimate human reaction time based on visual perception," *Healthcare Technol. Lett.*, vol. 4, no. 2, p. 73, 2017.

[25] S. Petermeijer, P. Bazilinskyy, K. Bengler, and J. de Winter, "Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop," *Appl. Ergonom.*, vol. 62, pp. 204–215, Jul. 2017.

[26] S. S. Borojeni, L. Chuang, W. Heuten, and S. Boll, "Assisting drivers with ambient take-over requests in highly automated driving," in *Proc. 8th Int. Conf. Automot. User Interface Interact. Veh. Appl. Automotive'UI*, 2016, pp. 237–244.

[27] A. Eriksson, S. M. Petermeijer, M. Zimmermann, J. C. F. de Winter, K. J. Bengler, and N. A. Stanton, "Rolling out the red (and green) carpet: Supporting driver decision making in automation-to-manual transitions," *IEEE Trans. Human-Mach. Syst.*, vol. 49, no. 1, pp. 20–31, Feb. 2019.

[28] A. Eriksson and N. A. Stanton, "Takeover time in highly automated vehicles: Noncritical transitions to and from manual control," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 4, pp. 689–705, Jun. 2017.

[29] J. Wan and C. Wu, "The effects of lead time of take-over request and nondriving tasks on taking-over control of automated vehicles," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 6, pp. 582–591, Dec. 2018.

[30] C. Gold, M. Körber, D. Lechner, and K. Bengler, "Taking over control from highly automated vehicles in complex traffic situations: The role of traffic density," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 58, no. 4, pp. 642–652, Jun. 2016.

[31] V. A. Banks, A. Eriksson, J. O'Donoghue, and N. A. Stanton, "Is partially automated driving a bad idea? Observations from an on-road study," *Appl. Ergonom.*, vol. 68, pp. 138–145, Apr. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0003687017302594

[32] 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.

[33] T. Suzaki *et al.*, "Steering behavior model of drivers on driving simulator through visual information," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, Dec. 2016, pp. 1–4.

[34] B. Hassan, J. Berssenbrugge, I. Al Qaisi, and J. Stocklein, "Reconfigurable driving simulator for testing and training of advanced driver assistance systems," in *Proc. IEEE Int. Symp. Assem. Manuf. (ISAM)*, Jul. 2013, pp. 337–339.

[35] S. Minhas *et al.*, "LEE: A photorealistic virtual environment for assessing driver-vehicle interactions in self-driving mode," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*. Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 894–900.

[36] A. Eriksson and N. A. Stanton, "Driving performance after self-regulated control transitions in highly automated vehicles," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 8, pp. 1233–1248, Dec. 2017.

[37] H. W. Lilliefors, "On the kolmogorov-smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 399–402, Jun. 1967.

[38] W. W. Daniel *et al.*, *Applied Nonparametric Statistics*. Boston, MA, USA: Houghton Mifflin, 1978.

[39] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric Statistical Methods*. Hoboken, NJ, USA: Wiley, 2013, vol. 751.

[40] B. Mok *et al.*, "Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 2458–2464.

[41] K. Zeeb, A. Buchner, and M. Schrauf, "Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving," *Accident Anal. Prevention*, vol. 92, pp. 230–239, Jul. 2016.

[42] R. D'Agostino, *Goodness-of-Fit-Techniques*. Abingdon, U.K.: Routledge, 2017.

[43] D. Lee and H. Yeo, "A study on the rear-end collision warning system by considering different perception-reaction time using multi-layer perceptron neural network," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2015, pp. 24–30.

Saad Minhas (Student Member, IEEE) received the master's degree (Hons.) in 3D Design for Virtual Environments in 2010. He is currently pursuing the Ph.D. degree with the School of Computer Science and Electronic Engineering, University of Essex. He also works as a Lead UX Designer with H3Space Ltd. His main research interests include advanced driver assistance systems and computer vision.

Aura Hernández-Sabaté received the degree in mathematics in 2002, and the Ph.D. degree *(cum laude)* in Computer Science in 2009. She is currently a member of the Computer Vision Center and an Associate Professor with the Computer Science Department, Universitat Autònoma de Barcelona. She has published more than 50 papers in refereed journals and international conferences. Her main research interests include performance evaluation and driver behavior analysis.

Shoaib Ehsan (Senior Member, IEEE) is currently a Research Scientist with extensive experience and major research accolades, including the BMVA Sullivan Doctoral Thesis Prize in 2013, and recognition as a "Researcher with Exceptional Promise" by the Royal Academy of Engineering in 2014. He has made significant contributions with his expertise in multiple fields, especially computer vision, embedded systems, and big data. He has received funding as a Researcher Co-Investigator for £12.69M in total, and a 'Named Researcher' for £5.5M in total.

Klaus D. McDonald-Maier (Senior Member, IEEE) is currently a Professor with the School of Computer Science and Electronic Engineering (CSEE), University of Essex, where he leads the Embedded and Intelligent Systems (EIS) Research Group. He is also a Visiting Professor with ESIGELEC, and the CTO of UltraSoC Technologies Ltd. He has authored or coauthored over 200 publications and 15 pending patents. His research has been funded by ESRC and EPRSC grants with substantial contributions from industry, as well as multiple EU grants, several other grants, and funded projects.